HOWARD UNIVERSITY

A Prosody-Enriched Dynamic Text Presentation Technique for Enhanced Reading of Electronic Text

A Thesis
Submitted to the Faculty of the
Graduate School

of

HOWARD UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

Master of Computer Science

Department of Systems and Computer Science

by

Eldon G. Marks

Washington, D.C.
December 2009

HOWARD UNIVERSITY
GRADUATE SCHOOL
DEPARTMENT OF SYSTEMS AND COMPUTER SCIENCE

THESIS COMMITTEE

____
John Trimble, Ph.D.
Chairperson

____
Kenneth Anderson, Ph.D
School of Education.

____
Harry N. Keeling, PhD.

____
Harry N. Keeling, PhD.
Thesis Advisor

Candidate: Eldon G. Marks

Date of Thesis Defense: 30th November, 2009.

ACKNOWLEDGEMENTS

I extend credit to my advisor and professors for providing me with guidance and insight to complete this project.

To my family and friends for their unwavering support and encouragement, thank you.

To my peers and to the persons who took time from their day to participate in the evaluation, many thanks.

Eldon Marks.

ABSTRACT

Increasing amounts of textual information are generated, exchanged and interfaced electronically in our progression towards a paperless world. The human-level method of processing this information must progress in tandem to address the reading efficiency bottlenecks experienced when trying to traverse this growing information space. Dynamic text presentation techniques such as Rapid Serial Visual Presentation (RSVP) offer the potential to induce high rates of reading through the temporal display of text, but at a significant cost of comprehension. This study explores the effectiveness of projecting *prosody* or the expressive properties of speech as visual effects on dynamic text as a means to sustain comprehension at higher-than-average rates of reading. The developed solution, a prosody-enriched dynamic text presentation technique, produces notable increases in comprehension in spite of low reading rates when compared to the traditional page format on electronic displays. Further, the majority of subjects that used the prosody-enriched technique indicated improved readability due to the visual projections of prosody on text.

Keywords: dynamic text presentation, Rapid Serial Visual Presentation, prosody, reading efficiency, reading rate, comprehension.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

CHAPTER 1. INTRODUCTION

Since the explosion of the social web, the natural, human-level methods of interfacing with, managing and learning from textual information have been greatly surpassed by the sheer volume and availability of electronic information. Early surveys done by Wright and Licorish (1988) state that electronic text will be increasingly accessed to the extent that it will rival or replace print-based text. Furthermore, due to the fast-paced nature of modern industry in which time is a critical resource, increasing demands are placed on the knowledge worker to rapidly read and comprehend electronically disseminated information at 'electronic speeds'.

In response to the preceding, dynamic text presentation, originally developed for reading on small displays, is now studied by researchers for its potential to allow readers to exceed aspects of reading encountered in traditional formats (Castelhano and Muter, 2001; Öquist, 2001; Beccue and Vila, 2004). The dynamic text presentation paradigm is one in which text is temporally displayed on screen as opposed to being spatially arranged on a traditional page. The temporal display of text generally removes the need for readers to make large sweeping eye movements as in traditional reading. It is this nature of temporally displaying text in a given dynamic text presentation technique which researchers believe creates the potential for reading enhancement.

1.1 The Problem

Research shows that the temporal display nature of dynamic text presentations also introduces added drawbacks that counteract its original potential for reading enhancement. For example, in Rapid Serial Visual Presentation (RSVP), a major dynamic text presentation technique, very high rates of reading may be imposed on readers, but at the cost of comprehension. The RSVP format displays words automatically at a fixed point on screen for predetermined intervals in quick succession (Forster, 1970; Aaronson and Scarborough, 1977; Potter, 1984; Rahman and Muter, 1999; Castelhano and Muter, 2001). Studies indicate that as the RSVP display rate increases beyond the average reading rate in traditional page formats, the comprehension levels in subjects drastically decline (Kang and Muter, 1989; Bernard, Chaparro and Russell, 2001). This limits RSVP to reading rates and comprehension scores similar to traditional page formats at best (Juola 1988; as quoted by Rahman and Muter, 1999).

Researchers attribute this limitation in RSVP to the invariance of RSVP word display times. While the temporal display of the text in RSVP is able to induce high rates of reading, the constant display rhythm disrupts the reader's natural projection of *prosody* on text, which is vital for comprehension (Fernández, 2007; Castelhano and Muter, 2001). Prosody is documented as the collection of properties evident in speech such as prosodic phrases, rhythm and focus which add communicative meanings to spoken language (Nooteboom, 1997; Frazier, Carlson and Clifton, 2006). Unlike speech, written text is virtually devoid of prosodic properties so readers must mentally project added prosody on text while they read (Fodor, 1998; Koriat et al. 2002; Ashby, 2006). The projection of prosody in silent reading is reported to influence the processing of the syntactic structure of a string of words to facilitate the integration of ideas (Fernández, 2007). In a study by Fernández (2007), the invariant rate of RSVP was shown to disrupt the projection of implicit prosody relating to the interpretation of ambiguous relative clauses in a sentence. The insertion of a pause after a relative clause mid-sentence was shown to alter this interpretation. Fernández's study demonstrated that interfering with the reader's perception of a prosodic phrase, through the odd placement of pauses in RSVP display times, affects semantics and, in turn, the comprehension of dynamically presented text.

1.2 A Brief Review of Existing Solutions

In relation to the preceding, Castelhano and Muter (2001) based their improvement of traditional RSVP on the concept of prosody. In experiment 2 of their study, they attempted to improve RSVP preference ratings by using an approach described as "*a hybrid of listening to speech and reading*". Their research produced a 'modified-RSVP' technique which emphasized the rhythm of word presentation. On relevant boundaries, pause durations were extended to a longer 460 ms ultimately varying the presentation in rhythm and allowing for emphasis to be placed on certain words. While the interspersed lengthier pauses may have reduced the overall reading rate, they significantly increased RSVP preference ratings over traditional RSVP as well as visibly increased the comprehension score over the traditional page format.

Further, in a previous study, a novel dynamic text presentation technique was produced which utilized prosodic phrasing. Aimed at enhancing the reading of electronic documents, the developed interface called the "Infojector" (Marks, 2008), dynamically presented chunks of text in a continuous, horizontal fashion while being spatially associated with the document that is read (see Figure 1.1).

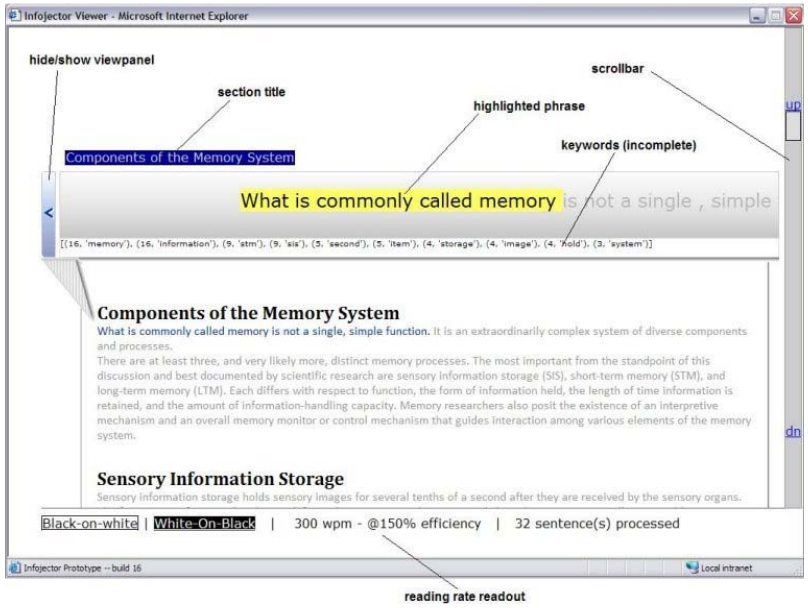

Figure 1.1 – Elements of the Infojector utility

In the dynamic text display, each sentence was chunked into phrases according to the natural language toolkit's (Bird, Klein and Loper, 2001) chunking algorithm to closely emulate the natural way readers chunk sentences when read or spoken. This chunking was representative of prosodic phrasing evident in speech. This prosodic phrasing method was chosen to make the transition from one phrase to the next more natural to the reader. To read, users were required to cycle through these highlighted phrases similar to that of a slideshow. In the evaluation of this technique, significant increases in reading rates were observed with a visible increase in comprehension levels compared to the traditional page format. Average reading performances of 178 wpm with 70.2 percent comprehension in the traditional page format rose to 232 wpm with 75.2 percent comprehension with the Infojector technique.

1.3 The Motivation for this Research

Results with the Infojector, as well as the aforementioned study by Castelhano and Muter (2001, experiment 2), revealed that the projection of prosody provides a potential to support comprehension in dynamic text presentations, possibly at higher-than-average rates of reading. The inclusion of rhythm in modified RSVP by Castelhano and Muter and the utilization of prosodic phrasing in the Infojector both demonstrated positive effects on comprehension of dynamically presented text. In either case, there was only a single applied prosodic property. Notably, in spoken language there exist other prosodic properties such as intonation, stress and focus (Nakatani and Schaffer, 1978; Castelhano and Muter, 2001) which further support the semantic processing and the integration of ideas. The purpose of this research is to assess the impact of incorporating multiple speech-level prosodic representations in a dynamic text presentation towards solving the outlined problem.

1.4 The Solution

A utility named the Prosody Enriched Dynamic Text reader (PEDText) was developed based on a formulated prosody-enriched dynamic text presentation technique. Under this technique, the traditionally auditory speech-level prosodic properties of stress, focus, rhythm and prosodic phrases were abstracted to produce their visual representational counterparts. These representations were designed to be visual effects applied to dynamically presented text in order to convey added elements of prosody during the presentation. The main objective in the utility's design was to exploit the benefit in dynamic text to induce high rates of reading, while at the same time address its drawback by synthetically projecting speech-level prosodic representations for comprehension support. This '*prosody-enriched dynamic text'*, is believed to not only relax the reader's task of mentally projecting prosody but extend the expressive dimension of text itself. The expected result is the sustainment and possible improvement of comprehension at higher-than average rates of reading, which translates to enhanced reading efficiency.

1.5 Thesis Outline

This thesis documents the formulation, implementation and testing of the PEDText reader to date. It begins by presenting a review of relevant background literature in Chapter 2. In this chapter, an overview is given of the process of reading and the concept of prosody, followed by an examination of reading efficiency and dynamic text presentation. In Chapter 3, the proposed concept of a prosody-enriched dynamic text presentation technique is presented and discussed along with the aspects relating to its design. Chapter 4 documents the implementation of the formulated model discussed in the previous chapter in terms of its interface and architecture. Chapter 5 describes the method and design of the user evaluation exercise together with an analysis of the obtained data. This is followed by a discussion of the results from the evaluation in Chapter 6. Finally, Chapter 7 concludes with insights and recommendations for future work reached on the basis of user evaluations.

CHAPTER 2. BACKGROUND

In this chapter, the natural human process of reading is first outlined in order to introduce the concept of reading efficiency – the fundamental focus of this research. Afterwards, a discussion of the deciding factors of reading efficiency which relate to legibility and readability attributes of the reading experience is presented. This foundation of reading efficiency factors is used to examine the enhancement of 'screen reading' or reading of electronic text – the scope of this research. In the examination of 'screen reading' enhancements, modes of traditional and dynamic text presentations are surveyed to highlight the position of this study within the scheme of related efforts.

2.1 The Reading Process

The reading process has been extensively studied over the years and according to Reichle et al. (2000), researchers have presented various generalized and specialized models of how the reading process works. The majority of these models identify that reading is an act of recognizing patterns and mentally processing them in some fashion (Öquist, 2001). As defined by Strydom and Du Plessis (2001), the reading process comprises interdependent stages and thought patterns that must take place simultaneously. These interdependent stages can be generally divided into (1) reception, (2) decoding and (3) learning. Related research has identified specific processes of reading which can be categorized under the aforementioned stages (Shepherd and Unsworth-Mitchell, 1999). Figure 2.1 shows a composite description of reading based on generalized and specialized models of the reading process.

	Reception	• ***Attention*** – an act of will to devote concentration for receiving information • ***Recognition*** – using the knowledge of alphabetical symbols to discriminate composed words and phrases. • ***Assimilation*** – the physical process of perception and scanning
	Decoding	• ***Intra-integration*** – acquiring basic understanding derived from reading the material itself without prior knowledge of the subject matter. • ***Extra-integration*** – involving the analysis, criticism of content based on the reader's past experiences
	Learning	• ***Retention*** – the ability to memorize information • ***Recall*** – the ability to recover information • ***Communication*** – the act of applying and synthesizing the acquired information

Figure 2.1 - The Stages of Reading - (Shepherd and Unsworth-Mitchell, 1999; Strydom & Du Plessis, 2001).

According to Strydom and Du Plessis (2001), while the stage of reception is taking place, the information is streamed in fragments to the intelligence areas of the brain where it is classified through association, understood and memorized. Little is known of the details regarding the processing aspects of reading beyond this description presented by Strydom and Du Plessis. What currently exists are competing ideologies of how orthography and phonology determine the semantics of what is read (Seidenberg and McClelland, 1989; Coltheart, Curtis, Atkins and Haller, 1993; as referenced by Coltheart, 2007). In response to these disputes around the processing aspects of reading, a viable approach to understanding how persons read may be to observe the more evident aspects of reading. The following section presents a survey the physical act of reading which is believed to provide some insight into how readers process textual information.

2.2 The Physical Act of Reading

A shallow observation of the physical act of reading reveals that a reader visually fixates on words within a sentence among sentences, scans from left to right to recognize the line of text which is mentally chunked for processing. According to research by Paulson and Goodman (2000; as quoted by Öquist, 2001), the eye makes brief jumps called *saccades* along a line of text. During these saccades, peripheral vision is used to locate the next portion of text upon which the eye briefly rests to create a *fixation.* On reaching the end of a line, the eye makes a *return sweep* (Huey, 1908; as quoted by Öquist, 2001) to the beginning of the next line and repeats the sequence of saccades, fixations and return sweeps. Research indicates that it is only during fixations that a reader cognitively processes information (Dodge 1900; as quoted by Öquist, 2001).

Existing models of the reading process generally assume that cognitive processing and the durations of fixations are directly related (Reichle et al., 2000; as quoted by Öquist, 2001). Notably, readers vary the lengths of the saccades and the durations of the fixations depending on the content being read (Landholt, 1891; as quoted by Öquist, 2001). Just and Carpenter (1980; as quoted by Öquist, 2001), discovered that content words, infrequent words and thematically important words attracted the longest fixation times and pauses. In addition, other researchers found that relatively longer pauses were made at sentence boundaries (Stine, 1990; as quoted by Rahman and Muter, 1999) and hold that post-processing of the sentence, involving the integration of ideas, takes place at these boundaries (Just and Carpenter, 1980; as quoted by Castelhano and Muter, 2001). Further, additional pauses were observed within sentences at major and minor clause boundaries as well (Stine, 1990; as quoted by Castelhano and Muter, 2001). This evident relationship between content and the physical reading act offers an insight into the processing aspects of reading. What follows is a discussion of *prosody* – a significant cognitive aspect of reading which is revealed through the variations of saccades, fixations and pauses of the physical reading process.

2.3 Prosody and Reading

The variable pauses and fixation durations of the physical reading process have been linked to linguistics and more specifically to prosody in linguistics (Nakatani and Schaffer, 1978; Wilkenfeld, 1980; Ashby, 2006; Fernández, 2007). The term prosody in modern linguistics refers to properties of speech which add communicative meanings to spoken language (Nooteboom, 1997; Frazier, Carlson and Clifton, 2006; Roach, 2009). Examples of these properties are (1) the *intonation* or controlled modulation of vocal pitch, (2) *lexical stress* or the elongation and contraction of syllable durations, (3) *focus* or the intentional fluctuations of loudness as well as (4) the overall *rhythmical* nature of the utterance (Nooteboom, 1997; Roach, 2009). In addition, prosodic properties also manifest in the grouping of words by the speaker so that the utterance is an audible sequence of *prosodic phrases* (Frazier, Carlson and Clifton, 2006; Roach, 2009). During fluent speech, prosody serves to extend the dimension of expression not encoded in grammar. For example, prosodic properties may be used to reflect the grammatical structure of what is said as well as the attitude or emotion of the speaker. In this regard, prosody is viewed as central to linguistics and language comprehension, in which the absence of prosody or inappropriate prosody produces an unpleasant or incomprehensible listening experience (Frazier, Carlson and Clifton, 2006; Roach, 2009).

Apart from spoken language, prosody is evident in reading as well (Selkirk, 2005; Ashby, 2006). Prosody in reading is reported to influence the processing of the syntactic structure of a string of words similar to the way it supplements the expression of spoken language (Fernández, 2007). During audible reading or reading aloud, the aforementioned prosodic properties can be clearly observed as the reader 'translates' text into speech (Selkirk, 2005). In silent reading, however, the presence of prosody may not be as obvious (Brown, 1958; Huey, 1968; Rayner and Pollatsek, 1989; Ashby, 2006). According to Huey (1908; as quoted by Ashby, 2006), the processes which facilitate both spoken language and silent reading are closely related. He noted in this observation, that when reading silently, one can 'hear' the text inside one's head as if it were spoken – an experience he called *inner speech*. Based on the interrelationship between spoken language and silent reading, the variable pauses and fixation durations of the reading process are seen as the physical evidence of a mirroring between the patterns of external speech and inner speech, which suggests the presence of prosody in silent reading (Brown, 1958; Huey, 1968; Rayner and Pollatsek, 1989; as referenced by Ashby, 2006).

Notably, the relationship between eye movements and the presence of prosody in silent reading was demonstrated in experiments by Ashby (2006). In these experiments, the frequency of fixations and their durations were measured to ascertain whether readers activate prosodic properties of lexical stress during silent reading. Results indicated that both high and low frequency words with two elongated syllables were fixated more frequently and were read more slowly than words with one elongated syllable. According to Ashby, these fixation measures indicate that readers impose at least the prosodic property of *lexical stress* during silent reading.

Both silent and audible reading require the reader to impose or generate prosody based on the interpretation of the words and sentences they read (Fodor, 1998; Koriat et al. 2002; Ashby, 2006). This requirement can be confirmed in the fact that text is devoid of explicit prosodic information (Fodor, 1998; Ashby, 2006). In this regard, written text is considered to have impoverished prosodic cues when compared to spoken language (Rayner and Pollatsek, 1989; Fodor, 1998; Frost, 1998; as referenced by Ashby, 2006). Nevertheless, audible readers are able to generate *intonation*, *phrasal rhythmic patterning* and *prosodic phrasing* from written sentences (Selkirk, 2005). Contrary to reading aloud, the degree of prosody in silent reading is subject to conflicting theories. Since written text does not explicitly encode prosodic information, Fodor (1998; as quoted by Ashby, 2006) claims that readers exploit linkages between spoken language and reading processes during silent reading. He implies that prosody in silent reading is not unlike that of spoken language. When readers *"mentally project"* prosody on printed text, syntactic ambiguity is resolved in a similar fashion to prosody in speech. Alternatively, Frost (1998; as referenced by Ashby, 2006) suggests a simplified manifestation of prosody in silent reading in which prosody is merely present to indicate the *"phonological gist"* of the word form. Regardless, researchers generally agree that prosody is central to linguistic processing, whether in spoken language, reading aloud or silently (Fodor, 1998; Frost, 1998; Frazier, Carlson and Clifton, 2006; Ashby, 2006). In light of this, prosody in silent reading can be viewed as integral to reading comprehension and ultimately reading efficiency in a cognitive sense. Having established this relationship, the following section continues by further discussing the concept of reading efficiency – the fundamental focus of this research.

2.4 Reading Efficiency

Throughout related researches, *reading efficiency* is viewed as a composite measure of reading rate and comprehension that indicates reading performance (Jackson and McClelland, 1979; Rahman and Muter, 1999; Castelhano and Muter, 2001). More specifically, reading efficiency is calculated as reading rate weighted by comprehension level (Cousin and Vinckenbosch, 2000). Reading rate is usually calculated in terms of the number of words the reader is able to perceive and scan per minute (wpm), whereas comprehension is measured by the percentage of correctly answered questions on the subject matter. The average recorded reading rate for English text on paper is approximately 250 wpm with a corresponding comprehension level of 60 percent (Cousin and Vinckenbosch, 2000). On screen, reading speed was recorded to be 20 – 30 percent slower with unchanged comprehension levels (Muter et al., 1982; Kang and Muter, 1984; Cousin and Vinckenbosch, 2000). While the majority of readers rarely improve beyond the aforementioned recorded threshold, skilled readers who are trained in methods such as speed reading are able to achieve efficiency in excess of 1000 wpm with near 85 percent comprehension on paper (Cousin and Vinckenbosch, 2001). Generally, the methods that allow skilled readers to achieve such high reading efficiency aim to condition the natural reading process for faster operation as well as compensate for some of the limiting factors encountered in the common reading experience (Buzan, 2000).

2.4.1 Reading Efficiency Factors

To a great extent, reading efficiency is impacted by a variety of factors which relate to the legibility and readability of textual information (Mills and Weldon, 1987; as quoted by Öquist, 2001). Readability related factors encompass the ease at which text can be comprehended by the reader (Mills and Weldon, 1987:331; as quoted by Öquist, 2001). For example, some readers are experienced and therefore are more fluent on a range of material as opposed to inexperienced readers. Likewise, readability can be affected by the content being read; whether it is properly written and comprehensive or flawed and unreadable. Legibility related factors, on the other hand, deal with the clarity of text as it affects recognition in reading. In this case, any visual aspects of textual information such as font size, font face or color, have the potential to affect word recognition through legibility (Muter, 1996). In addition, the way in which text is rendered or the *text presentation format* is yet another factor which affects legibility. For instance, reading from a flickering electronic display at a high resolution may be more difficult for recognition than reading from paper. Notably, the text presentation format has been shown to not only influence legibility but readability as well (Kang and Muter, 1989; Muter, 1996; Öquist, 2001; Castelhano and Muter, 2001; Fernández, 2007). For instance, as opposed to presenting text in a static, spatial fashion as in a traditional page, some text presentation formats dynamically display text in a temporal fashion. Some of these dynamically displayed text presentation formats have been reported to disrupt *prosody* during silent reading which negatively affects comprehension, and consequently the readability of text (Fernández, 2007; Castelhano and Muter, 2001). In brief, it seems unreasonable to consider all of the possible factors which either relate to readability or legibility; let alone to find an optimal combination of these factors to ensure good reading efficiency. The factor of *text presentation format*, however, presents potential to reliably influence reading efficiency since it relates to both readability and legibility aspects of text. The sections that follow continue by surveying the text presentation formats and relevant implementations under the context of reading efficiency.

2.5 Text Presentation Formats

The printed page of text is the *traditional* text presentation format which has been in existence for centuries. In this format, text is static and spatially presented often as a collection of several pages which are physically ‘flipped’ through in reading. With the advent of the first generation cathode ray tube (CRT) screens and computer technology, *electronic* text presentation formats began to emerge. These were initially screen-adapted renderings of print in the same static, spatial layout as a page of text with the exception that pages were navigated by virtually ‘scrolling’ through them as opposed to physically ‘flipping’ through them. With the introduction of small mobile devices, viewing screen-adapted renderings of print on the device’s small screen became inadequate due to the limited amount of textual information presented at a time. This restriction made it difficult and frustrating for users to read efficiently from mobile devices (Öquist, 2001). In response to this, dynamic text presentation formats emerged with the aim of improving user interactions with text on small displays. In dynamic text presentation, static electronic text is augmented by computer generated manipulations on its appearance, arrangement and placement within a limited display area. Researchers, subsequently discovered that the nature of dynamic text presentation formats allowed readers to match or even exceed aspects of reading encountered in traditional text presentation formats on both small and large screens (Castelhano and Muter, 2001; Öquist, 2001; Beccue and Vila, 2004). In addition, some researchers suggest that static text presentation formats are an outmoded technology compared to dynamic text presentation formats (Rayner and Pollastek, 1994). An overview of dynamic text presentation formats follows next along with an examination of related efforts in positively affecting reading efficiency in dynamic text presentations.

2.6 Reading Efficiency in Dynamic Text Presentations

Many formats for reading have been proposed under dynamic text presentation. Examples are the ‘moving window’ display (Just et al., 1982), ‘sentence-by-sentence’ presentation (Moore and Zarbrucky, 1995; Rahman and Muter, 1999), ‘Times Square’ (Chen et al., 1988; Kang and Muter, 1989; Juola et al., 1995; öquist, 2001) and Rapid Serial Visual Presentation or RSVP (Forster, 1970; Aaronson and Scarborough, 1977; Potter, 1984; Rahman and Muter, 1999; Castelhano and Muter, 2001). Of the various dynamic text presentation formats, RSVP exhibits the highest potential in optimizing reading efficiency (Juola et al., 1982; Masson, 1983; Potter, 1984; Juola et al., 1995; Muter, 1996; Rahman and Muter, 1999; Sicheritz, 2000; Goldstein et al., 2001; Castelhano and Muter, 2001).

RSVP was originally used by Forster (1970) as a tool to study the comprehension and processing of written language. Afterwards, RSVP was applied as a presentation technique for computer screens and became subject to many researches related to screen-based reading of text (Juola et al., 1982; Masson, 1983; Chen, 1986; Rahman and Muter, 1999; Castelhano and Muter, 2001; Öquist, 2001). While several variations of RSVP have been proposed, the common classification involves successively displaying words at a fixed point on screen at predetermined intervals (Forster, 1970; Aaronson and Scarborough, 1977; Potter, 1984; Rahman and Muter, 1999; Castelhano and Muter, 2001). The display technique of RSVP is believed to reduce cognitive load by eliminating eye movements in reading (Potter, 1984; as quoted by Castelhano and Muter, 2001) which many researchers view as a potential superiority to static displays (Juola et al., 1982; Masson, 1983; Chen, 1986; Castelhano and Muter, 2001). RSVP is also able to impose a high reading rate in readers (Rahman and Muter, 1999; Pelli and Tillman, 2007) due to interface designs that automatically sequence the display of words at a relatively high display rate. Comprehension, on the other hand, is yet to be enhanced or sustained by RSVP. Research shows RSVP to be limited to reading rates and comprehension scores similar to normal page formats at best (Juola 1988; as quoted by Rahman and Muter, 1999). When RSVP reading rates are increased beyond the average, comprehension deteriorated in the reader. Kang and Muter (1989) demonstrate that when RSVP presentation rates were increased from 100 words per minute to 300 words per minute, comprehension decreased from 70 percent to 52 percent. Further, Bernard, Chaparro and Russell (2001), conclude through their evaluation that comprehension also significantly decreased as RSVP display speeds increased from 250 wpm to 650 wpm. Both of these studies evaluated word-by-word RSVP with constant word display times.

The invariance of RSVP word display times is a widely established reason for low comprehension at higher than average RSVP reading rates (Rahman and Muter, 1999; Öquist, 2001; Castelhano and Muter, 2001). In normal page reading, researchers have observed that readers vary their fixation times based on whether the given word relates to content or function. The unnatural and invariant word display times of RSVP have been reported to disrupt the prosodic structure during silent reading which in turn negatively affects comprehension. In a study by Fernández (2007), the invariant rate of RSVP was shown to disrupt the projection of implicit prosody relating to the interpretation of ambiguous relative clauses in a sentence. The insertion of a pause after a relative clause mid-sentence was shown to alter the interpretation. This demonstrated that influencing prosody in reading, through variable pauses in RSVP display times, does indeed affect semantics and, in turn, the comprehension of dynamically presented text. In relation to the findings put forth by Fernández, researchers have aligned their focus towards developing dynamic text presentations that sequence the presentation of text based on features of the content being presented. Generally, these efforts have produced modified RSVP formats with variable display times determined by word length, sentence structure and punctuation, for example. What follows is a review of one such RSVP format which has been influential to this thesis.

2.6.1 Optimizing Reading of Dynamically Presented Text – An Evaluation

Castelhano and Muter (2001, experiment 2) based their improvement of traditional RSVP on the concept of prosody. They attempted to improve RSVP preference ratings by using an approach described as "*a hybrid of listening to speech and reading*". In this, they cited that speech is considered to have richer prosodic cues than written text and sought to better support *phrasal rhythmic patterning* in the dynamic presentation of text. Their research produced a 'modified-RSVP' technique which incorporated variable pause durations based on select content features. Firstly, they identified that function words were briefly fixated through their shape as opposed to the serial interpretation of each letter since they are frequently encountered. From this, they chose to reduce the display times of 11 high frequency words in certain conditions. Secondly, they considered the significance of pauses in the integration of information and chose to incorporate longer display times for words with punctuation. Thirdly, they focused on better supporting the syntactic structure of sentences through emphasizing certain words which establish the relationship between propositions of a sentence. They described that words such as 'because' or 'although' are types of words which may represent segmentation devices (that indicate syntactic structure), integration devices (that integrate current and previous information) or inference devices (that establish relationships between events in different clauses). By allotting longer display times for these words, they claimed that RSVP readers would better integrate information at clause boundaries. Lastly, they held that the lack of presentation controls in previous RSVP implementations frustrated users. Their modified RSVP allowed readers to pause the display with the option of regressing to the beginning of the sentence.

Evaluation results revealed that punctuation pauses implemented in their technique improved the user preference ratings for RSVP. However, the modified-RSVP technique achieved the lowest reading efficiency in comparison to plain RSVP and normal page formats. A closer look at the results indicates that while their modified RSVP achieved a marginally higher comprehension score than normal and plain RSVP formats, the relatively low reading rate impacted the resulting reading efficiency.

In spite of the results obtained, the research effort by Castelhano and Muter represents a promising move towards optimizing reading of dynamically presented text. Their concept of '*hybridizing spoken speech and reading*' is a viable approach in improving reading comprehension and ultimately enhancing reading efficiency in dynamic text presentations such as RSVP. While Castelano and Muter attempted to better support the mental projection of prosody on dynamically presented text during reading, their implementation focused only on the prosodic property of rhythm. In reading, readers not only project rhythm but, in theory, mirror the prosodic properties of speech which further support the semantic processing and integration of ideas. It is asserted that the fully applied concept of 'hybridizing spoken speech and reading' is key to sustaining comprehension at high rates of reading in dynamic text presentations. This hypothesis is the foundation of the designed prosody-enriched text presentation technique.

CHAPTER 3. DESIGN
PROSODY-ENRICHED DYNAMIC TEXT PRESENTATION

Creating a hybridization of spoken speech and reading required exploring two major research areas, namely prosody and text presentation. These areas then had to be combined in a way which renders prosodic properties of dynamically presented text in a visual form that enhances reading efficiency. This chapter describes the work done to design the proposed prosody-enriched dynamic text presentation format for enhanced reading efficiency.

It was reviewed that enhancing reading efficiency involves supporting good comprehension at higher-than-average rates of reading. Comprehension, in turn, is dependent on prosodic properties which add to the communicative depth of language, whether spoken or written. Unlike speech, written text is virtually devoid of prosodic properties so readers must mentally project added prosody on text for comprehension while they read. This mental projection of prosody is physically manifested as the pauses and fixations of the eye that come naturally when reading static text. The fluid motion of dynamic text presentations, while able to induce high rates of reading, interferes with the reader's natural projection of prosody on text. A resulting effect with dynamic text presentations is higher-than-average rates of reading with impaired comprehension which leads to a sub-par reading efficiency. Given the preceding, it is observed that the benefit of dynamic text presentations under the context of reading efficiency is also its drawback.

The main idea behind the proposed *prosody-enriched dynamic text presentation technique* is to exploit the benefit in dynamic text to induce high rates of reading while, at the same time, address its drawback by embedding prosodic representations inherently absent in text for comprehension support. As previously stated, the concept of this research is aligned with hybridizing speech and reading, as such, *speech-level* prosodic representations are designed to be synthetically projected on dynamically presented text. This synthetic projection of prosody not only aims to relax the reader's task of mentally projecting his own prosody, but to also extend the expressive dimension of text itself. The expected effect is the sustainment and possible enhancement of comprehension at higher-than-average rates of reading, resulting in enhanced reading efficiency.

Achieving enhanced reading efficiency through prosody-enriched dynamic text presentation not only involves projecting visual representations of speech-level prosody but the identification of content-based indicators of prosody (to sequence the representations), the choice of text display format (based on text display type, font specifics and text polarity) as well as the integration of an effective means of user interaction. These design aspects are examined next, which together describe the formulation of the proposed prosody-enriched dynamic text presentation technique.

3.1 Design Aspect I: Embedding Speech-level Prosody in Dynamically Presented Text

Speech-level prosodic properties are auditory in nature. For instance, intonation, lexical stress and focus are all manifested as audible frequencies during speech. Dynamically presented text, on the other hand, is entirely visual in nature. Accomplishing the synthetic projection of prosody on dynamic text given these disparate modes of information prompted a representational shift of speech-level prosodic properties from auditory to visual. This section examines the speech level-prosodic properties of (1) *intonation*, (2) *lexical stress*, (3) *focus*, (4) the overall *rhythmical* nature of the utterance as well as (5) *prosodic phrasing* with the aim of modeling their visual representations for improved comprehension in dynamic text.

3.1.1 Intonation

Intonation is described as the psychological perception of the variation of pitch during an utterance. It is also considered to be the *tune* of an utterance or the fundamental frequency (F0) produced by voiced speech. According to Rosenberger (1998), a high degree of precision is evidenced during intonation. Such precision is employed in conveying "*very fine shades*" of meaning that indicate the emotional state of the speaker which in turn substantially affects the interpretation of an utterance. The examples of intonational contour graphs shown in Figures 3.1 and 3.2 illustrate this point.

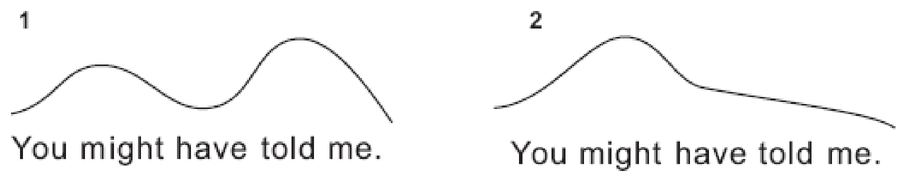

Figure 3.1 - Based on the intonation of "You might have told me", the speaker can imply indignation [1] or doubt [2] (Bollinger 1989; as referenced by Rosenberger, 1998).

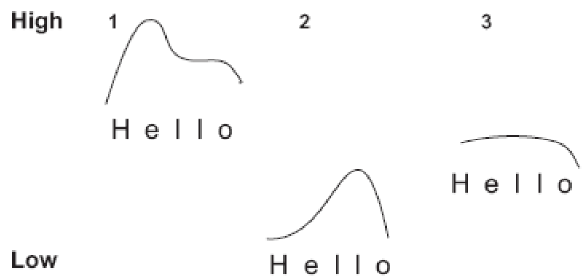

Figure 3.2 - Rosenberger shows that the common greeting "Hello" can convey mood based on the intonation used. The interpretations are as follows: [1] cheery, [2] sexual attraction; [3] indifference.

Apart from indicating emotion, intonation does little to influence the syntactic or grammatical nature of language. For instance, breaks in intonation by the speaker, whether intentional or due to speech errors may reveal a speaker's emotional state, but exhibit no completeness in a linguistic sense (Rosenberger, 1998).

Given that the significance of intonation is to convey the emotional connotations of utterances and does little for linguistic structure, its synthetic application to dynamic text in this study is considered to be unviable. It is observed that emotional tones conveyed by intonation in speech, are expressed differently in written text. For instance, written expressions such as rhetoric are commonly used in writing to convey emotional tones to the reader. The evident fact is that arbitrary text is incapable of indicating emotional undertones in any other way. As such, for a more natural integration with written text, focus was placed on more linguistically-oriented prosodic properties of speech.

3.1.2 Lexical Stress

In addition to intonation in any utterance, a speaker may stress on certain syllables with a greater emphasis than others. This is the prosodic property of *syllabic stress*, *lexical stress* or simply *accent*. Like intonation, lexical stress is perceived through distinct changes in the fundamental frequency (F0) contour, however unlike intonation it is considered to be linguistically-oriented. During lexical stress, elements of the fundamental frequency are associated with elements of text to convey a prominence relation in text (Ladd, 1980; as quoted by Rosenberger, 1998).

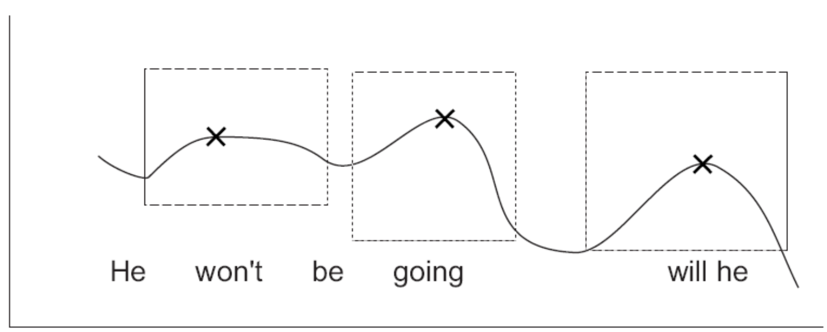

Figure 3.3 - An example of relative prominence created by lexical stress or accent (Rosenberger, 1998).

According to Rosenberger, the prominence effect gained from lexical stress can be abstracted through the *contrasts* created by the differences in the frequency contour.

This prominence effect can also be used to highlight a single word to establish contrast or comparison, or to draw focus to an entire phrase. In the latter sense, lexical stress and focus are one in the same.

3.1.3 Focus

The prosodic property of focus is observed through the changes in loudness or pitch of the utterance. Speakers often use high and low ranges of pitch for different semantic effects. For instance, high pitches are often associated with smallness and defenselessness whereas low pitches are used to convey dominance and power (Ohala, 1983; as quotes by Rosenberger, 1998). Rosenberger highlights that when speakers begin a new topic, their pitch is raised and when they approach the end of a phrase their pitch is lowered. Further, Rosenberger cites two differing representational models to describe the raising and lowering of pitch within a phrase: (1) the categorical or step-wise model (Bruce, 1977; Pierrehumbert, 1980; as reference by Rosenberger, 1998) and (2) the declination model (Collier and t'Hart, 1981; as referenced by Rosenberger, 1998). Figure 3.4 illustrates these models of focus on a phrase.

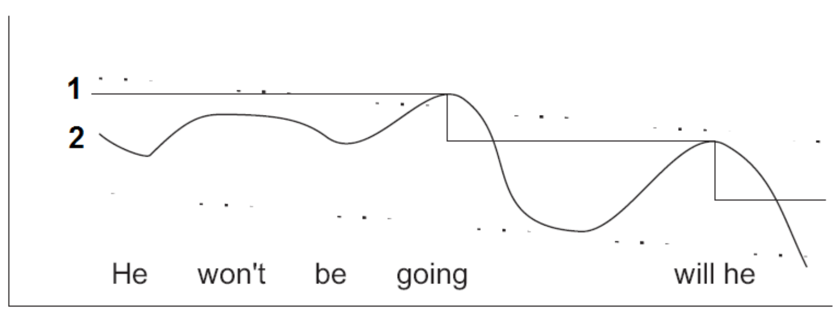

Figure 3.4 - The comparison of models of pitch range. [1] The step-wise depression of pitch during the course of an utterance, [2] the continuous linear declination of pitch range. As presented by Rosenberger (1998).

Rather than substantiate one model over the other, this study focuses on the general significance of pitch as it relates to elements of text in order to abstract the nature of this property for modeling. Through an observation of the contour graphs of lexical stress and focus (Figure 3.3 and Figure 3.4), as well as considering the nature of these prosodic properties, it is concluded that their representational effect is clearly *contrast*. The contrasting changes from one level of pitch to the next effectively convey prominence or emphasis in an utterance. In addition, the contrast created through changes in pitch is also manifested in lengthened durations of the utterance which in turn influences the prosodic property of rhythm.

3.1.4 Rhythm

Rhythm as a prosodic property is described as the result of other prosodic features. For instance, the syllabic prominence expressed through lexical stress elongates the duration of the utterance since the speaker is in essence *stressing* on the pronunciation of the syllable. Such an effect to produce rhythm is also encountered in the time it takes to change pitch as well as the expression of focus through loudness. Based on this observation, researchers assert that the rhythm in speech is determined by the duration patterns created by the periodic vocal stresses on content (Ladd, 1996; as referenced by Rosenberger, 1998).

The rhythmic patterning encountered in reading is also described to be based on content. When reading, it is the duration pattern created by the periodic visual saccades and fixations on content that manifest rhythm from text. As reviewed in the work by Castelhano and Muter, explicit text features such as punctuation, word types and propositions were used to indicate points within text where rhythm would vary or where lengthened fixations or pauses would be made. They stated that content words were fixated 83% of the time whereas function words were fixated only 35% of the time in normal page reading. This difference in fixation times relates to the prosodic properties of stress and focus where the duration of the utterance would be similarly lengthened when expressing emphasis or prominence.

3.1.5 Prosodic Phrasing

In speech, speakers usually group words into phrases delimited by pauses or intonational changes. Such phrases in speech are termed prosodic phrases and are demarcated by features of content collectively known as prosodic phrase boundaries (Kim and Oh, 1996). The prosodic phrasing of a sentence is reported to be important in understanding language. In cases where this prosodic structure is missing or inappropriate, speech becomes unnatural and negatively affects the meaning gathered from lengthy passages. An uttered sentence can contain a variety of prosodic phrase boundaries which may be influenced by syntactic structure, focus, the rate of speech, the need to breathe as well as the speaker himself (Kim and Oh, 1996). Prosodic boundaries that delimit phrases in spoken language are also present in written language; however, the written phrase is determined by a consistent grammar opposed to the relatively unpredictable utterance. Within the written sentence, readers have been reported to emphasize proposition words such as ‘because’ and ‘although’ for the purpose of supporting the integration of the phrases to which they belong (Townsend, 1983; as quoted by Castelhano and Muter, 2001). Such words in reading and speech have been described as either segmentation devices, integration devices or inference devices which generally denote the demarcation between two phrases or clauses and establish linkages among them. Proposition words within a sentence typically lie on written phrase boundaries and attract lengthier pauses. Consequently, written prosodic phrases are similarly indicated through rhythm due to the reported emphasis or degree of change in duration patterns on such words within a sentence.

3.1.6 Visual Representations of Speech-level Prosody

Based on the preceding assessment of speech-level prosodic properties, their visual representative counterparts were modeled with the exception of intonation. Since intonation is not linguistically-oriented and solely expresses emotional undertones which are already explicit in written text, its incorporation in dynamic text was viewed to be impractical and unnecessary. The linguistically-oriented nature of the remaining prosodic properties provided a common syntactic basis with written text that better supported the required representational shift from an auditory, speech-level property to a visual, dynamically displayed property.

3.1.6.1 The Representation of Stress and Focus

As discussed, the general nature of stress and focus is the indication of contrast. Through contrast, emphasis or prominence is created on words within a sentence. In speech, this contrast is conveyed through marked changes in pitch which in turn creates an effect of prominence or emphasis of the utterance as well. Given the common element of contrast in lexical stress and focus, a single visual representation was chosen that conveys their similar effect of *emphasis*.

The projected effect of contrast in the visual sense can serve as a practical indicator of emphasis. The application of contrast to dynamic text may be represented by a variety of means such as font-size, color, boldness. In addition consideration was made for the convention of enlarged font sizes or upper-case lettering for contrast that conveys emphasis in written text. For example, Ladd (1996; as quoted by Rosenberger, 1998), used the following joke with uppercase lettering in which a reporter and a bank robber have an exchange:

Reporter: "Why do you rob BANKS?"

Bank Robber: "Because THAT's where the real MONey is."

The use of uppercase lettering or a distinct change in the typeface is noted in similar expressions of shouting or emphasis. This convention works relatively well for static text, which is spatial in nature; however it is proposed that dynamic text must additionally support temporal contrast in the form of a lengthened pause to complement the emphasis of a dynamically displayed word. This proposed temporal contrast which essentially varies the display durations of words, can be viewed as the counterpart to the patterns of fixation durations in the physical act of reading as well as the lengthened durations of certain utterances in speech.

When considering the preceding, the devised visual representation of emphasis expresses the common element of contrast through combinations of color, font boldness, enlarged font sizes and a lengthened display time.

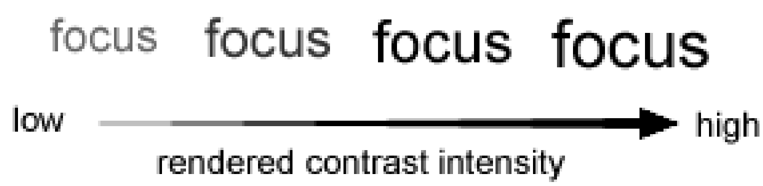

Figure 3.5 - The visual representation of contrast in text. The visual representation of contrast in text to portray emphasis may range from slight color contrast to pronounced color contrast and enlarged font.

3.1.6.2 The Representation of Rhythm

Rhythm in the design was based on the general duration patterns encountered in speech and reading. When dynamically presenting text, rhythm is projected in the varying specifications of word display times which aim to draw parallels to that of speech and reading. In retrospect, the prosodic element of rhythm required no representational shift considering that the element of timing, which expresses rhythm, is common to speech, reading and dynamic text presentations.

3.1.6.3 The Representation of Prosodic Phrases

Prosodic phrases in spoken speech are described as uttered groups of words separated by a noticeable pause. In a visual sense, a prosodic phrase can be viewed as a portion of a sentence truncated by a prosodic phrase boundary. The prosodic phrase boundary or the point of demarcation between phrases is similarly represented in the design by a noticeable pause given the temporal nature of the dynamic display.

3.2 Design Aspect II: Content-Based Indicators of Prosody

The devised representations of speech-level prosody *must* be mapped to the semantic nature of content in order to be effective in enhancing comprehension of dynamic text. The devised prosodic representations can be viewed as visual effects that rely on content-based indicators of prosody for their sequenced rendering throughout the dynamic text presentation.

In designing the required ‘mapping’ of content to visual prosodic representations, two attributes must be indicated for each mapping. Firstly, *the type of prosodic property* must be indicated for the current word within the scheme of content. In this regard, it is also considered that several properties may be applied in unison to a word such as rhythm and emphasis, for instance. Secondly, given that a prosodic property may be expressed in varying degrees, *the degree of intensity* must be indicated as well. Within the scope of this thesis, these attributes form a given *content-based prosodic indicator*.

It was identified that the required attributes of a content-based prosodic indicator can be reliably determined from the close inter-relationships that exist in prosodic properties. As previously discussed, the prosodic property of rhythm is the product of other prosodic properties. Based on this evidence, the indicators for the visual representations of emphasis and prosodic phrases may all be derived or ‘de-multiplexed’ from rhythmic patterns. More specifically, the degrees of change in duration patterns which give rise to rhythm offer enough information to indicate moments of contrast as well as the related intensity that not only denote emphasis but phrase boundaries as well. Based on previous discussions, the contrast expressed in stress and focus is in the form of a lengthened duration of the utterance based on the degree of emphasis applied; and the greater the intensity of emphasis, the lengthier the duration. To highlight this relationship the uttered durations of each word of an arbitrary sentence were recorded and graphically depicted, expressing the rhythmic pattern of that sentence (see Figure 3.6).

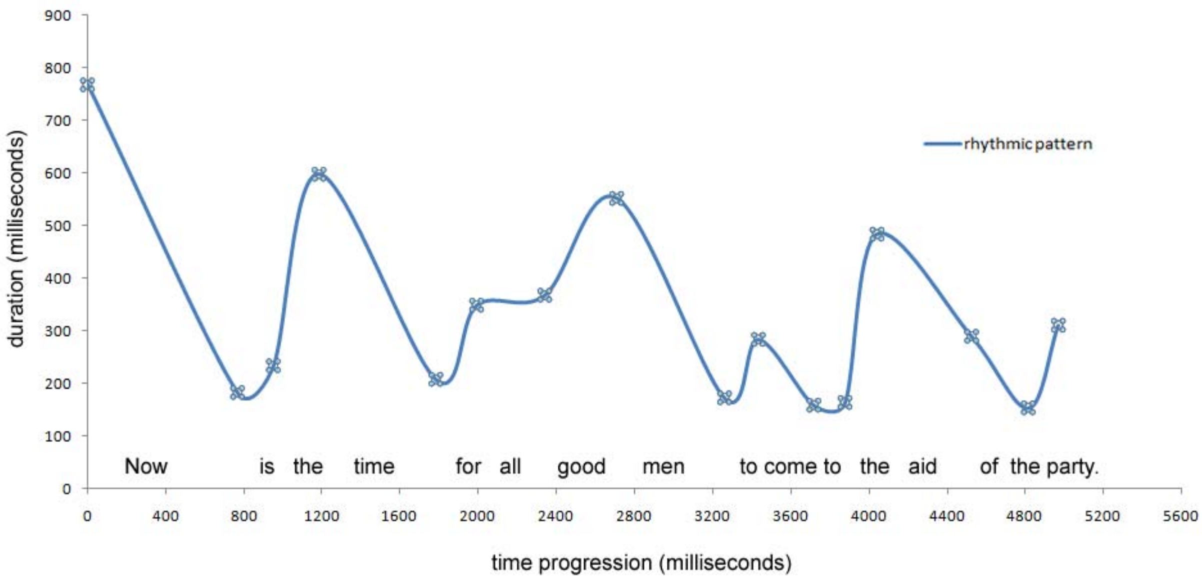

Figure 3.6 - The rhythmic pattern of an arbitrary sentence. The duration of each uttered word plotted against the timeline of the sentence highlights the relationship between the rhythmic patterns and the prosodic properties of emphasis and prosodic phrases.

From this rhythmic pattern, the presence of emphasis can be detected through contrasting changes in word duration. Figure 3.6 shows that the word '*Now*' has a significantly lengthened duration and is consequently contrasting to the neighboring words. The detection of this evident contrast is sufficient to determine the type of property to be emphasis since expressed emphasis creates durational contrast in speech and reading. Similarly, the degree of intensity of emphasis on this word is specified in the extent of durational contrast relative to other emphasized words.

Simple statistical analysis was applied to identify the contrasting words which should be emphasized based on a rhythmic pattern of text. Given that contrast is created through the relative durations of words in the rhythmic pattern, treating the patterns as a discrete statistical distribution and calculating the mean, provided a reference to relatively identify emphasized words. Since the mean represents the average duration of an uttered word in the sentence, only emphasized words will have a lengthier duration than that of the mean. A simple comparison of the recorded durations with the mean is sufficient for detecting emphasized words (see Figure 3.7).

Figure 3.7 - Identifying emphasized words in a rhythmic pattern. By treating the rhythmic patterns as a discrete statistical distribution, the calculated mean provided a reference point to relatively identify emphasized words. These words (shown in bold) possess a duration that exceeds that of the calculated mean.

After the emphasized words have been detected, their related intensity is then calculated based on the difference between the word duration and the calculated mean. The calculated mean is used in this case as a base reference to determine the relative degree of emphasis that must be applied to an emphasized word. Once the difference between the word duration and the mean is calculated, it is then normalized to a value that ranges between 0 and 1 to numerically represent the degree of emphasis to be applied. In keeping with the relative nature of contrast, the normalizing function uses the largest calculated difference between word duration and the mean as the denominator.

The normalizing function used is as follows:

$$\text{intensity} = d_n / \text{MAX}(d) \qquad \textbf{Equation 3.1}$$

where ***d*** *is the recorded differences between the duration of word n and the calculated mean*

Based on this normalized value, the visual representation of emphasis on an emphasized word may be rendered as slightly contrasting in color to highly contrasting in color, bold and uppercased according to four intensity ranges: +1 color contrast, +2 color contrast and +6 font size, +3 color contrast and +7 font size, +3 color contrast and +8 font size (see Figure 3.8).

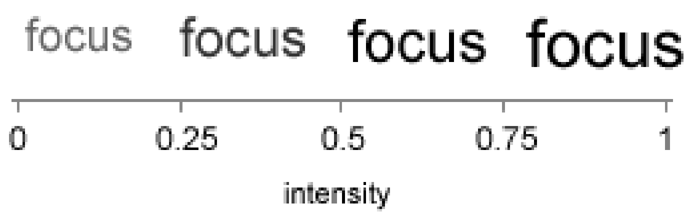

Figure 3.8 - The rendered intensity ranges. The visual rendering of emphasis is mapped to the intensity value through four intensity ranges: +1 color contrast, +2 color contrast and +6 font size, +3 color contrast and +7 font size, +3 color contrast and +8 font size.

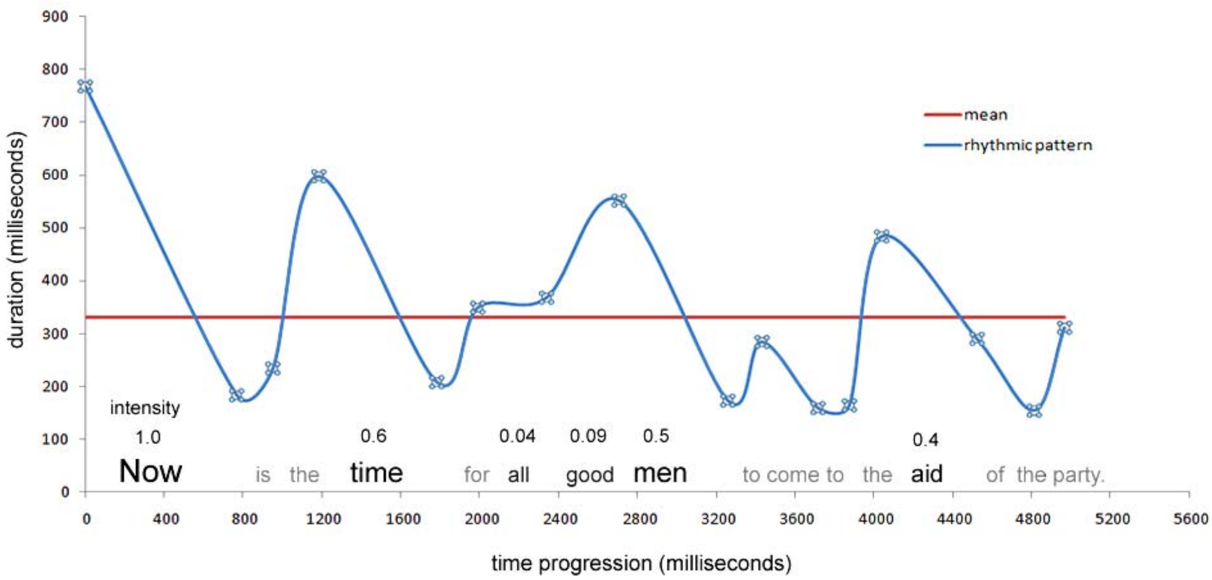

Figure 3.9 - The visual representation of emphasis. This figure shows the intensity ranges and related representations rendered on the sentence along with the calculated intensity values.

Similarly, the boundaries which separate prosodic phrases are expressed through a detectable pause or change in duration as well. More specifically, it can be observed that a distinguishable indication of a prosodic phrase is held in the relatively small temporal differences between uttered words that comprise the phrase. For example, in Figure 3.10, the prosodic phrase "*to come to the aid*" is made distinguishable since the temporal differences between the utterance of each word in the phrase is contrasted in duration by the longer pauses which delimit the phrase itself.

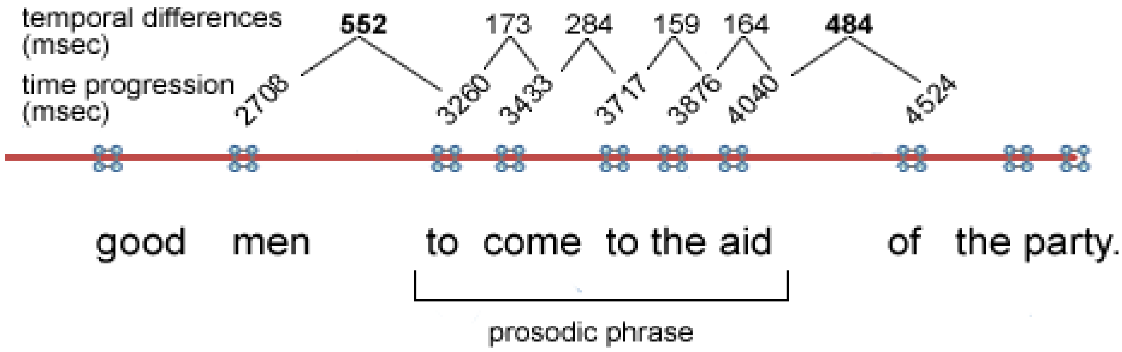

Figure 3.10 - The identification of a prosodic phrase from a rhythmic pattern. The prosodic phrase "to come to the aid" is made distinguishable in a rhythmic pattern since the temporal differences between the utterances of each word in the phrase are contrasted in duration by the longer pauses (in bold) which delimit the phrase.

Given the preceding examination, it is noted that composed rhythmic patterns can reliably indicate the type and intensity of the visually represented properties of emphasis and prosodic phrases. However, to achieve functional goals, these rhythmic patterns must be first composed based on the semantic nature of content. Devising a method that reliably accomplishes the composition of content-based rhythmic patterns for the projection of prosody in dynamic text is the design challenge.

Recent researches in RSVP-like dynamic text presentations vary word display times based on the features of content in order to remove the unnaturalness of invariant display times (Castelhano and Muter, 2001, Öquist, 2001). These methods range from simply increasing the word display time based on word length to examining the context of the word and determining the display time accordingly. Öquist (2001) referred to the former as content adaptation and to the latter as context adaptation. For the purpose of this research, a context adaptation method is required. Context adaptation seeks to emulate the large variations of fixation times and gaze durations on words when reading from paper. The goal of this emulation, according to Öquist (2001) is to mimic a reader's cognitive pace more adequately when processing text. Context adaptation considers that the required display time of a word should be proportional to the predicted fixation duration of that word in order to achieve a mimicking of the reading act. The method currently used for prediction of fixation durations in context adaptation generally involves the factoring of word frequencies, word types and punctuation. For example, Castelhano and Muter (2001) shortened the duration of common words to half the presentation time of other words while the duration for punctuation pauses was doubled. However, researchers express that there exists a larger variation in fixation durations in traditional reading. In some studies it has ranged between 100-500 ms (Rayner 1998; as quoted by Öquist 2001) whereas in others it has been found to vary as much as 50-1500 ms (Just and Carpenter 1980; as quoted by Öquist 2001). Considering the preceding, the fixed ratios of *common word duration* to *uncommon word duration* to *punctuation pauses* creates a restricted and mechanical reading effect as opposed to the intended 'natural' one. While such current context adaptation methods seek to synthesize variation, the range of achievable variation and accuracy is far from that of natural reading or speech which compromises the intended 'natural' effect expected from the text presentation.

For the intended goal of adequately projecting speech-level prosody on dynamic text, the composed rhythmic patterns must be as identical to that of an audible reader or speaker as possible. Since the focus of this research is not on developing a context adaptation method but on improving reading efficiency through projected prosody, the manual composition of rhythmic patterns for content is chosen for the purpose of evaluating the research concept.

3.3 Design Aspect III: The Dynamic Text Display Format

Of equal importance to projecting visual prosodic representations on text is the manner in which text itself is dynamically displayed during the presentation. The dynamic text display format is determined by those design factors of a dynamic text presentation that influence the legibility and readability of electronic text (So and Chan, 2009). These design factors, within the scope of this thesis, are the text display type, font specifics and display polarity. This section presents a survey of the aforementioned design factors of a dynamic text display in the formulation of an optimal display format given the project goals.

3.3.1 Text Display Type

Many text display types have been proposed including the 'moving window' display in which a reader views successive words in a text after pressing a button; 'Times Square' or 'leading' which involves the horizontal scrolling of text; 'sentence-by-sentence' presentation, in which individual sentences of a text are successively displayed; and RSVP, in which words are successively displayed on screen at predetermined intervals. The general focus of these aforementioned display types is presenting text for optimal readability within a limited display area.

Creating a more natural reading experience during the dynamic text presentation is beneficial to overall reading efficiency (Just et al., 1982; Stine, 1990; Castelhano and Muter, 2001, Fernández, 2007). Consequently, the design objective regarding text display type focuses on simulating the experience of fixations and saccades in the physical act of reading. For this design, an abstraction is made of a given fixation in the form of a momentary display of a word within a fixed focal region – much like word-by-word RSVP. This abstraction was extended when considering foveal vision in reading. Foveal vision refers to the central, focal region of vision that has 100 percent acuity. In the physical act of reading, the general, focal region of the reader has been reported to span eight or nine character spaces to the right of a fixation and about four to the left (Rayner and Pollatsek 1989; Robeck and Wallace 1990; Rayner and Serano 1994; Rayner 1998; as quoted by Oquist, 2001). It is widely established that only about five letters are seen with 100 percent acuity around the fixation point and the level of visual acuity gradually degrades outwards to the periphery of the reader's vision (Taylor, 1965; Hunziker, 2006) (see Figure 3.11).

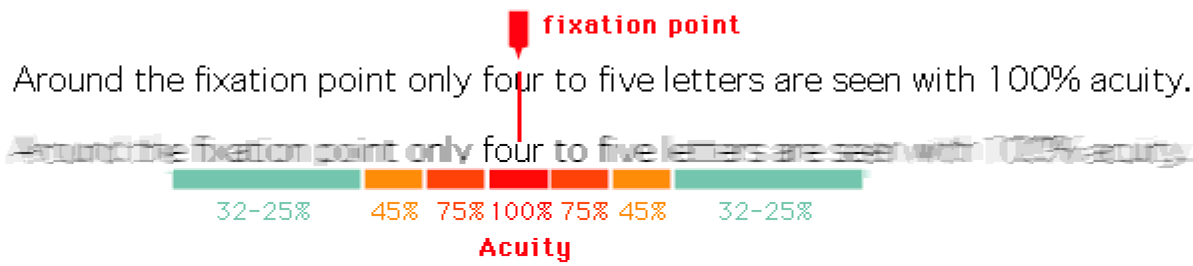

Figure 3.11 - The foveal vision in reading . During a momentary fixation in reading, a reader's foveal vision captures only four to five letters around the fixation point at 100% acuity. Taken from Im Auge Des Lesers, H.W Hunziker, 2006.

Notably, Hunziker, among other researchers (Chung, Legge and Cheung, 2003), considers peripheral vision to be an important form of input in reading. It is responsible for the pre-processing of surrounding text and facilitates the rapid calculations of muscle commands that reorient a reader's focus for a new fixation. This reorientation of a reader's focus occurs three to four times per second in the form of a *saccade*.

To abstract a saccade, a restricted amount of surrounding text is displayed in addition to the current word in focus to support peripheral pre-processing. The abstracted motion of a saccade within the design is enacted through the horizontal scrolling from the current word to the next word within the fixed focal region – somewhat similar to the 'Times Square' or 'leading' display type. The resulting display type design can thus essentially be described as a combination of the word-by-word RSVP and the word-shift 'Times Square' or 'leading' format.

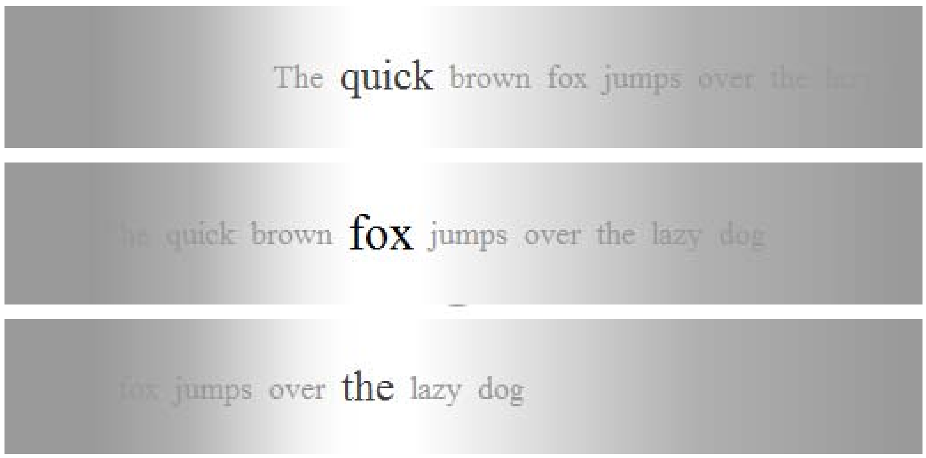

Figure 3.12 – The resulting display type. The figure depicts frames from the designed display type in motion. A single word is displayed in a fixed focal region while a restricted amount of surrounding text is also displayed to support peripheral pre-processing. The display scrolls horizontally from word to word, aligning the newly focused word in the fixed focal region.

3.3.2 Font Specifics

Apart from the display type which presents the textual information in a certain manner, the best font type and size that should be used when reading electronic text were also considered. Bernard, Liao and Mills (2001) investigated the legibility, reading time and general preference of select fonts at different size configurations. They evaluated Times New Roman, Georgia, Arial and Verdana font types at 14 and 12-point sizes on computer screens. Results indicated that the sans serif fonts of Arial and Verdana were preferred over the serif fonts of Times New Roman and Georgia. Further, the 14-point font size was more preferred than the 12-point font size and was found to be more legible. The larger size also increased reading speed for both fonts. The larger serif fonts were found to promote faster reading than the sans serif fonts; however, the sans serif fonts were more preferred. Overall, Bernard, Liao and Mills recommend larger font sizes as well as serif fonts when reading speed is a concern or sans serif fonts when user preference is a concern. For the design purpose, reading speed and legibility are of utmost importance. Therefore, Times New Roman at a large font size of 18-point was chosen as the default font specifics.

3.3.3 Display Polarity

Display polarity is another concern relating to text presentation on computer screens. Display polarity may either be positive (black letters on a white background) or negative (white letters on a black background). Studies show that on electronic displays, positive polarity was more comfortable to readers and produced faster searching times (Lee *et al.*, 2006). According to Taptagaporn and Saito (1990), reading from negative polarity apparently causes pupil dilation which then creates problems with focusing a viewed image on the retina. They concluded that positive polarity is "psychologically and physiologically" better than negative polarity on electronic displays. In keeping with the design requirements, a positive display polarity was chosen as the default polarity considering its aforementioned suitability regarding reading speed.

3.4 Design Aspect IV: Usability Requirements of Dynamic Text Presentations

One of the major shortcomings of well-known dynamic text presentation techniques such as RSVP is low user preference ratings. Studies show normal page, sentence-by-sentence and leading formats to be highly preferred over RSVP (Kang and Muter, 1989; Rahman and Muter, 1999; Castelhano and Muter, 2001). Some researchers have examined added features geared towards improving the usability of traditional RSVP such as completion meters (to indicate reading progress), bi-directional movement of text (for regressing) and the user-controlled presentation of text (Castelhano and Muter, 2001; Juola, 1988, Muter et al., 1988). In this section the usability requirements of dynamic text presentations are surveyed using the relatively disliked RSVP format as a basis. The intended design goal here is to incorporate an effective means of interaction for the proposed display type.

3.4.1 The Indication of Reading Progress

The spatial nature of traditional text allows a reader to judge what has been read and the amount of reading that remains. This spatial reference is obscured in RSVP which may result in the users feeling "lost" and uneasy when reading from RSVP displays (O'Hara and Sellen, 1997; as quoted by Castelhano and Muter, 2001). To mitigate this effect, some researchers have included completion indicators in RSVP displays that indicate the percentage of material currently read and what remains. In traditional page format, the spatial arrangement of text facilitates the use of incidental memory for locations within the page for added processing (Rothkopf 1971; as referenced by Rahman and Muter, 1999). Rahman and Muter (1999) state that completion indicators signify what portion of text has been read and may partially compensate for the absence of incidental location cues in small-window displays. They found that with the addition of completion indicators, user preference increased, however little evidence suggests that reading efficiency is affected in any way by this measure.

3.4.2 The Need to Regress in Reading

According to Rayner and Pollatsek (1994), a major usability flaw with RSVP is the fact that users are unable to go back and reread information as in normal silent reading. Through previous studies, they indicate that readers often make regressions in normal silent reading for clarification if they have misunderstood a segment of text; and the inability to regress in RSVP negatively impacts the reader's comprehension (Kennedy and Murray, 1984).

3.4.3 User-Controlled Presentation of Text

An additional usability flaw with RSVP reading is that it is subjectively demanding (Rayner and Pollatsek, 1994). As text is automatically presented at a predetermined display rate in RSVP, users are required to provide constant attention to text. Rayner and Pollatsek suggest that this creates fatigue in users and produces an unnatural reading experience when compared to the freedom of regressing and pausing in traditional reading.

As a proposed solution to this inadequacy of RSVP, researchers have explored methods of having the user control the presentation of text by pushing a button each time they want text to appear (Aaronson and Scarborough, 1976; Aaronson and Ferres, 1983; Just, Carpenter and Wooley, 1982; as quoted by Rayner and Pollatsek, 1994). However, studies highlight undesirable effects of this method. Firstly, since the reaction time to perform a button press with a finger is slower than that of moving the eye, the user controlled exposure time of each word is slowed to an average of at least 400msec (Rayner and Pollatsek, 1994). Aaronson and Scarborough (1976; as quoted by Rayner and Pollatsek, 1994) additionally observed that users make button presses at a relatively steady rate with little variability across different words. Further Rayner and Pollatsek proposed that since finger presses are less practiced than eye movements, the resulting effect is an unnatural experience which imposes an additional overhead on cognitive processes when using this user-controlled means of text presentation. They recommend that the desired mode of user-controlled presentation of text should be one in which "the reader can control the flow of words and regress in a natural manner so that reading could progress more rapidly than normal silent reading and yet not produce fatigue".

Considering the surveyed aspects of usability, designing an optimal means of user interaction given the project goals involves: (1) the inclusion of a completion indicator to combat reader disorientation and (2) incorporating a user-control method that not only permits regression, pausing and a variable display rate, but does so without the requirement of constant button presses.

What follows next is a description of the formulated prosody-enriched dynamic text presentation model based on the previously discussed design aspects of usability, text display format, content-based prosodic indicators and visual representations of prosody.

3.5 The Prosody-Enriched Dynamic Text Presentation Model

The four design aspects previously discussed in this chapter were composed into a model of the prosody-enriched dynamic text presentation technique. The major aspects of the designed model labeled I – IV in Figure 3.13 are summarized in the following:

II – The content based indication of prosody begins with the composition of rhythmic patterns on regular electronic text. For this composition a manual method is used to provide a natural rhythmic pattern based on the expressed durations of each word uttered by a reader. In the manual method, a human reads each word in a text aloud in a natural rhythm and the duration of each expressed word is recorded as “rhythmic patterned text”.

I - The marked-up text captioned as “rhythmic patterned text” is used to project the property of rhythm on text and is in turn used to derive other types of properties as well as their related intensity. The derivation of a property and its related intensity is accomplished through simple statistical analysis of the duration of each word in the rhythmic patterned text. For example, emphasized words are identified when their recorded duration exceeds the calculated average durations of all words. Their related intensity is then relatively calculated for each emphasized word by a normalized value called the intensity factor which is then used to apply the relevant intensity rendering from subtle color contrast to enlarged font size and high color contrast.

III – The detected property, its intensity as well as the “rhythmic patterned text” are all visually rendered and sequenced as prosody enriched text in the dynamic display format. The dynamic display format utilizes a hybrid of word-by-word RSVP and word-shift Times Square display type, a positive display polarity and a serif font with a 18-point size to render the prosody enriched text as prosody enriched dynamic text.

IV –The dynamic display format is manipulated by user controls that indicate the mode (paused or playing), direction (backwards or forwards) as well as the variable speed in both directions. The user controls are designed and incorporated to enhance user comfort. In addition, the model incorporates a completion indicator to indicate reading progress which has been shown to combat reader disorientation.

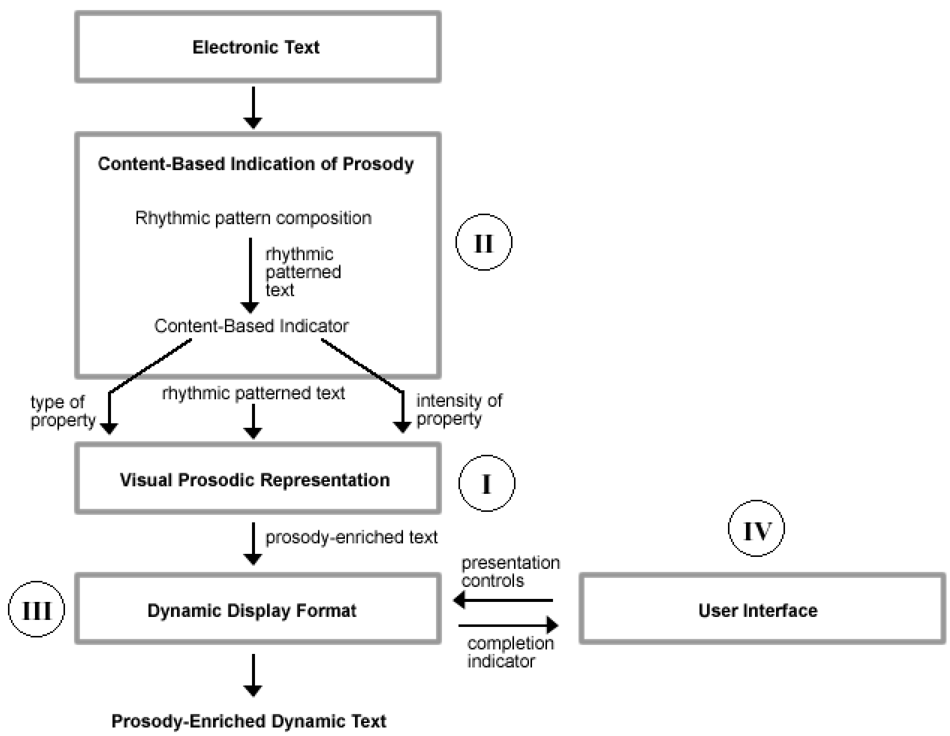

Figure 3.13 - The Prosody-Enriched Text Presentation Model. The model depicts the incorporation of all discussed design aspects: (I) visual representation of prosody, (II) content-based indicators of prosody, (III) the dynamic text display format and (IV) usability specifics.

CHAPTER 4. IMPLEMENTATION

PROSODY-ENRICHED DYNAMIC TEXT PRESENTATION

In order to evaluate the effectiveness of the proposed prosody-enriched dynamic text presentation technique, the designed model was implemented in the form of a software utility named the **P**rosody **E**nriched **D**ynamic **Text** (PEDText) reader. This chapter describes the implementation of the PEDText reader with respect to its interface and architecture.

4.1 Interface and Architecture

A combination of JavaScript, HTML and CSS, collectively known as Dynamic HTML or DHTML was used to implement the PEDText reader. The already available JavaScript Document Object Model as well as its scripting capabilities provided a convenient and powerful means to interface with text elements and dynamically manipulate their styling attributes for animation.

The developed architecture and related interfaces of the PEDText reader are made up of four major components based on the design model. These components are: (1) the composer engine, (2) the prosody renderer, (3) presentation controller and the (4) the display engine.

4.1.1 The Composer Engine

According to the model (see Figure 3.13), regular electronic text has to first be pre-processed to derive the rhythmic pattern that is used in the further detection of prosody within text. The general purpose of the composer engine is to derive and encode the rhythmic pattern based on supplied text which is representative of the individual word durations expressed by a human speaker. To ensure a highly natural derivation of rhythmic patterns from text, the implementation was adapted to rely on input from a human reader as opposed to operating automatically. The implemented composer engine was developed to capture from the reader the expressed durations of each word as well as to encode these durations in the text to produce "*rhythmic patterned text*".

Once regular text is inputted into the composer engine, it then tokenizes the text and generates a serial presentation of the text for the human reader (see Figure 4.1). In this serial presentation, the human reader is required to read the text out loud as he or she naturally would; however, after each word is read the human reader presses a button to cycle to the next word. With each button press, a timer component within the composer engine records the amount of time (in milliseconds) the reader has spent on the previous word and subsequently resets itself for the next word. The recorded amount of time as well as the associated word is referred to as a 'word token'. Each generated word token is represented as HTML markup text and appended in the correct sequence with existing word tokens (see Figures 4.1 and 4.2).

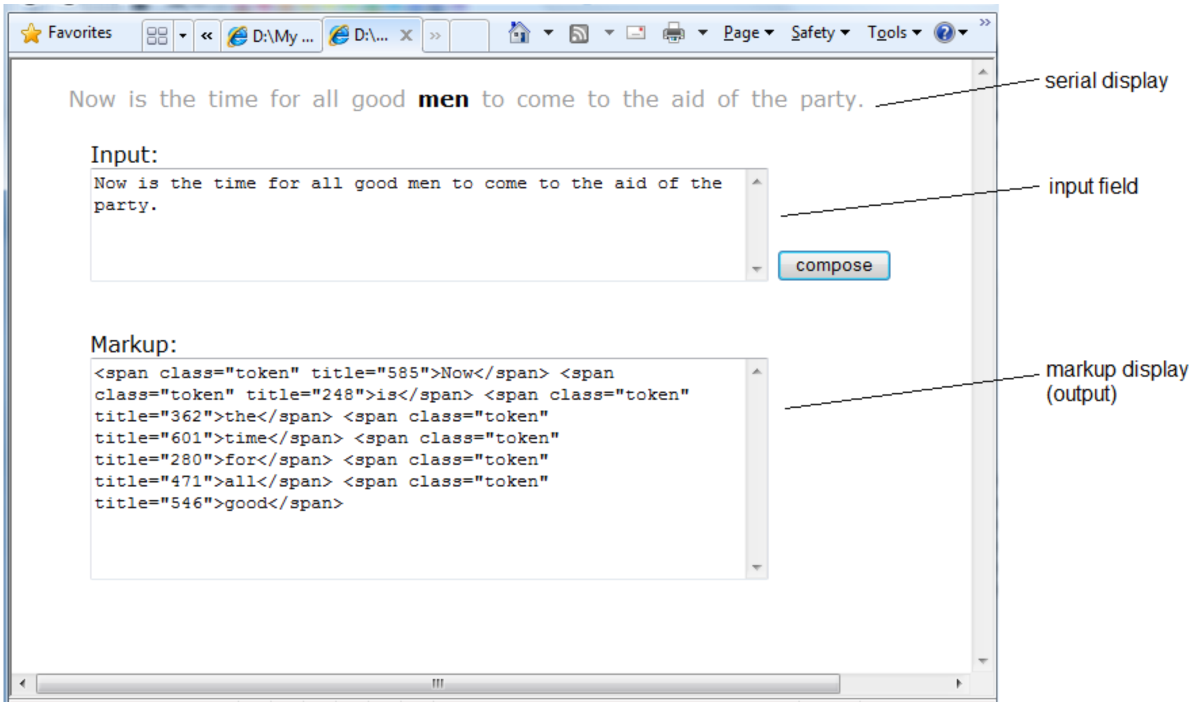

Figure 4.1 - The Composer Engine Interface. Regular text is inputted into the composer which generates a serial presentation that a human reader uses to ascribe spoken durations on each word to create a rhythmic pattern.

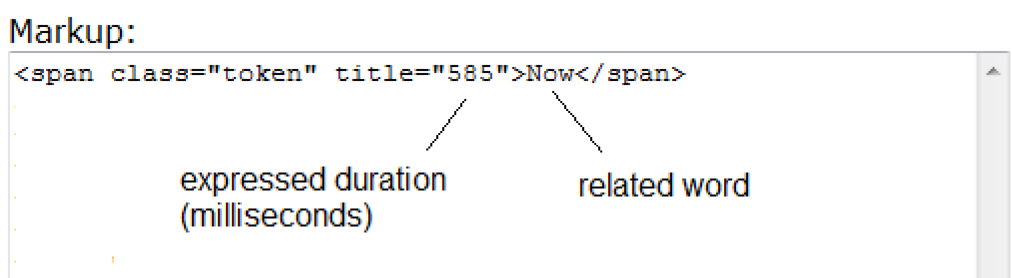

Figure 4.2 - The "Rhythmic Patterned" Markup. The recorded duration on each word as well as the word itself is generated in the form of markup text. Each 'span' element in the markup text represents a' word token'.

Once the word tokens are fully generated by the composer engine as markup text, they are organized by the display engine and further processed by the prosody renderer to produce prosody-enriched text for the dynamic text presentation.

4.1.2 The Prosody Renderer

The purpose of the prosody renderer is to analyze the rhythmic pattern expressed in the recorded word durations of all word tokens and based on this, apply the appropriate visual prosodic effect on the current token in focus. Upon initializing, the prosody renderer interfaces with the organized word tokens within the display engine and performs an analysis procedure that derives the average display time of all words. The data from this initial analysis is subsequently applied during the presentation itself to appropriately project the representational effects of rhythm, emphasis and defined phrases. When projecting rhythm, the prosody renderer dynamically computes the exposure time for each focused word in the display based on the recorded duration retrieved from the word token and the current rate reported by the presentation controller. It then sequences the horizontal scrolling in the display engine based on the computed exposure time of each word to produce the rhythmic effect. The projection of prosodic phrase boundaries is inherent in the projected rhythm through noticeable pauses. With respect to emphasis, the prosody renderer utilizes the average display time of all words to determine whether the focused word should be emphasized. If a word must be emphasized, the prosody renderer then calculates the intensity value which is then used to apply one of four possible intensity representations of emphasis.

4.1.3 The Presentation Controller

The presentation controller is an interface-related component designed according to the recommendations set out in the model. Focusing on ease of use as well as mitigating the negative effects of constant button pressing in previous user-controlled presentations, the presentation controller was implemented as a modified horizontal slider control.

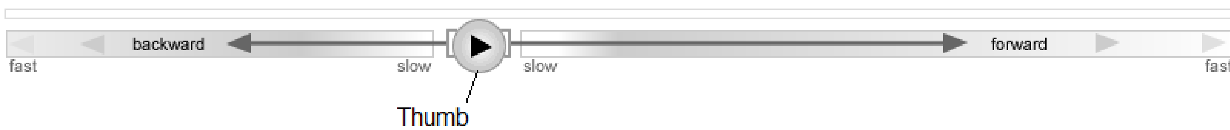

Figure 4.3 - The Presentation Controller. The designed controller allows a variable rate and presentation direction to be specified by the user without constant button presses.

Depending on the region of the slider's thumb, the presentation may regress, pause or progress. The horizontal slider additionally controls the rate of presentation in both directions. The presentation rate gradually increases as the thumb is moved along the slider, away from the center region. This rate is mapped to the location of the thumb on the slider by a quadratic equation (equation. 4.1).

$$\text{multiplier} = |\, 0.000036\mathbf{sx}^2 - 0.025737\mathbf{sx} + 5.573922 \,| \qquad \textbf{Equation 4.2}$$

The *multiplier* is based on the x-coordinate of the thumb (**sx**) along the slider.

The quadratic equation is used to taper the acceleration of the display rate along a manageable acceleration curve. The *a*, *b* and *c* values for this equation ensure a customized tapering. Through this acceleration curve, users would be given a wider range of display rates and would be able to find a comfortable display rate easily as opposed to a linear acceleration curve. The equation itself produces a multiplier value which is then used as a factor by which the display intervals of each word token are either lengthened or shortened, producing either a slower or faster display rate.

The user requirement to move the thumb along the axis was designed to be as effortless as possible. To accomplish this, the thumb is given a "sticky" attribute which binds it to the movement of the mouse, requiring no dragging or constant clicking but only lateral mouse movement to control the direction, state and speed of the presentation. To begin reading, the user is required to click once on the thumb which binds it to the lateral location of the mouse cursor. The user may optionally click once more on the thumb to pause the presentation and free the mouse.

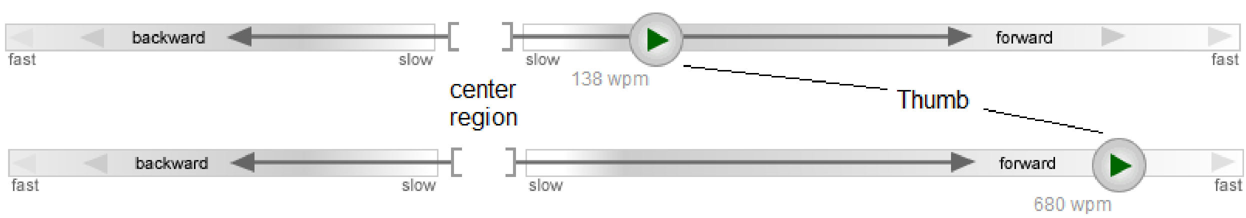

Figure 4.4 - Using The Presentation Controller. The presentation rate gradually increases as the thumb is moved further away from the center region. The thumb is also given a "sticky" attribute which binds it to lateral mouse movements removing the need for clicking and dragging.

4.1.4 The Display Engine

The display engine renders formatting, animates horizontal scrolling, provides completion indication as well as organizes and displays the word tokens in the dynamic presentation. Further it provides a programmatic reference of each word token and its duration for processing by the prosody renderer. It can be seen as the implementation of the dynamic display format as illustrated in the model. The implemented display engine accepts rhythmic patterned text and initially formats the text under the display type, polarity and font specifics outlined in the model. The display engine additionally collaborates with the prosody render and presentation controller to project prosodic representations on text as well as direct the rate and direction of the presentation respectively. Once input is received from the presentation controller regarding the direction and pace of presentation, the display engine then references the prosody renderer for added specifics on the applied prosodic effect and scrolling rhythm for the current word in focus.

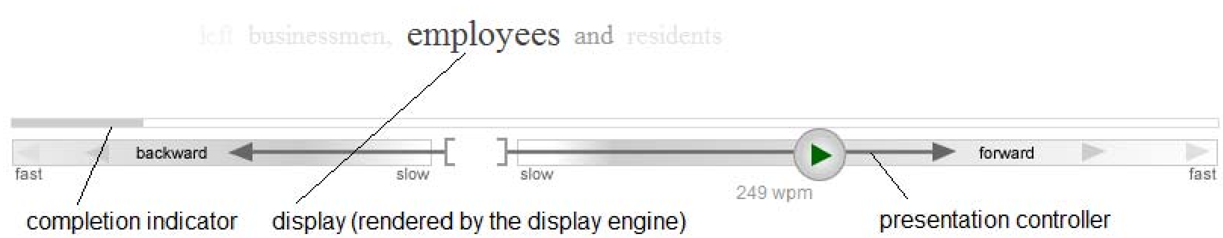

Figure 4.5 - The PEDText Reader. The completed PEDText reader based on the designed model.

CHAPTER 5. EVALUATION

PROSODY-ENRICHED DYNAMIC TEXT PRESENTATION

To assess the effects of the prosody-enriched dynamic text presentation technique, user evaluations were carried out. The aim of the evaluation was to see how the prosody-enriched text presentation compared to the traditional text presentation with respect to reading speed and comprehension. This chapter begins by outlining the details of the experiment used in the evaluation. This is followed by an analysis of the results gained from the experiment.

5.1 Method

5.1.1 Subjects

Twenty-four subjects participated in the experiment. The subjects were undergraduate university students with a mean age of 20. Fifteen were male while nine were female and all of the subjects were computer literate. They all participated in the experiment in exchange for extra credit towards an introductory programming course.

5.1.2 Setting

The evaluation exercise was carried out in a computer lab. The subjects were seated as shown in Figure 5.1.

Figure 5.1 - Subjects at the evaluation exercise.

5.1.3 Evaluation Method

The experimental method was used for the evaluation exercise. Under this method, the participants were divided into two groups, the treatment group and the control group. It was ensured that the groups were balanced with respect to numbers and testing variables such as environment and testing apparatus. The established dependent variables (D.V.) to be measured were reading rate and comprehension, whereas, the text presentation technique served as the independent variable (I.V.) and was altered for the treatment group. In the treatment group, participants read from the developed PEDText reader which uses the prosody enriched dynamic text presentation technique, while, in the control group, participants read from traditional electronic text – the baseline against which the results were judged.

The following null hypotheses were formulated for reading from the PEDText reader:

- Ho 1. The PEDText reader has no effect on reading rate compared to the traditional page format.
- Ho 2. The PEDText reader has no effect on comprehension compared to the traditional page format.

The following sections describe the methods used to evaluate reading rate and comprehension in the experiment.

5.1.4 Reading Rate Evaluation Method

The reading rate in words per minute for both groups was calculated by dividing the total words in an administered text by the total time taken (in minutes) to read the text in its entirety. The timing and calculation of the total words were designed to be handled programmatically in the experiment.

5.1.5 Comprehension Evaluation Method

To gather unbiased comprehension scores, the comprehension evaluation exercises were designed in such a way that the reader will have no reference to the material once s/he chooses to move on to the questions. With this in mind, comprehension texts in which passages must be referenced when answering questions such as those used in some standardized tests were not used.

As a metric to ensure that each chosen text was of sufficient difficulty, the widely Flesch-Kincaid readability test (Flesch, 1948) available in Microsoft Word was used. This readability test is widely used to assess comprehension difficulty when reading contemporary English passages. In the Flesch-Kincaid test, the readability score as well as the associated grade level are calculated based on the total words, sentences and syllables in each word. The readability score may be interpreted based on the table shown in Table 5.1.

Table 5.1

Flesch Readability Score Descriptions

Score (%)	Description
90.0–100.0	Very easy (easily understandable by an average 11-year-old student
60.0–70.0	Standard (easily understandable by 13- to 15-year-old students)
50.0 – 59.0	Fairly difficult
30.0 – 49.0	Difficult
0.0–30.0	Very difficult (best understood by university graduates)

Reading texts with a readability score between 40% and 50% were chosen which best applies to the undergraduate participants used in the evaluation.

5.1.5.1 Reading Texts

The texts were presented in the control group at 18-point ‘Times New Roman’ font. This was chosen to match the base font specifics of the PEDText reader used by those in the treatment group. The following table lists the pool of texts used and their related readability rating.

Table 5.2

Reading texts used in the evaluation

Passage	Title	Source	Words	Flesch Score
1	Airbus Crisis Over	BEC Business English Certificate Preliminary Exam	360	44.4
2	Which is the hardest language?	Intermediate comprehension exercise for ESL Exam	488	40.3
3	The Digital Divide	PET Preliminary English Test	224	50.5

5.1.6 The Evaluation Procedure

The procedure used to gather reading rate and comprehension scores in both groups utilized a repeated-measurement layout. In addition to an initial training phase to acquaint subjects with the given text presentation technique and evaluation methods, they were required to complete three rounds of reading passages and answering comprehension questions. Based on the applied repeated-measurement layout, the evaluation procedure was divided into seven phases: (1) training, (2) reading rate evaluation I, (3) comprehension evaluation I, (4) reading rate evaluation II, (5) comprehension evaluation II, (6) reading rate evaluation III and (7) comprehension evaluation III. Finally, to gather usability data on the PEDText reader, a questionnaire was administered only to the treatment group after they completed the seven outlined phases of the evaluation.

5.1.7 Instructions

All received instructions before the experiment began that explained the procedure, purpose and requirements of the exercise. Subjects in the treatment group were instructed to gradually increase the text presentation rate on the PEDText reader to their highest comfortable pace. They were informed that their reading rates would be automatically tracked and reported by the PEDText reader.

The subjects in the control group were instructed to click a button to signify when they began reading a passage as well as to click another button upon completion of their reading. It was explained that these buttons facilitate a tracking mechanism that calculates their reading rate.

5.1.8 Apparatus

The experiment was performed on desktop computers with 19” LCD displays bearing the resolution of 1280x1024 pixels. The evaluation exercise was implemented in the form of a web-based application viewable through Microsoft Internet Explorer. The web-based application was based on a distribution of Drupal 6.0 with the webform module enabled. Apart from providing centralized access to the evaluation exercise, this allowed a “wizard-like” progression of the various phases of the exercise as well as the utility of computing and organizing subject-related data.

5.1.9 Caveats

Two issues arose while performing the experiment which may affect results. The first of which is that the testing environment was not absolutely quiet while subjects were reading. Some subjects were conversing during the exercise which may have possibly been disturbing to other subjects. Secondly, despite the voiced and written procedure and instructions of the exercise, some subjects were unaware that there would be comprehension questions following the presented passage.

5.2 Results

The intention of this research was to evaluate the effects of the prosody enriched dynamic text presentation technique against traditional text presentation in the reading of electronic text. This section presents the results of the controlled experiment as well as the statistical analysis of the results under the categories of reading rate and comprehension scores. The T-Test was used as the test of significance for the data gathered from the treatment and control groups. The SPSS version 16 statistical package was used to generate results of the T-Test with a set significance level of 5% ($p<0.05$).

5.2.1 Reading rate

The reading rate data collected during the experiment is presented in this section. The results indicate that persons who read from the PEDText reader (treatment) had significantly lower reading rates compared to those who read from the traditional format (control).

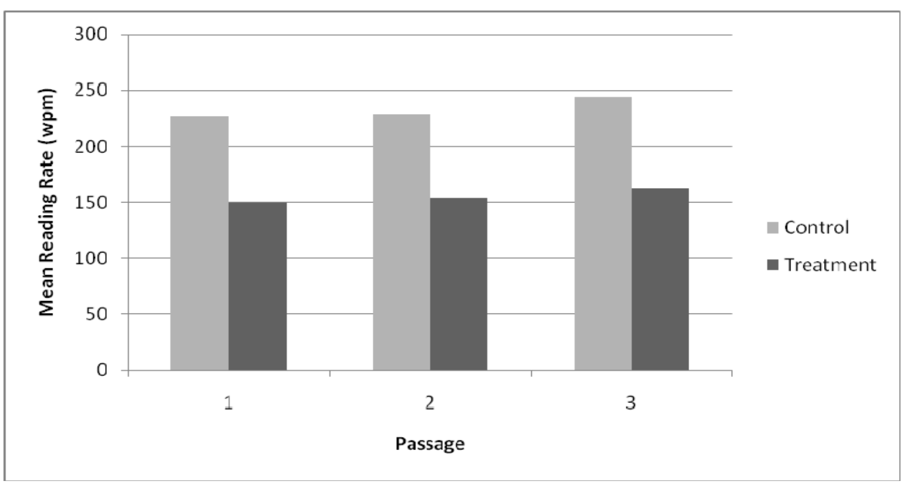

Figure 5.2 - Mean Reading rates from the treatment and control groups.

Table 5.3
Mean Reading Rates gathered from the experiment

	Control (N=12)		Treatment (N=12)			
Reading Passage	Mean Reading Rate (wpm)	Std. Deviation	Mean Reading Rate (wpm)	Std. Deviation	t-value	Sig. (p) (2-tailed)
1	226.9167	82.13786	149.5833	59.56579	3.513	0.005 *
2	229.1667	45.87698	154.1667	56.47338	3.331	0.007 *
3	243.6667	88.07985	161.8333	44.86309	2.581	0.026 *

(* $p< 0.05$)

As a result of the analysis carried out using the t-test, the resulting t-values were found to be significant ($p<0.05$) for all passages (see table 5.3) in favor of the control group. According to the results, the null hypothesis Ho. 1 is rejected. The PEDText reader significantly reduced the reading rates of the subjects compared to the traditional page format.

5.2.2 Comprehension

The comprehension data collected during the experiment is presented in this section. The comprehension score was computed as the percent of correctly answered multiple choice questions. The gathered results indicate that subjects who used the PEDText reader (treatment) visibly matched or exceeded the comprehension scores of those who read from the traditional text format (control) (see Figure 5.3).

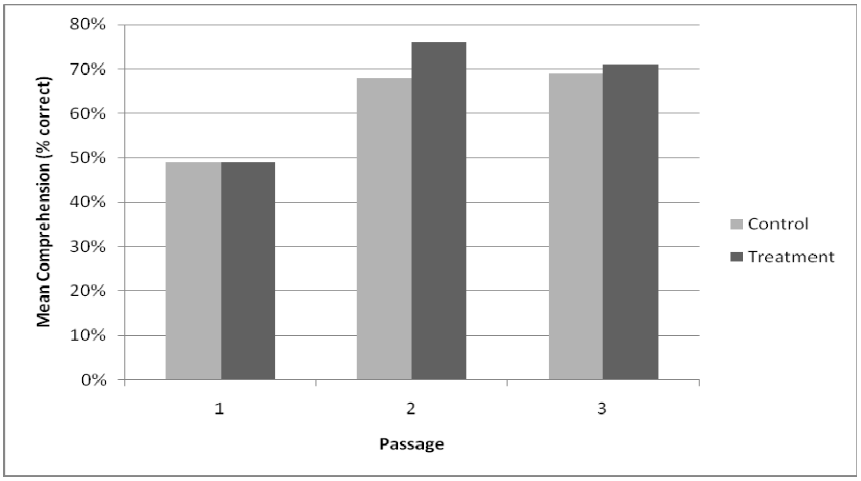

Figure 5.3 - Mean Comprehension scores from the treatment and control groups.

Table 5.4

Mean Comprehension Scores gathered from the experiment

	Control (N=12)		Treatment (N=12)			
Reading Passage	Mean Comprehension (% correct)	Std. Deviation	Mean Comprehension (% correct)	Std. Deviation	t-value	Sig. (p) (2-tailed)
1	49%	16.76215	49%	16.57193	0.015	0.988
2	68%	19.47026	76%	12.70260	-2.124	0.057
3	69%	17.29862	71%	14.43376	-0.290	0.777

(* p<0.05)

Based on the t-test analysis, carried out on the comprehension scores, the resulting t-values were found to be at a notable level of significance (p=0.057) for one of the three passages in favor of the treatment group. Based on these results, Ho. 2 is rejected. Although not truly significant, the PEDText reader produced a notably positive effect on comprehension in one of the three passages.

5.2.3 User Experience

The user experience data based on the PEDText reader is presented in this section. The responses gathered for each question in the treatment group questionnaire were presented as a calculated percentage of the total number of subjects in the treatment group.

Question 1: "How did the selectively enlarged words affect your reading?"

This question aimed to gather the effectiveness of the visual representation of emphasis in the design. The available response options were: (1) '[the selectively enlarged words] helped my comprehension'; (2) '[the selectively enlarged words] had no effect'; (3) '[the selectively enlarged words] confused me'. Results show that the majority of the subjects (50%) indicated that the visual representation of emphasis in the selectively enlarged words assisted their comprehension. In addition, only 8.33 % of subjects were reportedly confused by the visual representation of emphasis. The remainder of subjects (41.67%) indicated that the visual representation of emphasis made no difference (see Figure 5.4).

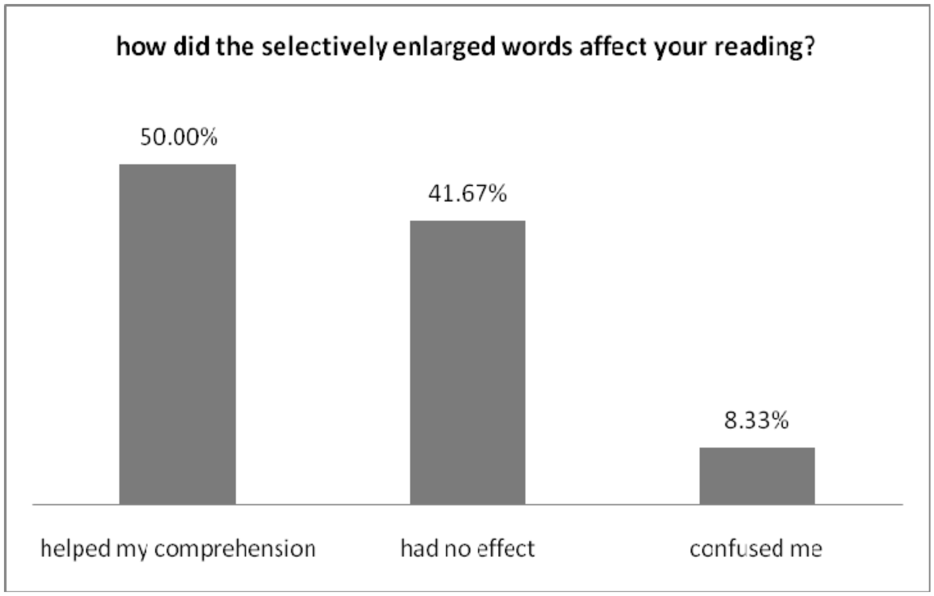

Figure 5.4 – Responses to Question 1 of the treatment group questionnaire.

Question 2: "How did the 'rhythm' of the moving words affect your reading?"

The purpose of this question was to assess the effectiveness of the projected representations of rhythm and prosodic phrase boundaries. The available response options were (1) '[the rhythm of the moving words] helped my comprehension'; (2) '[the rhythm of the moving words] had no effect'; (3) '[the rhythm of the moving words] confused me'. The results indicate that a majority of the subjects (58.3%) were reportedly aided in their comprehension by the projected rhythm. Only 8.33% of subjects were confused by this representation while the remaining 33.3% of subjects experienced no difference in their reading experience with the projected rhythm (see Figure 5.5).

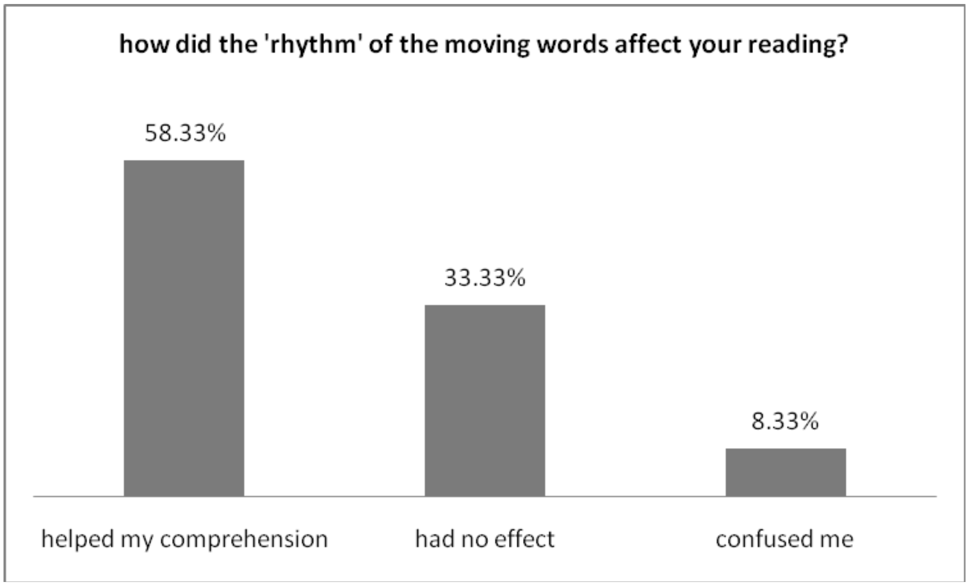

Figure 5.5 – Responses to Question 2 of the treatment group questionnaire.

Question 3: "What kept you from increasing the speed of the words?"

This question was designed to gather user feedback on their experience with the chosen display type and the user-controlled presentation rate. The available response options were: (1) 'it was difficult to read, even at lowest speed (< 150 wpm)', (2) 'I had no problem with adjusting speed above 200 wpm', (3) 'I was unaware the speed can be adjusted'. The third option was included to cater for any cases in which the subjects may have misused the reader. Results show that an equal number of subjects had problems with reading from display type at increased speeds as those who did not (41.7%). In spite of expressed directions at the beginning of the exercise, 16.7% of subjects were unaware that the speed could have been increased through the user-controls.

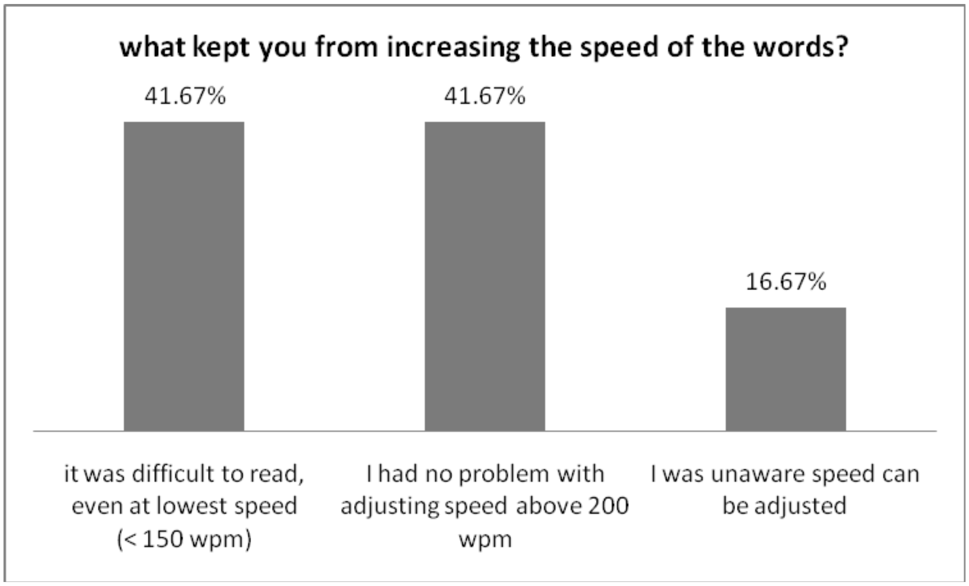

Figure 5.6 – Responses to Question 3 of the treatment group questionnaire.

CHAPTER 6. DISCUSSION

Results of the experiment indicate that subjects had significantly lower reading rates with the PEDText reader than in the traditional format. In spite of this, comprehension with the PEDText reader was higher, though not significantly higher than the traditional presentation format. More interesting is the positive user responses to the projection of emphasis and rhythm in the presentation technique. In this section these findings will be discussed along with related issues. After which, conclusions are drawn and directions are set out for future work of this thesis.

6.1 Reading rate

The significantly lower reading rates with the PEDText reader may be attributed to a combination of unfamiliarity with the reader and the user-controlled presentation of text. Based on the responses gathered from the post-test questionnaire, 41.7% of users indicated that the dynamically presented text was difficult to read, even at low speeds (< 150 wpm). As a result, these subjects never felt comfortable with increasing the presentation rate of the PEDText reader to a higher pace. This suggests that their unfamiliarity with the RSVP-like display type may have been to blame. Alternatively, an equal percentage of subjects indicated that they had no problem reading from the utility at a presentation rate of 200 wpm or more. Within this subset of subjects, reading rates in excess of 250 wpm were achieved by a select few. However this did little to improve the overall low reading rate of the group. It is likely that the unfamiliarity with the display type coupled with the provision of presentation rate controls and the ability to regress affected the majority of subjects and severely impacted the overall reading rate when using the PEDText utility. Previous studies by Juola (1988) and Muter *et al*. (1988) as quoted by Castelhano and Muter (2001) have reported that attempts to have users control the presentation rate of RSVP text displays have been “disastrous in terms of efficiency”. Muter *et al*. added a subject-controlled pacing feature to traditional RSVP along with the ability to regress with the aim of providing increased control to subjects. Results from this modification showed a significant reduction in reading speed without improvements in comprehension. Clearly, the PEDText reader, using a RSVP-like display type, produced a similar user reaction due to the granted level of control in the form of the PEDText presentation controller. The difference, however, is in the visibly improved comprehension scores with the PEDText reader.

6.2 Comprehension

In spite of low reading rates, comprehension scores with the PEDText reader were either matched or visibly higher than the traditional format. Passage 1 bearing a Flesch readability score of 44.4% at 360 words long received the lowest comprehension score of 49% with both the PEDText reader and traditional format. It is possible that the low comprehension score in both groups for the first passage may be a result of a caveat of the evaluation. Although instructions and a description of the exercise were expressed verbally and in written form at the beginning of the exercise, some subjects only became aware that comprehension questions followed each passage after they progressed to the comprehension questions of passage 1. Based on this, it would appear that until this point, subjects were not reading with intent. Passage 3, bearing a Flesch readability score of 50.5% and a length of 224 words was read with a negligible increase in comprehension with the PEDText reader where scores rose from an average of 69% in the traditional to 71% with the reader. However, the notable increase occurred when passage 2 was read. Passage 2 with a Flesch readability score of 40.3% at 488 words long received higher comprehension scores through the PEDText reader over the traditional format with a notable level of significance ($p=0.057$). Scores in passage 2 rose from an average of 68% with the traditional format to an average of 76% with the PEDText reader. Although the established level of significance at $p<0.05$ ruled out a significant comprehension increase within 95% certainty according to the t-test, the increase with the PEDText reader can be described as ‘marginally significant’ since it falls within the 90% certainty range of significance. In addition, given that passage 2 was the longest and least readable passage of the three, the comprehension gains in this light are quite encouraging.

6.3 User Experience

A post-test questionnaire administered to subjects who used the PEDText reader gathered data on their reading experience. The main purpose of the questionnaire was to assess the effectiveness of visual prosodic representations in the presentation on subject comprehension. Subject responses primarily indicate a positive effect on their experienced comprehension with the visual prosodic representations. Fifty percent of subjects indicated that the visual representation of emphasis in the form of selectively enlarged words aided their comprehension of the presented content while a negligible 8.3% reported confusion due to the effect. Further, 58.3% of subjects responded to the positive effects of the projected rhythm on their experienced comprehension with only 8.3% of subjects experiencing confusion from the effect. Together with the recorded increases in comprehension over the traditional format, these results are notable indicators to the effectiveness of the synthetically projected prosodic properties on dynamic text.

CHAPTER 7. CONCLUSIONS

This thesis investigated the effectiveness of visually representing multiple prosodic properties in a dynamic text presentation as a means to support or enhance comprehension at higher-than-average rates of reading. Two major discoveries were made in this research to devise a dynamic text presentation technique that exploits the high reading rate potential of RSVP while mitigating its negative effects on comprehension.

Firstly, the provision of self-paced controls coupled with the designed display type used in the prosody-enriched technique rendered no increases in reading rate, in fact, results show significantly low reading rates compared to the traditional page format. This outcome is consistent with earlier studies by Juola (1988) and Muter *et al.* (1988) that prove the incorporation of self-paced controls in RSVP negatively impacts efficiency. The self-paced controls in the prosody-enriched technique were provided to increase user comfort; however, the unfamiliarity with the chosen display type produced significant discomfort in most users which forced them to reduce their reading pace with the available controls. From this outcome, it is concluded that a proper balance of user comfort must be created between the designed display type and the degree of control given to a user in order to achieve reliable increases in reading rate.

Secondly, the visual representations of emphasis, prosodic phrase boundaries and rhythm incorporated into the prosody-enriched dynamic text presentation technique produced a notably positive influence on comprehension over traditional page formats. In relation to this result, the majority of subjects that used the prosody-enriched technique indicated improved readability due to the visual representations of prosody rendered in the dynamic text presentation. Furthermore, an effect of this improvement can be deduced from the notable increase in comprehension of the longest and least readable passage in the evaluation. It can then be concluded that the synthetic projection of prosody on dynamic text does indeed have positive effects on comprehension of dynamic text and indicates encouraging results when applied to content.

Overall, the current implementation of the designed prosody-enriched dynamic text presentation technique was unable to induce high rates of reading due to the design of the display type and the provision of self-paced controls. In spite of this, visually projecting prosodic representations proves to be a promising method of positively influencing comprehension in dynamic text presentations.

7.1 Future Work

Future work in this research comprises further refining the current display type to readily promote high rates of reading while facilitating the projection of visual representations of prosody on text. For this refinement, an iterative incremental prototyping approach will be used where each subsequent version benefits from feedback obtained from informal user evaluations. This iterative incremental prototyping approach should ensure the highest effectiveness with the refined reader. In addition, efforts will be made to replace the manual method of composing content-based rhythmic patterns with one that is automated. Finally, a future study is planned to assess the effectiveness of the refined prosody-enriched dynamic text presentation technique on reading texts with low readability ratings.

APPENDICIES

A. Control Group Reading Rate Results

	Passage 1		Passage 2		Passage 3	
IP Address	reading rate (wpm)	comprehension (% correct)	reading rate (wpm)	comprehension	reading rate (wpm)	comprehension (% correct)
138.238.128.202	207	33%	180	50%	203	40%
138.238.148.125	223	50%	176	38%	207	60%
138.238.121.7	390	50%	283	75%	296	80%
138.238.128.129	183	50%	226	38%	176	80%
138.238.128.128	225	17%	239	75%	144	80%
138.238.128.132	130	67%	235	75%	460	70%
138.238.128.124	380	33%	310	100%	292	80%
138.238.128.222	243	67%	294	63%	336	80%
138.238.128.218	245	50%	196	63%	182	50%
138.238.128.219	179	67%	230	63%	202	40%
138.238.128.220	172	67%	195	88%	215	90%
138.238.128.223	146	33%	186	88%	211	80%

A1. Control Group Background Data

IP Address	age	Sex	computer literacy	university level	mother tongue	do you speed-read?
138.238.128.202	26	Female	Medium	sophomore	english	no
138.238.148.125	25	Male	Medium	sophomore	other	no
138.238.121.7	19	Male	Medium	freshman	english	yes
138.238.128.129	18	Male	Medium	freshman	english	no
138.238.128.128	18	Female	Medium	freshman	english	no
138.238.128.132	18	Male	Medium	freshman	english	no
138.238.128.124	17	Female	Low	freshman	english	no
138.238.128.222	18	Male	Medium	freshman	english	no
138.238.128.218	18	Female	Medium	freshman	english	no
138.238.128.219	17	Male	High	freshman	english	yes
138.238.128.220	21	Female	High	sophomore	english	no
138.238.128.223	19	Male	Low	sophomore	english	no

A2. Treatment Group Reading Rate Results

	Passage 1		Passage 2		Passage 3	
IP Address	reading rate (wpm)	comprehension (% correct)	reading rate (wpm)	comprehension (% correct)	reading rate (wpm)	comprehension (% correct)
138.238.128.213	103	17%	135	60%	135	60%
138.238.128.203	185	33%	183	63%	171	70%
138.238.128.204	139	50%	139	63%	143	90%
138.238.128.205	136	50%	124	75%	134	80%
138.238.128.216	122	50%	118	75%	136	60%
138.238.128.207	107	83%	129	100%	156	70%
138.238.128.210	306	50%	314	75%	148	40%
138.238.128.212	129	67%	126	88%	127	70%
138.238.128.215	116	50%	135	75%	221	80%
138.238.128.217	110	50%	167	63%	253	60%
138.238.148.109	125	33%	98	88%	103	90%
138.238.148.112	217	50%	182	88%	215	80%

A3. Treatment Group Background Data

IP Address	age	Sex	computer literacy	university level	mother tongue	do you speed-read?
138.238.128.213	22	Male	High	senior	english	no
138.238.128.203	22	Female	Medium	sophomore	english	no
138.238.128.204	20	Female	High	junior	english	yes
138.238.128.205	18	Female	High	freshman	english	no
138.238.128.216	18	Male	High	freshman	english	no
138.238.128.207	21	Male	Medium	senior	other	no
138.238.128.210	21	Female	High	senior	english	no
138.238.128.212	18	Male	High	freshman	english	no
138.238.128.215	18	Male	High	freshman	english	no
138.238.128.217	18	Male	Medium	freshman	english	no
138.238.148.109	20	Male	High	sophomore	other	no
138.238.148.112	24	Male	High	sophomore	english	no

A4. Treatment Group Post-Test Questionnaire Results

IP Address	how did the selectively enlarged words affect your reading?	how did the 'rhythm' of the moving words affect your reading?	what kept you from increasing the speed of the words?	any additional comments?
138.238.128.213	they helped my comprehension	it helped my comprehension	I found it difficult to read words, even at the lowest speed.	It was pretty cool. The more I used it, the better I became at using it. I felt that I could increase the speed as I progress through the program.
138.238.128.203	they had no effect	it had no effect	I found it difficult to read words, even at the lowest speed.	
138.238.128.204	they had no effect	it helped my comprehension	I reached a decent speed (>200wpm), this does not apply to me.	
138.238.128.205	they had no effect	it had no effect	I did not know that the speed can be increased	
138.238.128.216	they confused me	it confused me	I found it difficult to read words, even at the lowest speed.	
138.238.128.207	they helped my comprehension	it helped my comprehension	I reached a decent speed (>200wpm), this does not apply to me.	I would not use this utility regularly, simply because the text is presented linearly, and the enlargement of some text and shrinkage of other text is not appealing to the eye. If all the text was presented in the same font face and size (preferably 16 or larger), it would be much more appealing to the eye for reading from a computer screen.
138.238.128.210	they had no effect	it helped my comprehension	I reached a decent speed (>200wpm), this does not apply to me.	
138.238.128.212	they helped my comprehension	it had no effect	I reached a decent speed (>200wpm), this does not apply to me.	

138.238.128.215	they helped my comprehension	it helped my comprehension	I found it difficult to read words, even at the lowest speed.	the environment was a little distracting (people talking, mainly) but atleast it was constant for all participants. i didn't remember part of the instructions for the first one, so i didn't remember that there'd be a comprehension test. perhaps if it was mentioned a second time, in bold, in color or something i think i would have noticed more of what i was reading. but i think that was really my fault. but it was simple enough to figure out. and i have ADHD and dyslexia so i definately would perfer to read like that. perhaps this tool could be used by people who have Trouble reading... i think there would be a big market behind that idea (i want 10%!) The fact that you can change the pace also makes it interactive, so that helped me stay "into" what was going on (this is, of course, after i learned that there was a comprehension test)
138.238.128.217	they helped my comprehension	it had no effect	I found it difficult to read words, even at the lowest speed.	
138.238.148.109	they helped my comprehension	it helped my comprehension	I did not know that the speed can be increased	
138.238.148.112	they had no effect	it helped my comprehension	I reached a decent speed (>200wpm), this does not apply to me.	

B. Control Group Evaluation Description

Home

PEDText Evaluation Overview - Control Group

The purpose of this experiment is to evaluate a novel reading utility that is designed to enhance reading efficiency over traditional pages of text.

The evaluation exercise has seven phases:

- Phase 1 - Training
  Here we allow you to get acquainted with the controls used for this evaluation.
  You may move onward with the evaluation once you are comfortable with these controls.
- Phase 2 - Reading Rate Evaluation I
  Here we test your reading rate with a traditional page of text.
- Phase 3 - Comprehension Evaluation I
  Here we test your comprehension based on the text you just read.
- Phase 4 - Reading Rate Evaluation II
  Here we test your reading rate with a traditional page of text.
- Phase 5 - Comprehension Evaluation II
  Here we test your comprehension based on the text you just read.
- Phase 6 - Reading Rate Evaluation III
  Here we test your reading rate with a traditional page of text.
- Phase 7 - Comprehension Evaluation III
  Here we test your comprehension based on the text you just read.

MAKE SURE YOU USE ONLY INTERNET EXPLORER for this exercise.

Begin Training

Thank you for your participation.

Phase 1 - Training - Reading Rate Evaluation ›

B1. Control Group Training Instructions for Reading Rate Evaluation

Home » PEDText Evaluation Overview - Control Group

Phase 1 - Training - Reading Rate Evaluation

This is the training phase. You may skip ahead if you are comfortable with the controls.

Description

Each passage you read will have "begin reading.." and "end reading.." buttons.
These are to compute your reading speed.

You must click on the "begin reading" button the moment you begin to read and click on the "end reading" button the moment you are finished reading.

At this time your reading rate will be displayed and you will be required to click on the "continue" button to progress to the questions related to what you just read.

click here then begin reading

Lorem Ipsum

You are reading dummy text for training purposes only. Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum. Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, consectetur, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "de Finibus Bonorum et Malorum" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, "Lorem ipsum dolor sit amet..", comes from a line in section 1.10.32.

click here when you finish reading

continue

B2. Control Group Training Instructions for Comprehension Evaluation

Home » PEDText Evaluation Overview - Experimental Group

Phase 1 - Training - Comprehension Evaluation

This is the training phase. You may skip ahead if you are comfortable with the controls.

Description

The comprehension questions will be multiple choice.
You will be required to answer all questions and then click on the "Grade Test" button.
Your test will then be scored and you will be allowed to proceed by clicking on "continue".

1. **This is a sample multiple choice question:**
   - True
   - False
2. **This is yet another sample multiple choice question.**
   - True
   - False
3. **Another sample multiple choice question.**
   - True
   - False
4. **This is a sample multiple choice question.**
   - True
   - False
5. **This is another sample multiple choice question**
   - True
   - False
6. **This is yet another sample multiple choice question.**
   - True
   - False
7. **This is a sample multiple choice question.**
   - True
   - False
8. **This is a sample multiple choice question.**
   - True
   - False

Grade Test

score: *

your comprehension score

continue

B3. Treatment Group Evaluation Description

Home

PEDText Evaluation Overview - Experimental Group

The purpose of this experiment is to evaluate a novel reading utility that is designed to enhance reading efficiency over traditional pages of text.

The evaluation exercise has seven phases:

- Phase 1 - Training
  Here we allow you to get acquainted with the controls used for this evaluation.
  You may move onward with the evaluation once you are comfortable with these controls.
- Phase 2 - Reading Rate Evaluation I
  Here we test your reading rate with our PEDText utility.
- Phase 3 - Comprehension Evaluation I
  Here we test your comprehension based on the text you just read.
- Phase 4 - Reading Rate Evaluation II
  Here we test your reading rate with our PEDText utility.
- Phase 5 - Comprehension Evaluation II
  Here we test your comprehension based on the text you just read.
- Phase 6 - Reading Rate Evaluation III
  Here we test your reading rate with our PEDText utility.
- Phase 7 - Comprehension Evaluation III
  Here we test your comprehension based on the text you just read.

MAKE SURE YOU USE ONLY INTERNET EXPLORER for this exercise.

Begin Training

Thank you for your participation.

Phase 1 - Training - Reading Rate Evaluation ›

B4. Treatment Group Training Instructions for Reading Rate Evaluation

Home » PEDText Evaluation Overview - Experimental Group

Phase 1 - Training - Reading Rate Evaluation

This is the training phase. You may skip ahead if you are comfortable with the controls.

Instructions

Each passage will be dynamically presented through the PEDText utility you see below.

To operate:

1. *Click on "play button". Your cursor will 'stick' to the button*
2. *Move the cursor and button to the right.*
3. *Stop moving the cursor when you have found the highest comfortable pace.*
4. *You can click the button at any time to pause the words. Or move the cursor to the far left to go back and re read*

Take your time to get acquainted with the utility.

reading rate: *

your calculated reading rate in words per minute (wpm)

continue

B5. Treatment Group Training Instruction for Comprehension Evaluation

Home » PEDText Evaluation Overview - Experimental Group

Phase 1 - Training - Comprehension Evaluation

This is the training phase. You may skip ahead if you are comfortable with the controls.

Description

The comprehension questions will be multiple choice.

You will be required to answer all questions and then click on the "Grade Test" button.

Your test will then be scored and you will be allowed to proceed by clicking on "continue".

1. **This is a sample multiple choice question:**
   - True
   - False
2. **This is yet another sample multiple choice question.**
   - True
   - False
3. **Another sample multiple choice question.**
   - True
   - False
4. **This is a sample multiple choice question.**
   - True
   - False
5. **This is another sample multiple choice question**
   - True
   - False
6. **This is yet another sample multiple choice question.**
   - True
   - False
7. **This is a sample multiple choice question.**
   - True
   - False
8. **This is a sample multiple choice question.**
   - True
   - False

Grade Test

score: *

your comprehension score

continue

C. Treatment and Control Group Background Data Questionnaire

age: *

state your age

sex: *

- ○ male
- ○ female

specify your gender

university level: *

select...

select your university level

computer literacy: *

- ○ low
- ○ medium
- ○ high

specify your level of computer literacy

mother tongue: *

- ○ english
- ○ other

specify your native language

do you practice any speed-reading techniques?: *

- ○ yes
- ○ no

C1. Treatment Group Post-Test Questionnaire

1. how did the selectively enlarged words affect your reading?: *

- ○ they confused me
- ○ they helped my comprehension
- ○ they had no effect

2. how did the 'rhythm' of the moving words affect your reading?: *

- ○ it confused me
- ○ it helped my comprehension
- ○ it had no effect

the 'rhythm' includes the pauses at punctuations and displaying key words for longer periods of time

3. what kept you from increasing the speed of the words?: *

- ○ I did not know that the speed can be increased
- ○ I found it difficult to read words, even at the lowest speed.
- ○ I reached a decent speed (>200wpm), this does not apply to me.

any additional comments?:

D. Passage 1 – Airbus Crisis Over

Airbus says it has turned the corner after a crisis connected to production problems and turmoil in the boardroom at its A380 super-jumbo project that has gone on for the past year. Speaking at the Paris air show, Louis Gallois, CEO of the European planemaker, said, "Airbus is back."

Airbus, which announced a raft of orders on the first day of the show, is competing with Boeing, its American rival, for the title of the largest planemaker in the world.

Boeing is expected to reveal the numbers of orders for its 787 Dreamliner soon. Airbus orders unveiled on Monday included Qatar Airways confirming a $16bn order for 80 A350 Airbus planes and ordering three A380 super-jumbos for about $750m.

Boeing and Airbus are also competing for orders from aircraft leasing firms. Orders from these companies - who rank highly among the biggest global buyers of aircraft - are often regarded as an indication of how successful a model will be in the long term.

Airbus also secured orders from US Airways that are worth $10bn for 22 of its A350 jets, 60 A320s and ten of its A330-200 wide-body planes.

A few months ago, Airbus unveiled a major cost-cutting programmed aiming to reduce the workforce in Europe by 10,000, as well as announcing a group restructuring. "I can tell you with full confidence that Airbus is back and fully back, as you have started noting yesterday as demonstrated by our first day announcements," said Mr. Gallois on the second day of the air show.

However, Boeing also announced a deal with General Electric (GE) on the show's first day. GE's commercial aviation services placed an order for six 777 Boeing freighters valued at around $1.4bn, to be delivered in the last quarter of 2008.

A Wall Street Journal website report, quoting the Delta operating chief yesterday said that Delta Air Lines were on the verge of ordering as many as 125 Boeing 787 jetliners by the end of this year. However, a spokesman for Delta later said that it had been having conversations "with several aircraft makers" and that "no final decision" had been made on future fleet purchases.

D1. Passage 1 Questions – Airbus Crisis Over

1. **The problems at Airbus:**
   - have been resolved completely.
   - are well on their way to being sorted out.
   - are far from being resolved.
2. **Airbus announced:**
   - a large number of orders on the first day of the show.
   - some orders on the first day of the show.
   - a few orders on the first day of the show.
3. **Qatar Airways ordered:**
   - 83 planes on Monday.
   - 80 planes on Monday
   - 3 planes on Monday.
4. **US Airways:**
   - placed an order for the new super-jumbo.
   - didn't place an order for the new super-jumbo.
   - may have placed an order for the new super-jumbo.
5. **Boeing:**
   - announced sales of the Dreamliner.
   - may sell some Dreamliners to General Electric.
   - may sell some Dreamliners to Delta Airlines.
6. **The Wall Street Journal website report:**
   - was definitely correct.
   - was possibly correct.
   - was definitely wrong

D2. Passage 2 – Which is the hardest language?

People often ask which is the most difficult language to learn, and it is not easy to answer because there are many factors to take into consideration. Firstly, in a first language the differences are unimportant as people learn their mother tongue naturally, so the question of how hard a language is to learn is only relevant when learning a second language.

A native speaker of Spanish, for example, will find Portuguese much easier to learn than a native speaker of Chinese, for example, because Portuguese is very similar to Spanish, while Chinese is very different, so first language can affect learning a second language. The greater the differences between the second language and our first, the harder it will be for most people to learn. Many people answer that Chinese is the hardest language to learn, possibly influenced by the thought of learning the Chinese writing system, and the pronunciation of Chinese does appear to be very difficult for many foreign learners. However, for Japanese speakers, who already use Chinese characters in their own language, learning writing will be less difficult than for speakers of languages using the Roman alphabet.

Some people seem to learn languages readily, while others find it very difficult. Teachers and the circumstances in which the language is learned also play an important role, as well as each learner's motivation for learning. If people learn a language because they need to use it professionally, they often learn it faster than people studying a language that has no direct use in their day to day life.

Apparently, British diplomats and other embassy staff have found that the second hardest language is Japanese, which will probably come as no surprise to many, but the language that they have found to be the most problematic is Hungarian, which has 35 cases (forms of a noun according to whether it is subject, object, genitive, etc). This does not mean that Hungarian is the hardest language to learn for everyone, but it causes British diplomatic personnel, who are generally used to learning languages, the most difficulty. However, Tabassaran, a Caucasian language has 48 cases, so it might cause more difficulty if British diplomats had to learn it.

Different cultures and individuals from those cultures will find different languages more difficult. In the case of Hungarian for British learners, it is not a question of the writing system, which uses a similar alphabet, but the grammatical complexity, though native speakers of related languages may find it easier, while struggling with languages that the British find relatively easy.

No language is easy to learn well, though languages which are related to our first language are easier. Learning a completely different writing system is a huge challenge, but that does not necessarily make a language more difficult than another. In the end, it is impossible to say that there is one language that is the most difficult language in the world.

D3. Passage 2 Questions – Which is the hardest language?

1. **The question of how hard a language is to learn is relevant to both first and second language acquisition:**
   - True
   - False
2. **Portuguese is definitely easier than Chinese.**
   - True
   - False
3. **A Japanese speaker may well find the Chinese writing system easier than a speaker of a European language.**
   - True
   - False
4. **The Hungarian alphabet causes problems for British speakers.**
   - True
   - False
5. **Hungarian is the hardest language in the world**
   - True
   - False
6. **Hungarian has as many cases as Tabassaran.**
   - True
   - False
7. **Many British diplomats learn Tabassaran.**
   - True
   - False
8. **writer thinks that learning new writing systems is easy.**
   - True
   - False

D4. Passage 3 – The Digital Divide

A recent survey has shown that the number of people in the United Kingdom who do not intend to get internet access has risen. These people, who are known as 'net refuseniks', make up 44% of UK households, or 11.2 million people in total.

The research also showed that more than 70 percent of these people said that they were not interested in getting connected to the internet. This number has risen from just over 50% in 2005, with most giving lack of computer skills as a reason for not getting internet access, though some also said it was because of the cost.

More and more people are getting broadband and high speed net is available almost everywhere in the UK, but there are still a significant number of people who refuse to take the first step.

The cost of getting online is going down and internet speeds are increasing, so many see the main challenge to be explaining the relevance of the internet to this group. This would encourage them to get connected before they are left too far behind. The gap between those who have access to and use the internet is the digital divide, and if the gap continues to widen, those without access will get left behind and miss out on many opportunities, especially in their careers.

D5. Passage 3 Questions – The Digital Divide

1. **More people in the UK do not intend to get internet access than before:**
   - True
   - False
2. **The majority of people in the UK are 'net refuseniks'.**
   - True
   - False
3. **Most of those without internet access want to get it.**
   - True
   - False
4. **The minority of the people surveyed in 2005 weren't interested in having internet access.**
   - True
   - False
5. **The main reason for not getting internet access is the cost**
   - True
   - False
6. **High speed intenet is not available everywhere in the UK.**
   - True
   - False
7. **Both costs and speeds are increasing.**
   - True
   - False
8. **Many people think that getting the costs down is the key to this problem.**
   - True
   - False
9. **The digital divide is widening in the UK.**
   - True
   - False
10. **Not having access to the internet will only affect people's careers.**
    - True
    - False

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.

Ashby, J. (2006). Prosody in skilled silent reading: evidence from eye movements. Journal of Research in Reading, 29 (3), 318-333.

Bernard, M. L, Chaparro, B. S. and Russell, M. (2001). Examining Automatic Text Presentation For Small Screens. Human Factors & Ergonomics Society 45th Annual Meeting, 637-639.

Bird, S., Klein, E., and Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media Inc.

Brown, R. (1958). Words and Things: An introduction to language. New York: Free Press.

Bruce, G. (1977). Swedish Word Accents in Sentence Perspective. PhD Thesis, Lund: CWK Gleerup.

Buzan, T. (2000). The Mind Map Book. London: Penguin Books, 1996. ISBN 978-0452273221.

Card, S. K., Mackinlay, J. D. and Shneiderman, B. (1999). Readings in Information Visualization. San Francisco CA: Morgan Kaufmann.

Castelhano, M. and Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. Behaviour & Information Technology, 20(4), 237-247.

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.

Chung, S.T.L, Legge, G. E. and Cheung, S. (2003). Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. Vision Research, 44(7), 695-709.

Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel distributed-processing approaches. Psychological Review, 100, 589-608.

Cousin, C. & Vinckenbosch, M.C. (2000). Is Speed Reading For You?. Retrieved from http://www.readingsoft.com.

Dodge, R. (1900). Visual perceptions during eye movement. Psychological Review, VII, 454-465.

Fernández, E. (2007). How might rapid serial visual presentation of text affect the prosody projected implicitly during silent reading?. International Congress of ABRALIN.

Flesch, R. (1948). A New Readability Yardstick. Journal of Applied Psychology, 32, 221-233.

Fodor, J.D. (1998). Parsing to learn. Journal of Psycholinguistic Research, 27, 339-375.

Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. Percep. Psychophys. 8, 215-221.

Frazier, L., Carlson, K. & Clifton, C. Jr. (2006). Prosodic Phrasing is central to language comprehension. Trends in Cognitive Sciences, 10(6) 244-249.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin, 123, 71-99.

Huey, E.B. (1968). The psychology and pedagogy of reading. Cambridge, MA: MIT Press. (Originally published 1908).

Hunziker, Hans-Werner (2006). Im Auge des Lesers foveale und periphere Wahrnehmung: vom Buchstabieren zur Lesefreude (In the eye of the reader: foveal and peripheral perception - from letter recognition to the joy of reading). Zurich : Transmedia. ISBN 978-3-7266-0068-6.

Jackson, M.D. and McClelland, J.L. (1979). Processing determinants of reading speed. Journal of experimental psychology, 108, 151–181.

Juola, J. F. (1988). The use of computer displays to improve reading comprehension. Applied Cognitive Psychology, 3, 87-95.

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.

Kim, Y. and Oh, Y. (1996). Prediction of Prosodic Phrase Boundaries Considering Variable Speaking Rate. Intl. Conf. on Spoken Language Processing 96, 1996.

Koriat, A., Sheffer, L., and Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. Journal of Experimental Psychology: General, 131, 147–162.

Ladd, D. R. (1980). The Structure of Intonational Meaning. Indiana University Press, Bloomington.

Lee, K.C., et al. (2006). The Study of Suitable Conditions for Human Reading. The 9th Asian Symposium on Information Display, 319-322.

Marks, E. (2008). The Infojector: Maximizing textual information assimilation using an enhanced information interface. The16th CAS Biennial Conference on Science and Technology: Vehicles for Sustainable Economic Development in the Caribbean, 115-122.

Masson, M. E. J. (1983). Conceptual processing of text during skimming and rapid sequential reading. Memory and Cognition, 11, 262 - 274.

Mills, C.B. and Weldon, L.J. (1987). Reading text from computer screens. ACM Computing Surveys, 19(4), ACM Press.

Moore, D. and Zarbrucky, 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. and Maurutto, P. (1991). Reading and skimming from computer screens and books: The paperless office revisited? Behavior & Information Technology, 10, 257–266.

Nakatani, L. H., and Schaffer, J. A. (1978). Hearing "words" without words: Prosodic Cues for Word Perception. Journal of the Acoustical Society of America, 63(1), 234-245.

Nooteboom, S. G. (1997). The prosody of speech: Melody and rhythm. In W J. Hardcastle & J. Laver (eds.), The Handbook of Phonetic Sciences, Oxford: Blackwell Publishers, pp. 640-673.

Ohala, J. J. (1983). Cross-language use of pitch: an ethological view. Phonetica 40, 1-18.

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).

Öquist, G. (2001). Adaptive Rapid Serial Visual Presentation. Master's Thesis. Language Engineering Program, Department of Linguistics, Uppala University.

Paulson, L.J. & Goodman, K.S. (1999). Influential Studies in Eye-Movement Research. International Reading Association. Available at: http://www.readingonline.org (December 2001).

Pelli, D. G., & Tillman, K. A. (2007). Parts, wholes, and context in reading: A triple dissociation. PLoS One, 2, e680.

Pierrehumbert, J. & Hirschberg, J. (1990). "The meaning of intonation contours in the interpretation of discourse." In Plans and Intentions in Communication and Discourse, Eds. P. R. Cohen, J. Morgan, and M. E. Pollack, Cambridge: MIT Press, 271-311.

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(1), 106-117.

Rayner, K. and Pollatsek, A. (1989). The psychology of reading. Englewood Cliffs, New York: Prentice Hall.

Reichle, E.D., Rayner, K. and Pollatsek, A. (2000). Comparing the E-Z Reader Model to Other Models of Eye Movement Control in Reading. CogPrints electronic archive for papers. Available at: http://cogprints.soton.ac.uk/ (December 2001).

Rosenberger, T.M.G. (1998). Prosodic Font: The Space between the Spoken and the Written. Master's Thesis. Media Arts and Sciences Program, School of Architecture and Planning, MIT.

Rothkopf, E.Z. (1971). Incidental memory for location of information in text. Journal of Verbal Learning and Verbal Behavior, 10, 608-613.

Seidenberg, M. S. and McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. Psychological Review, 96, 523-568.

Selkirk, E., (2005). Comments on Intonational Phrasing in English. In Prosodies: Selected papers from the Phonetics and Phonology in Iberia Conference, 2003. (Phonetics and Phonology Series.) ed. S. Frota, M. Vigario, M. and J. Freitas, Mouton de Gruyter.

Shepherd, P. and Unsworth-Mitchel, G. (1999). The Speed Reading Course. Retrieved from http://www.geocities.com/dbmmextra/src_Peter_Shepherd_Gregory_Unsworth-Mitchell.pdf.

So, C. Y. J. and Chan, A.H.S. (2009). Design Factors in Dynamic Text Display. Engineering Letters, 16(3).

Stine, E. A. L. (1990). On-line processing of written text by younger and older adults. Psychology and Aging, 5, 68-78.

Strydom, J. and Du Plessis, S. (2003). Foundational Reading Skills. Retrieved from http://www.audiblox2000.com/learning_disabilities/reading_skills.htm

Taptagaporn, S. & Saito, S. (1990). How display polarity and lighting conditions affect the pupil size of VDT operators. Ergonomics, 33, (2), 201-208.

Taylor, E. S. (1965). Eye Movements in Reading: Facts and Fallacies. American Educational Research Journal, 2, (4), 187-202.

Townsend, D. J. (1983). Thematic processing in sentences and texts. Cognition, 13, 223-261.

Wilkenfeld, D. (1980). Prosodic Encoding in silent Reading. In CUNY Forum, 7, 8, 230-236.

Wright, P. and Lickorish, A. (1983). Proof-reading texts on screen and paper. Behaviour and Information Technology, 2, (3), 227-235.

Wright, P. and Lickorish, A. (1983). Colour Cues as Location Aids in Lengthy Texts on Screen and Paper. Behaviour and Information Technology, 7, (1), 11-30.