THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

Is WORD SHAPE STILL IN POOR SHAPE FOR THE RACE TO THE LEXICON?

By

JESSICA C. HILL

A Dissertation submitted to the

Department of Psychology

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Degree Awarded:

Summer Semester, 2010

The members of the committee approve the dissertation of Jessica C. Hill defended on July 1, 2010*.*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ralph Radach

Professor Directing Dissertation

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Kenn Apel

University Representative

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Frank Johnson

Committee Member

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Janet Kistner

Committee Member

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Jeanette Taylor

Committee Member

The Graduate School has verified and approved the above-named committee members.

I dedicate this to…

…God, who gave me this opportunity and blessed me while doing it.

…my mother, who died before I finished but who has never stopped believing in me.

…my husband, who never complained about the late nights or my working on weekends.

…my children, who have (mostly) waited patiently for Mommy to finish her degree

and get a “grownup job.”

…my friends, who have cheered me on even when they had no idea what I was talking about.

**TABLE OF CONTENTES**

LIST OF TABLESvii

LIST OF FIGURESviii

ABSTRACTix

1. INTRODUCTION1

1.1. Background: Vision, Reading, and Oculomotor Research1

1.2. Eye Movements in Reading3

1.2.1. Basic issues in oculomotor control3

1.2.2. The perceptual span and parafoveal processing4

1.2.3. Using the boundary paradigm to study parafoveal processing in reading6

1.2.4. A meta-analytical examination of the boundary paradigm7

1.3. The Role of Word Shape in Reading10

1.4. Letter Similarity in Parafoveal Processing14

2. EXPERIMENT ONE: WORD SHAPE AND LETTER SIMILARITY18

2.1. Hypotheses18

2.2. Method19

2.2.1. Participants19

2.2.2. Materials and Design20

2.2.3. Apparatus21

2.2.4. Procedure22

2.2.5. Dependent Variables22

2.3. Results23

2.3.1. Temporal Analyses (“When”)23

2.3.1. Spatial Analyses (“Where”)25

2.4. Discussion26

3. EXPERIMENT TWO: WORD SHAPE AND CONTEXTUAL CONSTRAINT29

3.1. Method30

3.1.1. Participants30

3.1.2. Materials and Design31

3.1.3. Apparatus32

3.1.4. Procedure32

3.1.5. Dependent Variables32

3.2. Results32

3.2.1. Temporal Analyses (“When”)32

3.2.3. Spatial Analyses (“Where”)34

3.3. Discussion36

4. GENERAL DISCUSSION37

4.1. Background and Results Summary37

4.2. Theoretical Implications38

4.3. Methodological Importance41

4.4. Conclusions41

5. APPENDIX A: TABLES43

6. APPENDIX B: FIGURES54

7. APPENDIX C: META-ANALYSIS METHOD AND RESULTS REPORT69

7.1. Methods70

7.1.1. Search Strategy and Study Selection70

7.1.2. Study Characteristics and Variables73

7.1.3. Data Evaluation73

7.1.4. Effect Size Derivation74

7.1.5. Analytic Strategy75

7.2. Results75

7.2.1. Fixed-Effects Analyses75

7.2.2. Random-Effects Analyses76

7.2.3. Categorical Analyses76

7.2.4. Continuous Analyses79

7.3. Discussion83

8. APPENDIX D: LETTER SIMILARITY METHOD AND RESULTS REPORT85

8.1. Method86

8.1.1. Participants86

8.1.2. Materials and Design86

8.1.3. Apparatus86

8.1.4. Procedure87

8.2. Results88

8.2.1. Nonlinear Reduction—Trial Effects88

8.2.2. Nonlinear Reduction—Item Effects90

8.2.3. Dimensionality Reduction92

8.3. Conclusion93

9. APPENDIX E: SAMPLE EXPERIMENTAL MATERIALS FROM EXPERIMENT 194

10. APPENDIX F: EXPERIMENTAL MATERIALS FROM EXPERIMENT 298

11. REFERENCES106

12. BIOGRAPHICAL SKETCH112

**LIST OF TABLES**

Table 1. Summary of visual assessments for Experiment 144

Table 2. Experiment 1 counterbalance conditions45

Table 3. Mean temporal processing times for Experiment 146

Table 4. Mean spatial processing times for Experiment 147

Table 5. Summary of reading skill scores for Experiment 148

Table 6. Experiment 2 counterbalance conditions49

Table 7. Mean temporal processing times for Experiment 250

Table 8. Mean spatial processing times for Experiment 251

Table 9. Confusability metric52

**LIST OF FIGURES**

Figure 1. The boundary technique55

Figure 2. Word decoding during reading56

Figure 3. The Glenmore model57

Figure 4. Letter similarity study stimulus presentation58

Figure 5. First fixation duration by shape by similarity59

Figure 6. Single fixation duration by shape by similarity60

Figure 7. Gaze duration by shape by similarity61

Figure 8. Landing position by shape by similarity62

Figure 9. Fixation probability by shape by similarity63

Figure 10. First fixation duration by shape by context64

Figure 11. Single fixation duration by shape by context65

Figure 12. Gaze duration by shape by context66

Figure 13. Landing position by shape by context67

Figure 14. Fixation probability by shape by context68

**ABSTRACT**

Current models of normal reading behavior emphasize not only the recognition and processing of the word being fixated (n) but also processing of the upcoming parafoveal word (n+1).  Gaze contingent displays employing the boundary paradigm often mask words in order to understand how much and what type of processing is completed on the parafoveal word before the word is fixated.  The word “beach” might be masked by a homophone (“beech”), a semantically related word (“shore”), or a string of random letters (“tcsok”), which can be either similar (b as compared to d) or dissimilar (b as compared to x) to the letters of the original word.  Mask use is inconsistent and ill defined across studies.  It is common for words to be considered masked with random letters if either all letters in the word are masked with random letters or the same first, second, or third letters of the word are used with random letters masking the rest of the word, thus leading to conflicting results in the body of the reading literature.

This work examines the role of word shape and letter discriminability in two experiments in the context of experimental reading research. The first experiment manipulated mask similarity, either similar or dissimilar letters, and whether word shape information was provided or denied, while controlling for lexical frequency (high) and word shape frequency (rare). Results indicate a significant effect of letter similarity on the parafoveal preview benefit, but there was not an equivalent result for word shape. The findings for dissimilar masks impacting parafoveal processing more than similar masks are critical in terms of the methodology of reading experiments. Few researchers control for mask similarity in their studies as there was no indication, before now, that such a control was necessary. The lack of findings for word shape was attributed to the lack of a manipulation that tapped into higher order cognition. A second experiment was conducted that included a contextual constraint manipulation, either high or normal contextual constraint, while controlling for lexical frequency (high), word shape frequency (rare), and mask similarity (dissimilar). The results indicated a significant preview benefit and predictability effect; however, there was no significant finding for word shape. Thus, it seems that word shape still seems to be in poor shape for the race to the lexicon. Results are discussed in terms of models of normal reading.

**1. INTRODUCTION**

* 1. **Background: Vision, Reading, and Oculomotor Research**

Understanding vision and perception has been a consistent human preoccupation since ancient times. Plato, an early Greek philosopher, eloquently described his admiration for the power of the visual system, which gave humankind the ability to understand the universe surrounding them (Wade, 1998). Although a distrust of information gathered through the senses prevented Plato and his contemporaries from more fully considering the nature of vision, later philosophers, such as Aristotle, and mathematicians, such as Euclid, were more willing to address evidence gathered through the senses, and thus address visual perception as a fundamental part of interacting with the world around them (Wade, 1998). These philosophical lines of inquiry were maintained throughout the years, which inspired later generations of scientists. Later scholars recognized the limitations and inconsistencies in historical arguments that had been addressed using mostly deductive reasoning and began developing techniques that served as the foundation for developing the scientific method. For instance, early Greeks assumed that vision resulted from “visual rays” being produced within the eye and being projected onto “luminous objects” in the environment (Smith, 1992, p. 190). A first century Arab scholar named Ibn al-Haytham, questioned this argument based on his observations of the moon. He instead suggested that light emanates from objects in the environment; such light can also be reflected off other objects and later enter the eye making vision possible. Keeping in mind the technical limitations restricting his investigation, al-Haytham made significant contributions to the development of the scientific method as well as the science of vision during his investigations (Smith, 1992).

The ability for humankind to express itself in writing has coexisted with the philosophical questions related to vision. Such importance was placed on ancient writings that social class structures were developed around reading as a skill (Huey, 1908). Huey explains that the desire to read and understand the written word was historically so widespread that more primitive peoples were seen eating books to try and gain the understanding that they had seen in individuals who were able to read. Those who were able to read were venerated. While, in retrospect, it is quite obvious to see the driving and long-lasting passion of humans to understand their world through vision and to maintain knowledge of that understanding through the written word, it has only been relatively recently that a systematic and rigorous examination of information acquisition from written text has been possible.

The first optical apparatus for continuously monitoring eye movements was used by Dodge & Cline (1901), at a time when tachistoscopic word recognition research was also just beginning. Before the advent of modern computers, researchers relied on devices, such as the tachistoscope, to control stimulus display durations. This instrument is essentially a slide projector attached to a shutter similar to that of a camera (Benschop, 1998) and allowed precise timing of short stimulus presentation times. As technology advanced, additional methods for recording eye movements were developed, providing the foundation of modern eye tracking research. In their review of eye movement measurement techniques, Young and Sheena (1975) describe how

the corneal bulge produces a virtual image of bright lights in the visual field and region. Because the radius of curvature of the cornea is less than that of the eye, the corneal reflex moves in the direction of eye movement, relative to the head. Because it only moves about half as far as the eye, it is displaced opposite to the eye movement relative to the optic axis or the center of the pupil (p. 317).

As early as 1935, it was possible to use a film camera to record the reflection of light from the cornea of the eye (e.g., Buswell, 1935). Such systems were complicated and uncomfortable to use as they often required special lighting and bite bars to function appropriately. Additionally, it was difficult to analyze eye movements when any head movement was present. With the advance of miniature television systems and cameras, it was possible to develop more portable systems. Over the next 50 years, systems and analysis improved, yielding the advance systems we see today: remote trackers still using the corneal reflection to help determine eye position with increased precision and resolution.

As the technology underlying reading research has advanced, the methods used to investigate reading have kept pace. Rayner (1978) describes three “eras” of eye movement research. During the first era, the devices and techniques used to measure eye movements were rudimentary and crude. Despite the limitations of the technology of the time, reading researchers were able to make foundational observations (e.g., saccade latency, saccadic suppression) that remain relevant today. The “second era” of eye movement research occurred during the time of the behaviorists. Because of such influence, most investigations focused on “observable surface aspects” of eye movement behavior (Starr & Rayner, 2001, p. 156). Connections to the cognitive processes that underlie reading were not possible due to technological limitations as well as a scientific environment inhospitable to deeper questions related to cognition. Since then, there has been a relative ‘boom’ of research utilizing reading as a means to investigate information processing, attention, and perception, which constitutes the “third era” of eye movement research. This revival can be attributed to the development of smaller and more accessible computers as well as being the ripening fruits of the “cognitive revolution” (for a historical account see Miller, 2003). During this era, studies using precise measurements of oculomotor behavior as a test of psycholinguistic hypotheses became more and more prominent.

Radach and Kennedy (2004) have recently suggested that with the advent of computational modeling the field has moved into a fourth era. The fruits of the fourth era include a much more detailed and comprehensive understanding of information processing and eye movement control during reading, leading to a number of computational models of the reading process. These models differ, among other aspects, in the degree to which the interplay of linguistic processing and oculomotor control is dominated by cognitive vs. visuomotor constraints (see Radach, Reilly & Inhoff, 2007, for a detailed discussion).

**1.2 Eye Movements in Reading**

**1.2. 1. Basic issues in oculomotor control.** Humans have a narrow range of high-acuity central vision that drops off relatively rapidly into the periphery. In order to compensate for the narrow range of high-acuity fovealvision, the human eye has become highly mobile. As such, the eyes are able to make a variety of movements, the most common movement being the *saccade*, which is a quick, jumpy movement. A specific example of saccadic eye movements would be those encountered during reading; as you move from one word to another while reading this text, your eyes do not move smoothly across the page, instead they jump between points of saliency. As a reader moves across the text, they tend to move their eyes a distance of approximately seven to eight characters, which closely matches the average length of words (Findlay & Gilchrist, 2001). Critical issues in the successful and efficient navigation of text are twofold: (1) how the oculomotor system determines *where* to move the eyes next and (2) *when* the movements should be made.

The decision where to move the eyes next appears to be a relatively simple one as there are only a few options for saccade destinations while reading: the same word again, the previous word, the next word, or perhaps the word beyond the next word. These alternatives already account for the vast majority of local fixation patterns observed in reading (Radach & Kennedy, 2004). Some textual features help guide where a reader’s saccades land as they progress through text. Many modern languages use spaces to define the boundary between words, thus allowing a reader to perceive the length of a word accurately, yielding information about the optimal goal for the landing point of the next saccade. Typically, saccades into a word are intended to land at the center of the word or just to the left of the center, at a location called the *optimal viewing position* (Vitu, O’Regan, & Mittau, 1990). However, the eyes do not always land at this optimal position due to the contribution of several relevant factors, including the length and eccentricity of the target, resulting in landing at *preferred landing positions* about halfway between the word beginning and center (see e.g., Radach & McConkie, 1998, for a detailed discussion).

While there is general agreement about where the eyes move while reading, the question of how the oculomotor system determines when to move the eyes is still debated. Contributing to this debate is the relatively complex nature of the “when” decision. Not only does that decision include understanding the time course of cognitive processing, it also necessitates understanding of what information is used during the course of a fixation (Inhoff, Eiter & Radach, 2005). Recent evidence from ERP and basic attentional studies have revealed some basic limits for when eye movements are programmed and the amount of processing that must be completed; specifically, the majority of lexical processing must be completed within 100-150 ms to allow eye movements to be made “intelligently” (Radach & Kennedy, 2004, p. 11). Matters become more complex when reading in the context of a complete sentence is considered. For example, as Inhoff and Rayner (1986) demonstrated, processing time on a nearby word can be reduced by as much as 30 ms when that word is subsequently fixated.

**1.2.2. The perceptual span and parafoveal processing.** In addition to the issue of oculomotor control discussed above, a second major question of cognitive reading research concern the amount of information acquired during each fixation. Related to this are the issues of how this information is subsequently processed and how information from successive fixations is combined. These questions are of central importance for a comprehensive understanding of the reading process, but they are also interesting for cognitive science as a whole, because questions on the nature of attention allocation or serial vs. parallel information processing concern the human mind as whole.

Because word length is a significant determinant of landing position within a word, readers must acquire length information from the parafovea (a region in which text is still discriminable but outside the range of high acuity vision) before making a saccade. Just as length information can be acquired parafoveally, lexical information is also extracted. As such, readers are able to “preview” the upcoming word to a limited extent, allowing initial processing on (minimally) the first few letters of the upcoming word. This early processing is vital, as some letter combinations are more common (e.g., “al” is more common than “rh”) indicating potential ease or difficulty of identifying the upcoming word (Radach, Inhoff & Heller, 2004). For example, if an upcoming word is particularly difficult, a shorter saccade meant to allow a refixation on the same word may be programmed instead of a longer saccade into the next word allowing more time to process the difficult word.

The modern eras of experimental reading research discussed above have brought several interesting developments relevant to experimental reading research. In his 1978 review, Rayner already emphasized the role of progressively smaller and faster computers allowing the “on-line” recording of eye movements and the ease of analysis associated with those computers. With increased computer power, the display of reading material also grew more sophisticated. Instead of being able to present words or sentences on index cards or sheets of paper, stimuli could be presented on computer monitors. Precise measurement and control over the display of stimuli let researchers investigate cognition in a way that was impossible in the first and second eras. As such, elegant techniques, such as the moving window (McConkie & Rayner, 1975), the moving mask (Rayner & Bertera, 1979), and the boundary techniques (Rayner, 1975), were developed to provide fine-grained analyses of visual and linguistic processing. These techniques employ display changes that are contingent upon the position of the reader’s eyes, often called *gaze-contingent displays*.

Figure 1 illustrates the boundary technique. As the reader’s eyes cross an invisible boundary on the screen there is a change in the display. Such changes are implemented within a few milliseconds during a saccade and therefore invisible to the reader. By replacing the word of interest with a mask (e.g., river with stream), it is possible to understand how readers process a certain kind of information by denying them a normal preview of the parafoveal word. In most cases the critical dependent measure is an increase in viewing duration measures, such as gaze duration, when fixating a target word that was previously masked. The idea is that as a normal preview of the word was denied, processing that would have taken place on a prior fixation needs be done now, leading to the observed increase. Accordingly, the difference in viewing duration for condition with and without parafoveal mask is commonly referred to as ‘preview benefit’. This methodological elegance and potent inferential qualities make the boundary technique one of the most powerful tools in experimental reading research to date.

**1. 2. 3. Using the boundary paradigm to study parafoveal processing in reading.** There are many variations of the boundary technique (see Rayner, 1998); the classic technique, however, involves creating an invisible boundary at a fixed position within a sentence, usually at the end of pre-target word and masking the target word. Many different types of masks can be used to cover the target word (see Figure 2) including semantic, phonological, or orthographic, thus giving different information about the timeline of parafoveal processing.

The consensus in the field is that all types of information depicted in Figure 2 can be denied during parafoveal processing using a random letter mask and that the difference to a normal preview condition indicates the existence and relative importance of the respective type of parafoveal processing. In some cases masks are used that carry partial information about the upcoming target word, leading to viewing durations on the target that are between the full preview and full mask conditions. As an example, Pollatsek, Lesch, Morris & Rayner (1992) used orthographic and phonological (homophone) preview conditions, showing that phonological information can be used very early during word processing. It is interesting to note that, in general, different previews may affect different stages of processing, such that an orthographically coherent mask will aid early processing, whereas a semantic mask may ‘prime’ much later processing (Ferrand & Grainger, 1993; Grainger & Jacobs, 1999).

Despite the increasingly frequent use of the boundary technique by reading researchers, there are no standardized procedures for the location of the boundary, the creation of the masking material, and other methodological conditions during experimental sessions (Radach & Kennedy, 2004). This causes some difficulty in the interpretation and replication of results across studies and especially across labs. To develop a more comprehensive and systematic picture of these problems than was available before, a meta-analysis on the use of the boundary techniques was carried out.

**1.2.4. A meta-analytical examination of the boundary paradigm.** Meta-analysis is becoming a more common tool in psychology as well as other fields; it is, however, a rarely used tool in the experimental reading literature. To the best of our knowledge, only one meta-analysis, by Brysbaert, Drieghe, and Vitu (2005), has been conducted to date. This analysis asked and answered deep theoretical questions about models of reading in the context of fixation probability, the likelihood to fixate (or “skip”) a certain word. Models of eye movement control in reading are quite similar on the surface, including descriptions of how the eyes move across the page while the text is processed. However, a critical difference between the models is what event triggers the eyes to move. A cluster of models, which rely on strict serial attention (e.g., E-Z Reader; Pollatsek, Reichle, & Rayner, 2006), require that a certain amount of lexical processing of a word must be completed before attention and the eyes can move on to the next word. Other models, such as those assuming processing gradients (e.g., Glenmore, Reilly & Radach, 2006; and SWIFT, Engbert, Nuthmann, Richter, & Kliegl, 2005), do not demand as strict a relationship between lexical processing and eye movements. Here, the range of visual processing extends over the entire perceptual span and can include several words in parallel (see Radach, Reilly & Inhoff, 2007, for a detailed discussion).

Brysbaert and colleagues (2005) looked at the relative importance of word length, as a proxy for low level visual factors, and word frequency, which was a proxy for cognitive influence. Their results indicated that word length, and therefore visuomotor constraints, accounted for a much larger part of the variance found in all the word skipping studies than did word frequency. Further, they conclude that “educated guesses” are what is driving the next progressive saccade. Often, these educated guesses are based on incomplete information, thus resulting in erroneous eye movements skipping the intended word. Typically, these errors are corrected as soon as they are detected by either the execution of a regression to the word that was skipped or by the implementation of a longer fixation on the word currently fixated.

From a formal standpoint, the meta-analysis conducted by Brysbaert et al. (2005) was somewhat lacking in detail. However, the meaningful conclusions extracted from this frequently cited work make up for its technical limitations. Somewhat more importantly for the current synthesis, the Brysbaert and colleagues (2005) work demonstrates the applicability of meta-analytic techniques to the experimental reading research literature.

In line with the goals of the dissertation work presented below, a small-scale meta-analysis, based on 6 published studies including 11 experiments and 48 critical comparisons, was conducted to analyze the masking literature comprehensively as well as to examine the extent to which mask information interferes with lexical processing. The intention was to examine conditions of parafoveal information acquisition that might be significant determinants of performance in a large number of studies. Our two main candidates were parafoveal word shape and the similarity of the letter within parafoveal mask and target word. The focus on word shape was driven by theoretical considerations (see below for a detailed discussion), while our interest in letter similarity was more methodological, as we expected mask-target similarity to be major contributor to heterogeneous results in prior studies.

The meta-analysis was conducted in order to investigate mask type as a proxy for word shape violations and to determine how much word shape violation interferes with lexical processing during continuous sentence reading, thus providing an essential foundation from conflicting results from the literature for an explicit experimental examination of word shape. (See Appendix C for a detailed report on the methods and results of the meta-analysis.) It was expected that the effect size data would be best accounted for by a random-effects model as it is highly unlikely that much sampling error would occur considering the high accuracy and precision of modern eye trackers. Because the mixed-model provided the best fit, a conclusion about the predictors being inadequate must be made.

The amount of data excluded from a study was a significant predictor in the mixed model. Those studies with the highest amounts of excluded data (min = 7%, max = 38%, mean = 17%, SD = 8%) were also those with the lowest effect sizes (mean effect size for 38% excluded = -0.09; mean effect size for 7% excluded = 0.54). One explanation is that the trials being excluded were those showing an effect. Alternatively, there could have been a problem with the method (e.g., too difficult or challenging) that resulted in such a large amount of data being discarded. Whatever the cause, it is clear that a large amount of discarded trials is not advantageous in reading research.

Considering the amount of technological innovations associated with eye tracker development, there have been no large technological or methodological changes that can be attributed to the Year (i.e., the year the study was published) predictor. Interestingly, Year and Rayner (i.e., whether Rayner, a strong proponent of a serial attention model, is present as an author on the paper) are highly related. From examining scatterplots of the data, it seems as if a pattern of results from one lab was followed by responses from the Rayner lab, leaving the Rayner lab articles as the most recent ones, thus closely tying Year and Rayner. Furthermore, the majority of the effect sizes when Rayner is present as an author are also significantly lower than when Rayner is not present. While being a relatively poor theoretical proxy, this predictor’s significance provides evidence for the necessity of coding for theoretical viewpoint in any further meta-analysis of the reading literature, especially when parafoveal processing is being addressed.

The original question of mask type and word shape violations could not be addressed as expected. Surprisingly, the studies synthesized in the meta-analysis never included a condition in which a similar random letter mask was employed in comparison to a dissimilar random letter mask (see Section 1.4 for a theoretical discussion of mask similarity). As there were no similar and dissimilar random letter masks to compare, it can be maintained as a reasonable hypothesis that length and random letter masks do interfere with lexical processing especially in the early stages of processing as reflected by increases in first fixation duration when a mask is present. Word length manipulations were included in the category of word shape manipulations because changing the length of a word by either adding or removing letters also changes the shape of the word. Additionally, creating parafoveal masks out of random letters also has the potential to change the shape of a word should those masks be generated in a truly random manner. As these types of masks do change the shape of the word, the results imply that word shape violations are contributing to the effects seen (see Appendix C). In conclusion, as it is impossible to conclude from existing data precisely in what way word shape violations are affecting continuous reading, empirical research on this question is clearly needed.

While there is consensus about what constitutes a semantic mask (e.g., *river* covering the word *stream*) or a phonological mask (e.g., *beech* covering the word *beach*), there is less consensus regarding the nature of random letter masks. Although it may seem intuitive to assume that a random letter mask indicates all letters of the target word will be masked with randomly generated letters, this is not always the case. A random letter mask may be defined in several ways. First, it may be defined as a mask entirely composed of random letters (e.g., basket/ulxrop). Alternatively, it may also be defined as random letters covering all but the initial letter of the target word (e.g., basket/blxrop). Finally, when several letters of the target word are retained at the beginning of the mask (e.g., basket/ baxrop/basrop), the label of ‘random letter mask’ may also be applied. To complicate the issue further, random letter masks may use letters that are either similar (e.g., b may be masked by d) or dissimilar (e.g., b may be masked by x) to the target word.

Additionally, since this type of masking can also disrupt the processing of word shape (a hotly contested issue in the experimental reading literature, see below) or may even be confounded with word shape (Paap, Newsome, & Noel, 1984), further difficulties in interpretation arise. Considering the amount of preprocessing that occurs before a word is fixated, it has been proposed that the “shape” of an upcoming word can help predict what the word will be (Besner, 1989; Faust, Babkoff, & Avidor-Reiss, 2000; Reichle, Pollatsek, & Rayner, 2006; Rudnicky & Kolers, 1984; see Lete & Pynte, 2003, for a recent discussion). Words composed of lowercase letters can have a definite “shape” created by the combination between ascending (e.g., t), descending (e.g., g), and neutral (e.g., o) features. The words “little” and “pepper” possess distinctive word shape whereas “seems” has a neutral shape. Shape information is sometimes described as being processed parafoveally and in a prelexical fashion (Reichle et al., 2006). Specifically, “the low-spatial frequency information that is available from parafoveal and peripheral vision provides information about word length and shape, and the boundaries between words” (Reichle et al., 2006, p. 6).

**1.3. The Role of Word Shape in Reading**

When comparing results across studies, it becomes apparent that masks of different types significantly affect lexical processing in the parafovea. Similar to factors like word length and lexical frequency, word shape is a variable that may play a role in parafoveal processing and is perhaps now often confounded with other influences. In the meta-analysis mentioned above, there were significant amounts of variation unaccounted for by the predictor variables used, thus implying that additional factors are affecting reading sentences with parafoveally-masked words present. It is possible that shape violations of upcoming words may be the source of this extra variation. Additionally, there is a lack of research focused on understanding the methodological differences in using different types of masks and the effects those can have on such basic measures as first fixation duration and gaze duration. This is concerning due to the conflicting reports of word shape effects in the literature, and the potentially vital role of word shape in reducing the number of candidate words during lexical processing (Walker, 1987).

Even though the idea that word identification can be facilitated by the nature of its shape seems intuitive, there have been contradictory findings accompanied by a heated debate in the reading literature. This debate is often characterized in reference to Paap, Newsome, and Noel (1984) who examined the contribution of word shape to reading by addressing the “holistic” nature of shape processing. It was hypothesized that rare word shapes were more informative than common word shapes because rare shapes had fewer word candidates than common words. Both the shape and frequency of the target words were manipulated, and the amount of ascenders and descenders within the word as well as letters within the word (e.g., *cellar*, a rare shape, and *recall*, a common shape) were also matched (Paap et al., 1984, p. 422). Because the authors were unable to find evidence that rare word shapes were identified faster than common word shapes, they concluded that word shape information was lost early in the identification process. Critics point out that a weakness associated with the Paap et al. (1984) work concerns the stimuli used in the experiment. Instead of using an existing computer font, Paap et al. (1984) developed a 5 x 5 matrix for the body of each neutral letter leaving only two “dots” to distinguish ascenders

and descenders from the neutral body (Walker, 1987). This type of “impoverished stimuli” led to less word shape information being available parafoveally than would normally be encountered in regular reading.

Determining how much information is obtained parafoveally as well as when exactly such information is obtained, are critical questions to reading researchers. The evidence and argument provided by Paap et al. (1984) shaped the foundation of most models of normal reading; specifically, despite E-Z Reader’s (Reichle et al., 2006) foggy description of word shape as a type of low-level visual information gathered parafoveally, there is not a known model that explicitly includes the processing of word shape in the computational portion of the model. As a result, most theories posit that word identification relies on individual letter identification within a word (Lete & Pynte, 2003). From a computational standpoint, this assumption makes sense. Considering that “the average skilled reader knows somewhere between 30,000 and 50,000 words, the computation involved in solving shape invariance for each word is going to be a lot more costly than solving it for each letter of the alphabet” (Grainger, 2007, p. 3). A good example depicting the prioritization of letter identification in a computational model of normal reading is the Glenmore model (Reilly & Radach, 2006; see Figure 3). This model contains a word identification unit in which letter information is extracted within a window of processing (the input vector), corresponding to the current perceptual span. Based on the spatial separation between words, letters are grouped into clusters and the corresponding activation is fed further into word nodes. The processing of letter information within the input proceeds with a rate determined by eccentricity relative to the current fixation. In this process, no consideration is given to a potential role of word shape on either the level of letter or word processing. However, the exclusion of word shape processing by theorists and computational modelers might have been premature, as recent empirical evidence has challenged Paap et al.’s (1984) conclusions and thus the conceptualization of word shape processing during reading.

Common ways to distort word shape include alternating case text (e.g., dRaGoN v. dragon), alternating size text (e.g., dragon v. dragon), and using entirely uppercase letters (e.g., DRAGON v. dragon). In order to examine the contribution of word shape in lexical decision tasks, Perea and Rosa (2002) conducted two experiments using alternating size (Experiment 1) and entirely uppercase letters (Experiment 2). The results for experiment one indicate an interaction between font size type (same size v. alternating size) and lexical frequency. High frequency words showed a smaller effect of font size type than low frequency words. The authors suggest that the larger letters within the word were inappropriately grouped into units and disrupted word identification. In order to avoid problems associated with alternating size text, a second experiment was conducted using upper- or lowercase stimuli of the same size. Results from experiment two were similar to those of experiment one, including the interaction between case type and lexical frequency. Perea and Rosa (2002) suggest that their findings do not support the idea that high frequency words can be accessed via a “fast word-level channel” (p. 20) thus denying the role of word shape in lexical decision tasks. Despite this claim, the authors emphasize that word shape (or “visual familiarity”) does influence lexical decision performance and that this influence seems to occur rather late in the process of word identification, which they attribute to top-down/bottom-up stimulus mismatch (Besner, 1989).

While distorting text is a common way to disrupt word shape information, it is also relatively ecologically invalid. As readers navigate the textual world, it is uncommon (although not impossible) to encounter situations in which significantly distorted text must be decoded (e.g., entering a confirmation code into a website). As an empirical tool, such changes to font size or case can provide valuable insights; however, it also creates results that can be difficult to generalize. An equivalent situation would be equating the results of a basic visual search task to an applied driving situation; while the visual search task may provide important insights into the nature of attention, working memory, and driving simulator performance, it is difficult to directly generalize to “real world” behavior of driving on city streets.

Given these considerations, the work of Lete and Pynte (2003) becomes particularly relevant as they also investigated word shape and lexical frequency, yet they did not distort the text of their stimuli. In a series of single word reading experiments, these authors investigated the roles of both word shape and lexical frequency in the context of lexical decision and naming tasks while also exploring whether word recognition is based on “supraletter features” (i.e., word shapes are stored as units) or on individual letter features. The authors found that word shape information facilitated different types of processing based on task demands. Specifically, priming lexical information, as in the lexical decision task, showed a facilitation for words with the most common shape (i.e., many lexical candidates sharing the same overall shape—e.g., the word *recall*). A different pattern of results emerged for the naming task; in particular, rare shaped words (e.g., the word *cellar*) were facilitated by priming in this case. Lete and Pynte (2003) conclude that word shape does play a role in word recognition, but that role does not necessarily include the outline of a word, rather it includes a variety of features made up of patterns of single or multiple letters. While the authors did not conclude that word shape processing was an individual stage or unit in the word recognition process, they did emphasize the assistance such information can provide in the parsing of written language. Further, word shape was described as a “complex variable” (p. 941) and such complexity highlights the varied ways in which word shape may affect lexical access and the underlying representations of words.

Even though Lete and Pynte (2003) did not rely on distorted text, their results are still somewhat limited due to the isolation of words required for the tasks employed. In addition, their experiments were conducted in French and contained an extra category of word shape: marked letters (i.e., those with accents). An alternative explanation to the reported results could be that the addition of marking/accenting creates a greater possibility for rare word shapes, and thus a greater chance to discover an effect. If this is the case, a more stringent test of the role of word shape would be to use words in a language that is unaccented and also in a situation where the words are not isolated.

Further, a critical confounding variable in Paap et al.’s (1984) first experiment was the confusability of the letters they used to mask the word. Confusability (or: similarity) is a somewhat more subtle variable, which shows a lack of reported information in most published studies similar to that of word shape information. Frequently, studies report the use of random letter masks; yet, they fail to report how similar or confusable the masks are. This is particularly troubling because similar masks often have overlap with the target word (i.e., they share letters; e.g., bench v. berok). While there have been a few attempts at creating matrices of confusability (e.g., Bouma, 1971; Courrieu, Farioli, & Grainger, 2004; Jacobs, Nazir, & Heller, 1989), this information is underutilized in the experimental reading literature.

**1.4. Letter Similarity in Parafoveal Processing**

As mentioned above, a critical methodological issue related to masking using the boundary technique (Rayner, 1975) is the similarity (or confusability) of the letters used to mask the target word (see Appendix E for an example of a similar letter mask). Early studies examining the confusability of letters degraded their stimuli by increasing the viewing distance or otherwise making the letter stimulus more difficult to perceive (Bouma, 1971) in order to obtain adequate cell values for a confusion matrix (Courrieu et al., 2004). For example, Bouma (1971) used a tachistoscope to present participants with lowercase alphabet letters in Courier 10 font at varying distances and retinal eccentricities. Results were presented in a confusion matrix in which errors in letter identification are recorded (Grainger, Rey, & Dufau, 2008) and also into salient perceptual groupings produced from the perceptual distance mappings (see, for example, Figure 4, Bouma, 1971, p. 464). In addition to the confusion matrix, Bouma (1971) assessed distinctive features of letters, which yielded sixteen factors including the height of the letter, whether it had a gap to the left of the right, and oblique angles within the letter.

More recent work (Courrieu et al., 2004; Jacobs et al., 1989) focuses on non-degraded stimuli in a way that is more immediately applicable to reading research. For example, Jacobs and colleagues (1989) investigated letter features utilized in recognition with an additional purpose of creating a confusion matrix to index letter similarity. Participants were asked to fixate a central uppercase letter which served as a cue. After fixating the cue, it was removed and two lower case letters were presented 6 degrees of visual angle (18 character spaces) from the central fixation point. A saccade was then made to the lowercase letter matching the previously appearing central uppercase cue. Saccade latencies and directional error rates serves as performance measures and were used to construct similarity matrices. Additionally, the authors conducted a feature analysis to determine both the local (e.g., “lower gap”) and global (e.g., “circular envelope”) letter features that may contribute to letter similarity (cf. Jacobs et al., 1989, p. 101). Results indicated that the ratio between the height and width of the letters were the primary discrimination strategy used by participants. A critical contribution of this paper is the use of saccade latencies as a metric for letter similarity.

While the Bouma (1971) and Jacobs et al. (1989) studies provided a foundation for understanding letter perception, there were significant limitations to the generalizations that can be made from their results. Bouma (1971) presented letters in isolation and at varying distances, situations which occur only rarely in normal reading. Jacobs et al. (1989) presented letters to be compared to a central cue, however, they used a font manufactured for reading studies on monitors that unfortunately minimized the contribution of ascending features (e.g.,t) and descending features (e.g., y) (see Jacobs et al., 1989, p. 95, for examples). Additionally, there was a confound between the shape of the central cue and the shape of the letter to be identified. Specifically, the capital letter ‘A’ is quite distinct from the lowercase letter ‘a’ whereas the capital letter ‘S’ is very similar to the lowercase letter ‘s,’ thus it is possible that priming effects occurred.

The later work by Courrieu, Farioli, and Grainger (2004) improved upon the methods used by Bouma (1971) and Jacobs and colleagues (1989) by presenting letters in pairs and also presenting only lowercase letters in a commonly used computer font. Additionally, the authors included lateral masking of their stimuli. This is a major improvement as letters within words laterally mask their neighbors. Laterally masked stimuli are more difficult to recognize and take longer to be perceived (Huckauf et al., 1999). In the Courrieu et al. (2004) work, the authors examined foveal letter confusability under conditions of lateral masking using a go-nogo paradigm. After transforming manual reaction times using monotonic embedding, the authors used Euclidian distances in a multidimensional scaling analysis that yielded 25 meaningful factors related to similarity and contrasting classes somewhat similar to those reported by Bouma (1971). Interestingly, Courrieu and colleagues (2004) qualify their results by noting that “the visual similarity of letters will depend to some extent on the particular type font tested” (p. 910). The typeface utilized in their study was Arial 12-point, bold font, which may partially explain some of the differences seen in the letter feature classes reported by Bouma (1971) and Jacobs et al. (1989).

Unfortunately, the studies reviewed above have several limitations in their application to experimental reading research, especially on parafoveal processing. For example, Courrieu and colleagues (2004) used a foveal discrimination task requiring manual responses, which is quite far removed from a normal sentence reading situation. Jacobs et al. (1989) utilized eye movements as the major dependent measure, however, the positioning of their stimuli at 6 degrees of visual angle from the central cue is also far greater a distance than normally found in single sentence reading. Also, the font type and size (e.g., Courier New, 12-point) used in many typical reading experiments is not addressed by the previous work. Many researchers use non-proportional fonts with equal width assigned to all letters in order to control the visual angle of the stimuli they have created for their studies. Given this state of affairs, it was decided that for the planned dissertation work to be successful and significantly advance the field, a novel letter similarity metric needed to be developed (see Appendix D for a detailed report on the methods and results of pilot work to create such a metric). Therefore, to assist in creating stimuli for the dissertation project, a task was designed that maintained the most advantageous features of both Courrieu et al. (2004) and Jacobs et al. (1989) as mentioned above.

As seen in Figure 4, a combination of the methods used in Courrieu et al. (2004) and Jacobs et al. (1989) was employed. Specifically, the concept of using saccade latencies and requiring an eye movement response from participants was imported from Jacobs et al. (1989) while avoiding the potential confound of using upper case central cues. It is important to note that the distance between the central cue and the target letters was much smaller than that in Jacobs et al. (1989); instead of using the 6 degree distance, a shorter distance of slightly less than 2 degrees was used as that is within the range of typical saccades seen in a normal sentence reading task. Lateral masking was imported from Courrieu et al. (2004) as to approximate the lateral masking of letters within a word while reading.

The results were analyzed using a nonlinear reduction techniques introduced by Hinton (2005) in order to establish a more fine grained understanding of letter similarity. The analysis demonstrated the viability of using saccade latency in a cued target-distractor paradigm as a measure of letter confusability. It has provided a rich picture of visual letter processing, suggesting features of the process that need to be focused on in the construction of a more realistic model of the low-level visual aspects of reading. For example, openness versus closedness of a letter and their spatial frequency appear to play roles. Moreover, ‘closedness’ of the type in the letter ‘**o**’ seems to have a distinct advantage over other patterns closed when presented in the parafovea. This is evidenced by the ‘**og**’ versus ‘**go**’ asymmetry, for example (See Appendix D Figure 8.3). Most of the highly confusable pairs followed what would be expected from intuition, such as ‘**pq**’ or ‘**bd**’. However, there were some notable exceptions to that general rule, such as ‘**cx**’ or ‘**sh**’. The most likely explanation for this counterintuitive pairings is that the parafoveal lateral masking employed made letter discrimination more difficult in cases where letters had angled components, like the diagonal lines composing the ‘**x**’ or the slanting portion of the ‘**s**’. The resulting metric was employed in the creation of stimuli for the current experiments. It was essential to create this pairwise similarity table using the specific font and size of the letters that would be used in the current reading experiments. Thus, it was possible to insure that similar and dissimilar letter features were specific to the stimuli produced for the dissertation.

**2. EXPERIMENT ONE: WORD SHAPE AND LETTER SIMILARITY**

Based on the considerations outlined above we designed an English sentence-reading study manipulating both word shape and the similarity of the random letter mask without overlap between the target word and the similar mask. Further, it is desirable to determine the effects of the similarity of letter masks and word shape on the parafoveal preview benefit. Despite sparse reports of word shape and similarity, to the best of my knowledge, never before has word shape information and mask similarity been manipulated in conjunction experimentally. Thus, both were manipulated concurrently in this experiment, avoiding the potential for similarity to become a confounding variable as seen in Paap et al. (1984).

**2.1. Hypotheses**

Based on Lete and Pynte (2003), an effect of word shape is expected to be found. As such, it is hypothesized that rare word shapes will be the most informative, and thus will show less disruption in lexical processing even when masked. Further, it is assumed that similar random letter masks will be less disruptive to reading than dissimilar random letter masks. It is also expected that there will be an interaction between word shape and mask similarity; rare word shapes and similar letter masks will be the least disruptive to reading of all the experimental conditions. Despite differences in shape features due to language, similar results to Lete and Pynte (2003) in a continuous sentence-reading task are expected to be found.

Information on changes in the parafoveal preview benefit due to the similarity of random letter masks is often unavailable in the experimental reading research literature. Such information is critical to understanding how a mask interferes with lexical processing. Consider the nature of a dissimilar random letter mask composed of letters dissimilar to the target word. In such a case, early lexical processing, such as orthographic processing, would be disturbed due to the word recognition system being unable to cope with the randomness of the input being provided. Alternatively, a random letter masked composed of letters similar to or easily confusable with the target word would not cause as much disturbance, due to overlapping letter features, during initial processing, instead the interruption occurs later. In the former case, it is hypothesized that the mask is less similar and potentially less word like, stopping processing early. The penalty for having to postpone or restart processing should be relatively larger as compared to the latter case in which much further processing has occurred and any further processing can benefit from that which is already done (e.g., narrowing down potential word candidates).

Theoretically, this hypothesis is quite interesting. The newest modification of the E-Z Reader model (Pollatsek et al., 2006) discusses two pivotal, and potentially mutually exclusive, assumptions related to saccadic suppression while reading. In a recent attempt to simulate the parafoveal preview benefit in reading, Pollatsek and colleagues (2006) examined a hypothetical preview condition by not allowing lexical processing on the target word to start before the target word was first fixated. Apparently, the underlying assumption is that a random letter mask prevents parafoveal linguistic processing at all. For the present study, the prediction follows that different types of random letter masks should not make a difference. In contrast, the Glenmore model (Reilly & Radach, 2006) would treat a nonword letter string like an extremely low frequency word. This model predicts that initial letter activation based on parafoveal information would occur and therefore the disruption in processing the critical string would be a function of letter similarity. More specifically, following the argument made above and assuming that parafoveal information supports an early phase of letter and word processing, dissimilar letters should have a more detrimental effect.

**2.2. Method**

**2.2.1. Participants.** Thirty five Florida State University college undergraduates from General Psychology courses that include an option for experiment participation as a part of course credit and community members were recruited for participation in this experiment. All participants had normal or corrected-to-normal vision (see Table 1 for a summary of vision scores). Participants were either native speakers or had spoken and read English for at least 10 years. The sample was composed of 52% females with a mean age of 22 years. No information on race or ethnicity was recorded as there is no reason to assume differences in oculomotor behavior based on either race or ethnicity.

A power analysis was conducted using the G\*power program of Faul, Erdfelder, Lang, and Buchner (2007) in order to determine the necessary sample size for Experiment 1. Using information from previous studies, a medium (approximately .5) to low (.2 to .3) effect size was assumed for the proposed experiment. Results of the power analysis show that 32 participants are required to achieve conservative power of .95 of this magnitude with an alpha level of .05. In order to have an equal number of participants in each counterbalance condition (see Materials and Design section below), this number was increased to 35 participants.

**2.2.2. Materials and Design.** A within-subjects 2 x 2(similar or dissimilar random letter masks x word shape information provided or denied) was used in this experiment with a legal preview control condition. One hundred single line sentences of no more than 83 characters were constructed, each containing a manipulated critical target word. All target words were six to eight words in length to decrease the probability that they will be skipped (Brysbaert et al., 2005). Target words were controlled for lexical frequency (high) and word shape frequency (rare) in order to maximize the chances of discovering an effect of word shape. Lexical frequency was determined using the CELEX database of word frequency (Baayen, Piepenbrock, & van Rijn, 1993). Word shape frequency was established by analyzing the CELEX database for shape information using the Walker (1987) protocol. Words chosen as target words were in the bottom 10% of word shape frequency categories. These words had a mean word shape frequency of 11.8 (SD = 8.2, minimum = 1.0, maximum = 29.0) whereas the mean word shape frequency of the entire database was 131.5 (SD = 149.5, minimum = 1.0, maximum = 666.0). All potential target words used in these experiments had no more than 4 other words sharing the same shape, although the majority of the words were unique in shape. Specifically, out of 299 potential target words, 35 words shared their shape with one other word, 6 words shared their shape with two other words, and one word shared its shape with three other words. In order to increase the chances of readers attaining word shape information parafoveally, the pretarget word was at least five letters in length. Potential targets meeting the above criteria but containing repeated letters (e.g., pepper) were excluded as information extracted from repeated letters may have given participants additional parafoveal information when the word is masked (e.g., yayycn) thus allowing easier identification of the parafoveal string and possibly leading to less interference from the mask. All sentences were controlled for contextual constraint; specifically, all sentences had a low level of word predictability. Neither the pretarget nor target word occupied the beginning or ending positions within the sentence.

Utilizing the boundary technique (Rayner, 1975), five viewing conditions were created for the experimental sentences (see Appendix E for example sentences and Table 2 for how stimuli lists were counterbalanced):

1. a no-change condition in which the target word was visible throughout the trial;
2. a similar-provide condition, in which the mask was composed of random letters similar to the target word and that maintain the shape of the target word (e.g., blanket 🡪dtcrial);
3. a dissimilar-provide condition, in which the mask was composed of random letters dissimilar to the target word and that also maintain the shape of the target word (e.g., blanket 🡪 fkwxdvb);
4. a similar-denied condition, in which the mask was composed of random letters similar to the target word and that denies word shape information (e.g., blanket 🡪 oxcrmax); and
5. a dissimilar-denied condition, in which the mask was composed of random letters dissimilar to the target word and that also denies word shape information (e.g., blanket 🡪 wsrxzva).

All masks were created in such a way that they were unpronounceable. This guaranteed that no mask was created which is a pseudoword/nonword, which might create an additional source of variance due to the processing of parafoveal sound information (Pollatsek, Lesch, Morris, & Rayner, 1992). As the participants reads each line of text and the eye crosses the invisible boundary present in each sentence, in this case the space between the pretarget and target words, the mask was removed and the target word will become visible.

There were 20 sentences per condition (totaling 100 sentences across the five conditions) and 60 filler sentences in Experiment 1. Because there are five conditions within Experiment 1, five counterbalance lists were created (see Table 2), and sentence presentation was randomized within each counterbalance list. For example, the set of 20 sentences that served as the no-change condition in one counterbalance list served as the similar-provide condition for another counterbalance list.

**2.2.3. Apparatus.** The EyeLink 2K measurement system (SR Research Ltd., www.eyelinkinfo.com, Mississauga, Ontario, Canada) was used in this experiment. The system uses infrared video-based tracking technology to compute the center and size of the pupils in both eyes. It has a relative spatial resolution of 0.02° RMS at 2000 Hz, and its absolute accuracy is below 0.25, depending on calibration (see McConkie, 1981, for details on calibration; Inhoff & Radach, 1998, for measurement issues). It allows accurate tracking without attaching parts to a participant’s head, allowing for maximally unobtrusive data collection over extended periods. All participants were tested on a Dell computer, used to control stimulus timing and record participant’s eye movements, saccadic accuracy, saccadic latency, saccadic trajectory and deviations, amplitude, and gain with a 22” monochrome CRT monitor (1024 by 768 at 200 Hz) on which the stimuli were presented. Eye movements were sampled twice per millisecond.

Although it is possible to record binocular eye movements, data was recorded from the right eye only during this experiment. The reason for this is that there is very little difference in the positions of the eyes in adult readers, and the difficulty in calibrating both eyes is high. Thus, recording from one eye is advantageous. The on-line saccade detector of the eye tracking system was set to detect saccades with an amplitude of 0.15° or greater, using an acceleration threshold of 8000°/sec² and a velocity threshold of 30°/sec. All stimuli were presented in 12-point Courier New font face. The distance between the display screen and the reader’s eyes will be set to 65 cm, thus causing each letter to subtended 0.33° of visual angle laterally.

**2.2.4. Procedure.** Participants were tested individually. Before completing any oculomotor tasks, the participant was assessed for near and far visual acuity. During the eye-tracking portion of the experiment, participants will be seated in a comfortable chair 65 cm from a 22-inch CRT monitor. A chin and forehead rest was used to stabilize the head, and participants were asked to avoid making any large head movements. Then the eye-tracking camera was adjusted until a clear corneal reflection was present. After setting the threshold for detecting the pupil, the system was calibrated. Calibration was performed before each block of trials and a drift correction before each trial to ensure high data quality with mean average deviation in validation trials under 0.33°. Once the calibration procedure was successfully completed, the experimental session started.   
 Participants were instructed to read for comprehension in order to answer such questions as they were presented during the experiment, which occurred every six trials. Participants then began with six practice sentences in order for them to become accustomed to both the type of stimuli and the procedures of the experiment. At the end of each trial, participants pressed a button to continue or to move on to the comprehension question screen. After the six practice trials, the experiment began. After completing the eye-tracking portion of the experiment, reading skill level was established using the Passage Comprehension, Word Attack, and Word Identification subtests of the Woodcock Reading Mastery test (Woodcock, 1998).

**2.2.5. Dependent Variables.** Critical dependent variables necessary to measure the contribution of word shape on parafoveal processing include first fixation duration, single fixation duration, gaze duration, landing position, and fixation probability. Fixations on a word were defined as a fixation on either the blank space at the beginning of the word or any of the letters within the word. First fixation duration was defined as the initial fixation on a word regardless of any later fixations. Any subsequent fixations on the word that occur before they eye leaves the word, but excluding saccade durations, were added to the first fixation duration to calculate the gaze duration. Total viewing time was defined as the summed duration of all fixations made on the word. The following conditions resulted in the exclusion of data from the analyses: (1) fixations must last between 70 and 800 ms, all others will be excluded; (2) initial fixations on the target word that are greater than three standard deviations from the mean; (3) fixations preceded by progressive saccades greater than twenty characters; or (4) any trials including blinks or loss of tracking during target word fixation.

**2.3. Results**

The results from Experiment 1 were analyzed with a one-way ANOVA using the variable Condition as a factor. Technically the use of one-way ANOVA is related to the fact that only one control condition is included in the design, leading to five different preview conditions. The underlying theoretical rationale is that the effect of a general preview factor is tested with all five levels hypothetically ordered along a dimension of how severely parafoveal preview is compromised. This approach is not ideal from a methodological point of view, but typical for publications in the field (see e.g. Pollatsek, Lesch, Morris, & Rayner, 1992, for a similar analysis with much more heterogeneous preview conditions). Throughout the following section, the main emphasis will be on planned comparisons especially with respect to the question whether preview effects are different between the shape provided / denied as well as the dissimilar / similar conditions.

**2.3.1. Temporal Analyses (“When”).** Table 3 shows the means for the temporal measures analyzed in Experiment 1. Fixation duration (see Figure 5) is the first fixation made on the target word whether or not there are any later fixations on that word. Single fixation duration (see Figure 6) is a subset of fixation duration that includes all cases in which only one fixation is made on the target word and, in such cases, the single fixation duration will equal the value of the gaze duration. Gaze duration (see Figure 7) is the total first pass reading time on the target word including the first fixation duration and any subsequent fixations before the eyes leave the word.

While it was expected that the parafoveal preview benefit would be seen in the earliest stages of processing, as reflected in elevated fixation durations and single fixation durations, this was not reflected in the results obtained as only the single fixation durations showed significant relationships. Even though there was not a significant effect of condition on fixation duration, there was a non-significant trend (F(4,136) = 2.148, p = .078) for words masked parafoveally to have longer viewing times than the legal preview condition. Planned comparisons reveal that this trend is driven by the slightly larger difference between the dissimilar mask/shape intact condition and the legal preview condition, a difference of 12 ms (t(34) = -1.694, p = .099 ). All other conditions had much smaller increases in viewing times when compared to the legal preview condition, none of which were significantly different nor demonstrating a non-significant trend: similar mask/shape intact v. legal preview: 9 ms; similar mask/shape denied v. legal preview: 1 ms; dissimilar mask/shape denied v. legal preview: 5 ms (see Table 3 for a summary of means and standard deviations). None of the similarity or shape planned comparisons showed any significant differences or non-significant trends between the shape or similarity conditions (all p values > .1). While unexpected in terms of this experiment, this result is not unusual in the literature as gaze duration is the measure with which the parafoveal preview benefit is typically assessed.  
 Despite the lack of a significant effect in the larger category of first fixation duration, in the specific subcategory of initial fixation measures, single fixation duration, there was a statistically significant effect of condition (F(4,136) = 5.022, p = .032). As seen in Figure 6, there seem to be relatively large differences between all conditions except the legal preview v. the similar mask/shape denied condition, which only had a 9 ms difference. Planned comparisons reveal, however, that the only significant difference is between the similar letter mask with shape information intact condition and the similar letter mask with shape information denied condition (t(34) = 2.257, p = .031), in which there was a 11 ms increase in viewing times for the similar letter mask with shape information intact condition. All other planned comparisons were not significant (all p values >.1).

There was also a robust parafoveal preview benefit obtained for the gaze duration measure (F(4,136) = 6.872, p < .0005). When the target word was masked with dissimilar letters and word shape information was provided, viewing times seemed to be longer than for a legal preview as indicated by planned comparisons. In comparison to the legal preview, dissimilar masks with shape information were viewed on average 40 ms longer and when shape information was absent, target words were viewed an average of 22 ms longer. Analyses demonstrated that similar masks were not significantly different from the legal preview condition with those containing shape information averaging 20 ms longer viewing times than the legal preview condition and those without shape information averaging only 11 ms longer viewing times (see Table 3 for a summary of means and standard deviations). Planned comparisons revealed two significant contrasts: (1) between the dissimilar letter mask with shape information provided condition and the dissimilar letter mask with shape information denied (t(34) = 2.339, p = .025) in which viewing times for the shape denied condition were 18 ms shorter than when shape information was provided; and (2) between the similar letter mask with shape information provided condition and the dissimilar letter mask with shape information provided condition (t(34) = -2.334, p = .026) in which similar letter masks resulted in viewing times 20 ms shorter than when the parafoveal mask was composed of dissimilar letters.

**2.3.2. Spatial Analyses (“Where”).** Table 4 shows the mean spatial measures analyzed in Experiment 1. Landing position (see Figure 8) is the character position within the target word that is fixated. For example, if the letter “c” were fixated in the word “bucket,” the landing position would be 3, reflecting the third letter within the word. Fixation probability (see Figure 9) is the probability, expressed as a proportion, that the target word will be fixated.

It was not anticipated that the manipulation of the shape or similarity of the mask would affect the decision of where to move the eyes or the likelihood that a target word would be fixated. While there was not a significant effect of landing position based on condition, there was a non-significant trend present in the analyses (F(4,136) = 2.270, p = .065). There was a tendency for readers to move less deeply into words that were parafoveally masked, .1 to .3 character spaces to the left of the average landing position (see Table 4 for a summary of means and standard deviations). There were no significant contrasts resulting from planned comparisons (all p values > .1).

Target words were more likely to be fixated if there were masked parafoveally than if the preview had been legal. Due to violations of the assumption of sphericity (Mauchly’s W(9) = .470, p = .004), a conservative adjustment was applied to the degrees of freedom, specifically the Greenhouse-Geisser correction, in order to account for such assumption violations. The resulting analysis of fixation probability was still significant (F(3.140,136) = 3.416, p = .018) after the adjustment was made. There were no significant contrasts resulting from planned comparisons (all p values > .1).

**2.4. Discussion**

The purpose of Experiment 1 was to investigate both the role of letter similarity in the creation of parafoveal masks, as well as the use of low-level word shape information in the processing of items in the parafovea. In the case of word shape information, inconsistent findings have been reported (e.g., Paap et al, 1984, v. Lete & Pynte, 2003). Most models of normal reading do not include word shape as a significant contributor to lexical processing (e.g., Pollatsek et al., 2006; Reilly & Radach, 2006), thus evidence regarding the processing of word shape information would be a valuable theoretical contribution to the experimental reading research literature. In the case of letter similarity in parafoveal masks, little to no research has been conducted (see Appendix C). As there is little consensus in the creation of masks, it is methodologically important to understand the impact of letter similarity on the parafoveal preview benefit, as such an understanding may reconcile inconsistent findings between laboratories insofar as the development of stimuli is concerned. For example, in some cases, similar letter masks are composed entirely of letters distinct from the target word, while in others there is overlap between the letters of the target word and the letters composing the mask. Despite the drastically different nature of these masks, they are both considered similar random letter masks. Unifying such disparate operational definitions would be methodologically desirable.

It was hypothesized that dissimilar letter masks would disturb processing more than similar letter masks and that those disturbances would be reflected in longer viewing times. It was also hypothesized that denying word shape information would cause larger disturbances in processing than when that information was provided. The results for Experiment 1 showed a general preview benefit across most temporal measures, most importantly gaze duration, the measure typically reflecting the preview benefit. Viewing times tended to be longer when words were masked parafoveally. For example, the first fixation duration measure indicated that a dissimilar letter mask with shape information intact tended to yield the longest viewing times when compared to a legal preview. Additionally, the single fixation duration measure also specified a dissimilar letter mask with shape intact as the source of significantly longer viewing times, although all other conditions had longer viewing times with the exception of the similar letter mask with shape denied (see Table 3). Finally, an analysis of gaze duration demonstrated that a dissimilar letter mask with shape information provided resulted in the longest viewing times.

While it is clear that dissimilar letter masks resulted in increased viewing durations, it is unknown whether this effect is modified or qualified by the presence or absence of word shape information, as there were no significant differences specifically related to the word shape manipulation. It was thought that the analysis of spatial measures could help clarify the relationship between letter similarity and word shape. An analysis of landing position did not yield significant results. However, a non-significant trend indicated that readers did not enter as deeply into a word when it was masked parafoveally, especially in the similar letter mask with word shape intact condition. Further, fixation probabilities were examined and suggest that the significant relationship seen was driven by the similar letter mask with shape intact condition as well. When these results were inspected in conjunction with those of the temporal measures, clear conclusions were difficult to derive.

Despite providing an advantageous environment in which a word shape effect could occur, a significant result for that manipulation did not emerge. A potential reason for this may be that this particular set of manipulations were not as ideal as they were initially conceptualized to be and did not sufficiently tap the mechanisms underlying word shape processing. The work of Lete and Pynte (2005), for example, did not rely solely on low-level manipulations as did the current experiment. Instead, they also included factors that required the use of higher level cognition by participants. In Experiment 2, Lete and Pynte (2005) employed a lexical decision task in which participants had to decide whether the stimulus they were presented with was a word or a nonword. They found that lexical decision times were significantly decreased, but this effect only held true for rare word shapes. In another experiment, common word shapes facilitated response times in a naming task. It seems clear that they uncovered different word shape effects based on the amount of lexical processing required by the task demands. In the task requiring relatively more higher-order cognition, the lexical decision task, only rare shapes were facilitated. Considering the manipulations in the current experiment required processing more in line with Lete and Pynte’s (2005) naming experiment, the failure to find a significant word shape effect can possibly be explained by a lack of word shape frequency manipulation. Thus, word shape information alone may not be as useful to word identification as word shape combined with other informational cues, such as orthography, syntax, or contextual constraint, its utility may be increased (Lete & Pynte, 2005).   
 Related work by Faust and Babkoff (2000) highlights the contribution of additional cues to the effect of word shape in a setting that is more similar to single sentence reading than the Lete and Pynte (2005) work. In the Faust and Babkoff (2000) study, participants were presented with an incomplete sentence beginning that was either neutral (“The next word is”) or highly predictable (The clown amused the crowd in the”) in conjunction with either a neutral (“xxx”) or informative word shape outline (see Faust & Babkoff, 200, p 53). Their results indicate that having a word outline shape present yielded significantly reduced lexical decision times as compared to the neutral place marker. Additionally, having highly predictable sentence context in conjunction with the word outline shape resulted in decreased lexical decision times from the word outline condition alone. This finding is particularly relevant to the field of experimental reading research as the parafoveal preview benefit can be modulated in size by contextual constraint. In cases where sentences are highly predictable, there is a much larger preview benefit observed than in cases where sentences do not have the same level of predictability (Ballota, Pollatsek, & Rayner, 1989; Rayner & Well, 1998). Therefore, it is possible that by combining a contextual constraint manipulation (i.e., predictability) with a word shape manipulation in the context of single sentence reading more subtle word shape effects may be seen.

**3. EXPERIMENT TWO: WORD SHAPE AND CONTEXTUAL CONSTRAINT**

Upon determining that there was a tendency for dissimilar letter masks to be more detrimental to lexical processing in Experiment 1, a second experiment was designed to understand the effect of word shape on cognitive processing during continuous reading. One of the most prominent methods to do this is by manipulating the *contextual constraint* (i.e., predictability) of the sentence frame in relation to the target word. For example, the sentence “the janitor mopped the floor” would be considered one in which there is high contextual constraint or predictability, whereas the sentence “the janitor mopped the bench” would be considered to possess low predictability. The way in which predictability is typically assessed is by presenting participants with the sentence up to the target word, in this case “the janitor mopped the”, and presenting them with a blank to fill in the word that comes next. Balota, Pollatsek, and Rayner (1985) used a 5-point scale to assess the predictability of their materials resulting in a mean rating of 4.47 for high predictability words and 2.32 for low predictability words (p. 371).

Predictability modulates the parafoveal preview benefit; specifically, in cases of high predictability there is more parafoveal preview obtained from upcoming words (e.g. Inhoff & Rayner, 1986; Glover, unpublished Master’s Thesis; Sawyer, unpublished Master’s Thesis). In a factorial design manipulating both contextual constraint and the similarity/relatedness of items in the parafovea, Balota, Pollatsek, and Rayner (1985) demonstrated that predictability influenced not only the amount of time readers fixated the target word during the first pass, but also whether they skipped that word and the duration of fixations on the word immediately to the right of the target word (i.e., spillover effects). Within identical sentence frames, target words were created that either would fit the frame of the sentence in a predictable manner or would fit the sentence but be relatively unpredictable. For example, take the sentence “The banker loaned the businessman some more (money/tools) for his new project” (Balota, Pollatsek, & Rayner, 1985, p. 372). In this example, “money” is the predictable target word, and “tools” is the unpredictable target word. The target words (money/tools) were also masked by a visually overlapping similar mask (e.g., moncg/toohz), a semantically related mask (tools/money), a visually dissimilar overlapping mask (toohz/moncg), and an anomalous mask (house). The results indicated that the highly predictable letter string mask had a much higher skipping rate (.11) than when the preview strings provided an unpredictable letter string or anomalous word (approx .01). Furthermore, gaze durations for the identical and visually similar condition were relatively normal in sentences containing a high predictability target whereas all other masks and all results from the low predictable condition were significantly larger.

From these results, Balota et al. (1985) concluded that contextual constraint modulates the amount of parafoveal information that is acquired during reading, resulting in increased time to achieve lexical access while the reader does a “double take” (p. 376). Interestingly, there was no effect on first fixation duration but only gaze duration, suggesting that predictability affects later stages of processing, up to and including post-lexical processing. Further evidence for an effect on later processing comes from single sentence reading that indicates that only low predictability words had significantly larger second pass times (Kliegl, Grabner, Rolfs, & Engbert, 2004). Some have proposed that this effect arises from the transitional probabilities between words (i.e., word-to-word contingencies) leading to shorter fixations on words which commonly occur together in a particular order (e.g., “rely” is usually followed by “on”; McDonald & Shillcock, 2003a). By being able to create probabilities of word-to-word contingencies, readers are able to tap such statistical information and create almost immediate predictions about words not yet encountered in the text (McDonald & Shillcock, 2003b). It appears likely that in reality such low level causes of predictability co-exist with more high level top down contextual effects (Frisson, Rayner, & Pickering, 2005).

**3.1. Method**

**3.1.1. Participants.** Thirty-six Florida State University college undergraduates from General Psychology courses that include an option for experiment participation as a part of course credit and community members were recruited for participation in this experiment. The size of the sample is again based on a power analysis that indicated that assuming a medium (approximately .5) to low (.2 to .3) effect size, 32 participants are required to achieve conservative power of .95 of this magnitude with an alpha level of .05 (G\*power; Faul, Erdfelder, Lang, & Buchner, 2007). In order to have an equal number of participants in each counterbalance condition (see Materials and Design section below), this number was increased to 36 participants. All participants had normal or corrected-to-normal vision. Participants were either native speakers or had spoken and read English for at least 10 years. The sample was composed of 49% females with a mean age of 21 years. No information on race or ethnicity was recorded as there is no reason to assume differences in oculomotor behavior based on either race or ethnicity.

**3.1.2. Materials and Design.** A within-subjects 2 (contextual constraint condition) x 3 design (parafoveal preview condition) was used in this experiment. Seventy-two double line sentences of no more than 85 characters were constructed, each containing a manipulated critical target word. As in Experiment 1, all target words were six to eight words in length to decrease the probability of skipping (Brysbaert et al., 2005). Target words were controlled for lexical frequency (high), word shape frequency (rare), and were masked with dissimilar letter masks, which cause the most interference in processing during reading as established in Experiment 1. Both lexical frequency and word shape frequency were determined in the same manner as in Experiment 1. All target words and masks maintained the same characteristics and controls as specified in Experiment 1.

Utilizing the boundary technique (Rayner, 1975), six viewing conditions were created for the experimental sentences (see Appendix F for example sentences and Table 6 for counterbalance lists):

1. a predictable no-change condition in which the target word was visible throughout the trial, and the sentence context was predictable following the target word;
2. an unpredictable no-change condition in which the target word was visible throughout the trial, and the sentence context was not predictable following the target word;
3. a predictable-provide condition, in which the shape of the target word was also maintained in the mask, and the sentence context was predictable following the target word;
4. an unpredictable-provide condition, in which the shape of the target word was also maintained in the mask, and the sentence context was not predictable following the target word;
5. a predictable-denied condition, in which the shape of the target word was not maintained in the mask, and the sentence context was predictable following the target word;
6. an unpredictable-denied condition, in which the shape of the target word was not maintained in the mask, and the sentence context was predictable following the target word.

There were 12 sentences per condition (totaling 72 sentences across the six conditions) different from those used in Experiment 1 and 48 filler sentences. Because there are six conditions within Experiment 2, six counterbalance lists were created (see Table 6), and sentence presentation was randomized within each counterbalance list. For example, the set of 12 sentences that served as the predictable no-change condition in one counterbalance list served as the predictable-provide condition for another counterbalance list. An equal number of participants (in this case, six) were exposed to each condition.

**3.1.3. Apparatus.** The same equipment from Experiment 1 was used in the current experiment.

**3.1.4. Procedure.** The same procedure was used as that in Experiment 1 with the exception that the vision tester was no longer used.

**3.1.5. Dependent Variables.** The dependent variables measured in Experiment 1 were used in the current experiment.

**3.2. Results**

The results from Experiment 2 were analyzed with a repeated measures ANOVA using two factors: shape and context. Throughout the following section, the main emphasis will be on planned comparisons especially with respect to the question whether preview effects are different between the shape provided/denied as well as the normal/high context conditions.

**3.2.1. Temporal Analyses (“When”).** Table 7 shows the mean temporal measures analyzed in Experiment 2. As mentioned before, fixation duration (see Figure 10) is the first fixation made on the target word whether or not there are any later fixations on that word. Single fixation duration (see Figure 11) is a subset of fixation duration that includes all cases in which only one fixation is made on the target word and, in such cases, the single fixation duration will equal the value of the gaze duration. Gaze duration (see Figure 12) is the total first pass reading time on the target word including the first fixation duration and any subsequent fixations before the eyes leave the word.

There were main effects for both the word shape factor (F(2,70) = 22.996, p < .0005) and the context factor for first fixation duration. (F(2,70) = 5.462, p = .025) The initial fixations that readers made on target words were 21 ms longer when the word was masked and word shape was intact or 22 ms when word shape was denied than when the preview was legal, a statistically significant finding as revealed through planned comparisons (F(5,215) = 4.454, p = .001). Additionally, the first fixation duration was 13 ms longer in the case of legal preview when comparing high and normal contextual constraints, a significant finding through planned comparisons (F(1,35) = 5.462, p = .025). These results are desirable as they confirm that the intended manipulations were successful. Aside from the robust parafoveal preview effect seen between legal and illegal preview, there was only 1 ms difference between masks in which word shape was intact and denied (n.s., p > .05). Further, there was no interaction between the word shape factor and the context factor (n.s. , p > .05).

In the case that readers only fixated the target word one time (i.e., single fixation duration), viewing times were 23 ms longer when a legal preview was provided than when the word was masked with shape intact or between 19 and 27 ms with shape denied, thus driving the main effect for shape seen in Experiment 2 (F(2,70) = 26.242, p < .0005). This finding was also supported through planned comparisons (F (5,215) = 6.092, p < .0005). There was no difference in the cases when word shape information was provided and denied (n.s., p > .5). While disappointing, this was not a surprising outcome considering the results of Experiment 1. The main effect of context was also significant (F(1,35) = 5.097, p = .030). Viewing time was 13 ms longer in the case of legal preview in the normal context condition than in the high context condition although there was no difference between normal and high context when words were masked in any way (n.s. , p > .05). Interestingly, there was a non-significant trend toward an interaction between the word shape factor and the context factor in terms of the single fixation duration measure (F(2,70) = 2.92, p = .061). While viewing times were generally equal for high and normal context in the shape intact condition, the trend indicates that viewing times tended to be somewhat longer in the shape denied/normal context condition than in the shape denied/high context condition (see Figure 11). This finding is particularly interesting as it supports the supposition that word shape information would be utilized in situations involving higher-order cognition.

Main effects for both shape and context were found for the gaze duration measure (Shape: F(2,70) = 27.531, p < .0005; Context: F(1,35) = 15.594, p < .0005). Gaze durations were 36 ms longer in the normal context condition and 47 ms longer in the high context condition when the word was masked and word shape was intact. When word shape was denied, gaze durations were 39 ms longer in the normal context condition and 46 ms longer in the high context condition when the word was masked and word shape was intact. The difference between the legal preview and masked conditions was statistically significant as revealed through planned comparisons (F(5,215) = 6.769, p < .0005). The interaction between shape and context was not significant for gaze duration (F(2,70) = .562, p > .05).

**3.2.2. Spatial Analyses (“Where”).** Table 8 shows the mean spatial measures analyzed in Experiment 2. As mentioned previously, landing position is the character position within the target word that is fixated (see Table 8 for a summary of means). Fixation probability is the likelihood, expressed as a proportion, that a target word will be fixated (see Table 8 for a summary of means). As with Experiment 1, it was not anticipated that the shape or context manipulations would affect the decision of where to move the eyes as reflected in the landing position measure. The context manipulation, on the other hand, is likely to have an effect on the fixation probability of the target word.

Landing position did not seem to be influenced by either the shape (F(2,70) = 2.355, p = .102) or the context manipulation (F(1,35) = .010, p = .921). Landing positions were shifted .2 character spaces to the right in the normal context condition and were also shifted .2 character spaces to the right in the high context condition when the word was masked and word shape was intact (see Table 8). When word shape was denied, landing positions were not shifted at all in the normal context condition and shifted .1 character spaces to the right in the high context condition when the word was masked and word shape was intact. The interaction component of the analysis was not significant (F(2,70) = .401, p = .901).

The pattern of results obtained for landing position was also seen for fixation probability. There was not a significant main effect for either factor, although there was a non-significant trend seen for shape (F(2,70) = 2.936, p = .060). Readers were more likely to skip target words when they were not masked than when they were. There was also a non-significant trend for the shape by context interaction. Because assumptions of sphericity were violated for the interaction analysis, a relatively conservative adjustment, the Greenhouse-Geisser correction, was used to account for the violated assumption (F(1.713,70) = 2.868, p = .072). Target words were more likely to be fixated in the normal context condition than in the high context condition particularly in the legal preview condition. This trend toward an interaction seems to be driven by the reduced fixation probability in the legal preview/high context condition, t(35) = 1.909, p = .064, as this is the only comparison that neared significance.

**3.3. Discussion**

Experiment 2 was designed to investigate whether additional context cues would facilitate the processing of word shape in single sentence reading. Work by Faust and Babkoff (2000) examined the contribution of additional cues to the effect of word shape and determined that providing word shape information yielded significantly reduced lexical decision times as compared to situations where such information was denied to the reader. It can be concluded that by manipulating both contextual constraint (i.e., predictability) and word shape in the context of single sentence reading, the likelihood of finding an effect of word shape would be increased. Finding an effect of word shape would be theoretically interesting on several levels. First, a significant finding would help resolve historically inconsistent findings in the literature (see Paap et al., 1984, and Lete & Pynte, 2005, for examples) that have, secondly, impacted the creation and implementation of computational models designed to simulate normal reading behavior (e.g., Reichle et al., 2006; Reilly & Radach, 2006).

It was hypothesized that the results of Experiment 2 would be similar to those of Faust and Babkoff (2000) in that having redundant information provided by both high contextual constraint (e.g., “the janitor mopped the floor) and word shape would result in the least impact on processing during reading. Conversely, situations in which contextual constraints were low (e.g., “the janitor mopped the bench”) and word shape information was denied would cause increases in the dependent measures. Additionally, it was expected that there would be both typical context effects and typical parafoveal preview effects. Indeed, all viewing time measures obtained supported the context and preview effects. In all cases, viewing times were longer in the normal context condition than in the high context condition and also when words were masked parafoveally than when a legal preview was provided. Unlike the temporal measures, none of the spatial measures were significantly affected by the manipulation of context.

Consistent with the results of Experiment 1, there was no effect of landing position as only dissimilar letter masks were used in Experiment 2. This result is also consistent with recent findings reported by Vainio, Hyona, and Pajunen (2009). In a study that specifically examined the effects of contextual constraint on both temporal and spatial measures during sentence reading. The authors report that the predictability manipulation yielded shorter viewing times in the high contextual constraint position, there was no effect on landing position, thus implicating predictability in the “when” decision but not the “where.” While there was not a significant effect of context on fixation probability in the current experiment, there was a non-significant trend present in the data. Specifically, in the high context condition, target words tended to be skipped 7% than in the normal context condition. The small difference observed is likely caused by a ceiling effect due to relatively long target words (see Brysbaert et al., 2005, for a detailed discussion). Although the data supported the context and preview hypotheses, there was no evidence supporting the word shape hypotheses. There was not a significant effect of word shape on any data, neither temporal nor spatial.

There are several possible reasons why the strong shape manipulation included in Experiment 2 did not produce significant findings. It may be the case that word shape is not utilized in English as much as in other languages; for example, experiments that produced significant findings tended not to include English stimuli: the Lete and Pynte (2005) study was in French and the Faust and Babkoff (2000) study was in Hebrew. Studies that did utilize English stimuli tended to not to find an effect, such as the Paap et al. (1984) work. The additional shape information in these languages, provided by additional accents and markers both above and below the text, could give readers additional outline information other than the ascending and descending features seen in English. It may also be the case that the word shape effect seen in these studies might be task dependent. In the Lete and Pynte (2005) study, words were presented in isolation for brief periods of time. Faust and Babkoff (2000) presented participants with only partial sentences, which is also quite removed from normal sentence reading. Critically, participants had much more time to solve the task than is available during normal reading. On the other hand, Paap and colleagues (1984) had participants read in a natural situation while conducting a proofreading task. When comparing these studies, in light of the results of the current work, one might conclude that word shape information can become relevant when additional information from surrounding words or the context of a full sentence is unavailable in tasks that have more of an offline character. The present experiment, however, was the first to study the issue in a normal sentence reading situation, and, despite a strong manipulation and favorable conditions, no word shape effect emerged.

**4. GENERAL DISCUSSION**

**4.1. Background and Results Summary**

Since the dawn of humankind, there has been a fascination with vision and how the visual system functions (Wade, 1998). With the advent of writing systems, that fascination has extended to how readers process the written word (Huey, 1908). Initial observations of visual behavior were restricted to direct observation of the eyes; however, as technology has advanced, so have the devices used to investigate eye movements (Young & Sheena, 1975). Despite early, crude investigations, the body of reading literature has expanded exponentially with the introduction of small and portable computers (Rayner, 1978). Modern reading researchers have taken advantage of this advanced technology and are building upon the foundation of earlier eras of research to focus on understanding reading on a more functional level through computational modeling (e.g., Pollatsek et al., 2006; Reilly & Radach, 2006).

Despite methodological and theoretical advances, there are still areas of importance that either have not been fully investigated or need to be reevaluated. For instance, while there is a broad empirical base supporting the parafoveal preview benefit, little is known about how differences in the methodological implementation of various mask types can affect the size of the benefit. Additionally, inconsistencies in empirical findings are common in the literature. Such inconsistencies are theoretically diverse and include, for example, questions regarding whether the parafoveal preview benefit extends to word N+2 and whether word shape information facilitates parafoveal processing. The results from the current experiments were intended to address some of these methodological and theoretical concerns.

Experiment 1 highlighted the methodological importance of mask creation and demonstrated for the first time that mask similarity has a significant impact on the preview benefit. The similarity of masks to the target words they obscure can significantly impact common measures of viewing time, such as gaze duration; specifically, when the mask employs letters that are dramatically dissimilar to the target word, longer viewing times occur. This result occurred across all critical temporal measures in Experiment 1. Contrary to our initial expectation, there was no support in Experiment 1 for the hypothesis that word shape information is used in lexical processing and that such information, when denied to the reader, would result in longer viewing times. Experiment 2 was designed to extend the word shape findings from Experiment 1 as additional contextual cues may facilitate the use of word shape information. Robust parafoveal preview and contextual constraint effects occurred in the temporal measures. When words were masked in the parafovea, viewing times were longer than if the word had not been previously masked. Additionally, viewing times were significantly shorter in the high context condition than in the normal context condition regardless of the presence of masked words. Critically, despite a strong manipulation as established by previous work (e.g., Faust & Babkoff, 2000; Lete & Pynte, 2005; Paap et al., 1984) and very favorable conditions, such as controlling for low shape frequency, high lexical frequency, and including a contextual constraint manipulation, there was no effect of the word shape manipulation observed in any of the dependent measures. These results will be discussed in terms of both methodological and theoretical factors including the need for stringent control in stimulus creation and how processing is allocated parafoveally.

**4.2. Theoretical Implications**

The results from both experiments are enlightening both methodologically and theoretically across several issues important to current cognitive reading research. First, and perhaps most critically, it is now possible to address how letter similarity and word shape information contribute to the decisions [[1]](#footnote-1)of “when” and “where” to move the eyes in the context of single sentence reading. Similarity seems to affect when to move the eyes, but not where. Word shape, on the other hand, seems to affect neither decision.

It is more expedient to begin with a discussion of the “where” decision. None of the measures recorded during the current experiments affected the landing position on the previously masked target word. These results are not surprising as recent evidence has established that landing position is not effected by contextual constraint manipulations (Vainio, Hyona, & Pajunen, 2009) even though viewing time measures were affected by that manipulation—there were shorter viewing times in the high contextual constraint position. As there was no effect on landing position, the “when” decision is implicated in contextual constraint, as in Experiment 2, but not the “where.”

An interpretation of the current results in the context of the “when” decision is somewhat more complex. Models of normal reading provide the context in which to understand why a similar letter mask in the parafovea affects the decision of when to move the eyes. These models all make the assumption that linguistic processing is spatially distributed, and most importantly, that there is substantial parafoveal preprocessing of the upcoming text. In the E-Z Reader model (Pollatsek et al., 2006), a serial attention model, early lexical processing occurs including a familiarity check that informs the rest of the model of the potential difficulty of the upcoming word once the currently fixated word has achieved lexical access. In the Glenmore model (Reilly & Radach, 2006), a processing gradient model, all words contained within the perceptual span are processed together, but differ in the amount of saliency each item has at any given time. Just as these models vary in how attention is distributed and how much preprocessing of upcoming text occurs, each model handles the parafoveal masking of upcoming target words differently. The Glenmore model (Reilly & Radach, 2006) can more flexibly handle random letter masks by treating them as exceedingly low frequency words. On the other hand, the E-Z Reader model (Pollatsek et al., 2006) stops processing if a word cannot reach lexical access. This means that this model would predict no difference between similar and dissimilar letter masks. The results from Experiment1 indicated significant differences between the legal preview and dissimilar mask conditions, a finding that is more in line with processing gradient models, like Glenmore and SWIFT.

While the letter similarity manipulation affects the decision of when to move the eyes, the word shape manipulation did not. Neither experiment produced results that indicated a significant effect of word shape on any of the variables measured. This was the case even though it is possible, in theory, to extract word shape information within the perceptual span, as often two upcoming words fall within its range. Paap and colleagues (1984) elegantly describe why the word shape hypothesis continues to be favored, in terms of continuous reading, despite a body of evidence pointing to letter identification as a major contributor to parafoveal word recognition. They mention that “the possibility that the shapes of upcoming words in the text are preprocessed when they are in parafoveal view is very seductive because a word's holistic shape does not seem to require the fine acuity needed to resolve the specific features required for letter recognition” (Paap et al., 1984, p. 426). Despite this possibility, data from the current experiments indicate that word shape is not utilized in parafoveal word recognition.

There have been several hypotheses proposed that address why word shape is in “poor shape.” Paap and colleagues (1984) proposed an uncertainty-reduction hypothesis, which states that the only time that word shape will assist in recognition is when a word is unique in shape or shares that shape with relatively few other words. In the case where there are many words that share the same shape, the pool of candidate words that could match in shape is not sufficiently reduced, thus processing will proceed with letter identification. This hypothesis rests on the so-called *shape invariance assumption* that somewhere in the word recognition system is stored the shapes of all words ever encountered. In order to lessen the chances of shape neighbors interfering with shape recognition, the target words included in this study were chosen to minimize the number of shape neighbors they possessed, emphasizing the use of targets with no neighbors at all.

Despite trying to create the most advantageous environment possible to elicit a word shape effect, there was no effect of word shape on the preview benefit. On explanation for failing to find the effect may be because so few words have a unique shape. Consider Walker’s (1987) analysis of the Kucera and Francis (1967) corpus. This analysis identified unique shape in less than 3% of all words. In the current experiments, 299 potential target words were identified out of a database of over 54,000 words. Taken together, this evidence leads to the conclusion that the uncertainty-reduction hypothesis is of little explanatory value. It seems unlikely that word recognition system would rely on such a costly and cumbersome process to help identify word candidates when more economical and reliable parafoveal information, such as initial n-gram frequency, is available (see Radach, Inhoff, & Heller, 2004, for a discussion). Considering the limited resources allocated to the parafovea while a foveated word is being processed (Henderson & Ferreira, 1990), it makes sense that parafoveal preprocessing is supported by the most efficient and informative features of the text.

Another argument against the use of parafoveal word shape information is that the majority of studies providing significant effects of word shape are employing experimental conditions that are remote from normal reading whether it is single word reading (Lete & Pynte, 2005) or report data on reading in other languages (e.g., Hebrew, Faust & Babkoff, 2000). In other languages, there are sometimes additional cues, such as accents above or below the letter as in French, or special characters that indicate the beginning or endings of words, such as in the unspaced language of Thai. These additional cues may be more amenable to a word shape manipulation than English. Readers of English most likely rely on textual cues that are more available, like initial bigram or trigram frequency in which certain letters appear more often together at the beginning of words than others (e.g., “th” v. “rh”). This argument is perhaps the most powerful of all because of the strong evidence of individual letter information being used by the perceptual system—despite being able to use such fine grained information, more coarse shape information is not used.

A final point can be made in terms of the composition and implementation of computational models of reading. A recent study by Inhoff, Radach, Eiter, and Juhasz (2003) examined the role of word length in lexical processing during normal reading. Similar to the assumptions made about word shape, word length is often thought to reduce the amount of potential candidate words, thus aiding in word recognition. To test this assumption, the authors masked parafoveal word length information. Because they did not find a reliable effect of the word length manipulation, they concluded that knowing exactly how many letters are in a word is not used in word recognition. Instead, it is more likely that a coarse-grained process that generally identifies whether a word is short or long is at work to guide the programming of the subsequent eye movement. When taken in conjunction with the current findings, they provide consistent findings about what sort of low level information is used by the visual system and should be included in models of reading.

These results are important as they inform the modules included in models of reading. The E-Z Reader model contains two components of lexical processing, L1 and L2. During the L1 stage, a “familiarity check” is completed allowing the next eye movement to be programmed upon completion. The L2 stage, on the other hand, allows attention to move to the next word once that process is complete. In addition, newer versions of the model now also contain a preattentive module. This module takes in raw visual information and maintains it during saccades, where it is replaced by new low-level visual information as the sensory system sends it to the brain. They estimate that the time it takes for such information to reach the brain is 50 ms, thus there is a lag between what the eye takes in and what the brain is processing. It is likely that word shape and length information would be processed during this early stage. As the current experiments rule out the use of word shape information and the Inhoff, Radach, Eiter, and Juhasz (2003) research rules out length information, it now calls into question of what information is actually processed in this early stage for readers of English.

**4.3. Methodological Importance**

Several studies have investigated the confusability (i.e., similarity) of alphabet letters in an effort to better understand how the visual system processes and identifies letters (Pelli, Burns, Farell, & Moore-Page, 2006). The results from Experiment 1 indicated a significant impact of letter similarity on parafoveal preview. As this is a factor not often controlled in studies on reading, this result calls into question the current lack of standards for random letter mask creation. Critically, similar and dissimilar letter masks are not equivalent—dissimilar masks are significantly more powerful than their similar counterparts. This is important because it gives insight into the stages of processing at which such information is used. Specifically, dissimilar letter masks seem to be more effective on earlier stages of processing, it indicates a significant disruption in the letter recognition process (e.g., the L1 stage of the E-Z Reader model, Reichle et al., 2006). Similar masks do not seem to be as disruptive to early processing, rather they show a significant effect on later processing as reflected in the gaze duration measure.

Significant findings for the similarity factor in Experiment 1 underscore the utility of having a methodological tool that would enable researchers to create and control letter masks. The pilot study conducted prior to creating the stimuli for these experiments resulted in a confusability metric based on saccade latencies (see Table 9). This provides the reading research community with a tool to be used to avoid potential confounding as seen in the Paap et al. (1984) study could have been avoided.

**4.4. Conclusions**

Despite having given word shape a fair chance, in both of the current experiments it has been shown that there is no influence of word shape on parafoveal processing even under conditions that are considered favorable. It seems that the findings of Paap et al. (1984) are correct and that more recent findings from the single word reading literature are not applicable to experimental reading research. From the viewpoint of spatially distributed processing in reading, these early conclusions have stood the test of time, at least in terms of reading in English. This is important for models of reading because word shape is not a factor that needs special consideration. Despite multiple sources of well-controlled, empirical evidence, it is not likely that the debate will end soon. As Paap and colleagues note, the idea that word shape is acquired parafoveally is “seductive” (p. 426) and not likely to be abandoned without more lively debate. One direction this debate might take is by examining word shape in the context of bilingualism. By having readings fluent in a language that contains more word shape (e.g., Hebrew or French) as well as in English, it may be possible to determine how much shape information is utilized by each language respectively. Additionally, a powerful way to investigate the contribution of word shape to lexical processing in English is by providing misleading shape information in the parafovea (e.g., “little” masked by “qnqqno”). By providing misleading information, it is possible that inappropriate word candidates will be selected and processed thus causing massive errors in late processing s reflected in the gaze duration measure.

**5. APPENDIX A: TABLES**

*Table 1. Summary of visual assessment taken during Experiment 1 including the type of measure, values considered normal, and the mean scores for all participants with the standard deviation noted parenthetically.*

|  |  |  |
| --- | --- | --- |
| Measure | Acceptable Scores | Mean Score |
| Near Vision | 20/20 to 20/40 | 22/20 (6) |
| Far Vision | 20/20 to 20/40 | 23/20 (5) |
| Far Lateral Phoria | 4 to 12 | 9 (2) |
| Near Lateral Phoria | 4 to 12 | 9 (2) |
| Far Vertical Phoria | 3 to 5 | 4 (1) |

*Table 2. Counterbalance conditions for Experiment 1.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Senten-ces | Counter-balance List 1 | Counter-balance List 2 | Counter-balance List 3 | Counter-balance List 4 | Counter-balance List 5 |
| Set 1 | No-change | Dissimilar-Denied | Similar-Denied | Dissimilar-Provide | Similar-Provide |
| Set 2 | Similar-Provide | No-change | Dissimilar-Denied | Similar-Denied | Dissimilar-Provide |
| Set 3 | Dissimilar-Provide | Similar-Provide | No-change | Dissimilar-Denied | Similar-Denied |
| Set 4 | Similar-Denied | Dissimilar-Provide | Similar-Provide | No-change | Dissimilar-Denied |
| Set 5 | Dissimilar-Denied | Similar-Denied | Dissimilar-Provide | Similar-Provide | No-change |

*Table 3. Mean processing times for temporal measures in Experiment 1. All values are in ms followed parenthetically by the standard deviation. For each cell, the number of subjects included is 35.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | First Fixation Duration | Single Fixation Duration | Gaze Duration |
|  |  |  |  |
| Legal Preview | 285 (31) | 294 (35) | 347 (52) |
| Similar Letters—Shape Intact | 294 (35) | 314 (39) | 367 (55) |
| Dissimilar Letters—Shape Intact | 297 (38) | 317 (47) | 387 (55) |
| Similar Letters—Shape Denied | 286 (31) | 303 (36) | 358 (47) |
| Dissimilar Letters—Shape Denied | 290 (38) | 311 (39) | 369 (54) |

*Table 4. Mean processing times for spatial measures in Experiment 1. All values are in ms followed parenthetically by the standard deviation. For each cell, the number of subjects included is 35.*

|  |  |  |
| --- | --- | --- |
|  | Landing Position | Fixation Probability |
|  |  |  |
| Legal Preview | 3.5 (.6) | .97 (.05) |
| Similar Letters—Shape Intact | 3.2 (.6) | .99 (.02) |
| Similar Letters—Shape Denied | 3.4 (.5) | .98 (.03) |
| Dissimilar Letters—Shape Intact | 3.3 (.6) | .99 (.03) |
| Dissimilar Letters—Shape Denied | 3.4 (.6) | .98 (.03) |

*Table 5. Summary of scores from the Woodcock Reading Mastery Test in Experiment 1, including the subtest given, the maximum score for each subtest, and the mean score for all participants with the standard deviation noted parenthetically.*

|  |  |  |
| --- | --- | --- |
| Measure | Maximum Score | Mean Score |
| Word Identification | 106 | 101 (4) |
| Word Attack | 45 | 40 (3) |
| Passage Comprehension | 67 | 59 (4) |

*Table 6. Counterbalance lists for Experiment 2.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Senten-ces | Counter-balance List 1 | Counter-balance List 2 | Counter-balance List 3 | Counter-balance List 4 | Counter-balance List 5 | Counter-balance List 6 |
| Set 1 | Predictable-no-change | Unpredictable-Denied | Unpredictable-Provide | Predictable-Denied | Predictable-Provide | Unpredictable-no-change |
| Set 2 | Unpredictable-no-change | Predictable-no-change | Unpredictable-Denied | Unpredictable-Provide | Predictable-Denied | Predictable-Provide |
| Set 3 | Predictable-Provide | Unpredictable-no-change | Predictable-no-change | Unpredictable-Denied | Unpredictable-Provide | Predictable-Denied |
| Set 4 | Predictable-Denied | Predictable-Provide | Unpredictable-no-change | Predictable-no-change | Unpredictable-Denied | Unpredictable-Provide |
| Set 5 | Unpredictable-Provide | Predictable-Denied | Predictable-Provide | Unpredictable-no-change | Predictable-no-change | Unpredictable-Denied |
| Set 6 | Unpredictable-Denied | Unpredictable-Provide | Predictable-Denied | Predictable-Provide | Unpredictable-no-change | Predictable-no-change |

*Table 7. Mean processing times for temporal measures in Experiment 2. All values are in ms followed parenthetically by the standard deviation. For each cell, the number of subjects included is 36.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Fixation Duration | Single Fixation Duration | Gaze  Duration |
|  |  |  |  |
| Normal Context—Legal Preview | 276 (39) | 287 (41) | 317 (51) |
| Normal Context—Shape Intact | 297 (45) | 310 (49) | 353 (57) |
| Normal Context—Shape Denied | 298 (39) | 314 (45) | 356 (54) |
| High Context—Legal Preview | 263 (40) | 267 (43) | 293 (50) |
| High Context—Shape Intact | 293 (37) | 310 (51) | 340 (58) |
| High Context—Shape Denied | 293 (37) | 306 (44) | 339 (64) |

*Table 8. Mean processing times for spatial measures in Experiment 2. All values are in ms followed parenthetically by the standard deviation. For each cell, the number of subjects included is 36.*

|  |  |  |
| --- | --- | --- |
|  | Landing Position | Fixation Probability |
|  |  |  |
| Normal Context—Legal Preview | 3.8 (.7) | .96 (.06) |
| Normal Context—Shape Intact | 3.6 (.8) | .96 (.05) |
| Normal Context—Shape Denied | 3.8 (.7) | .97 (.06) |
| High Context—Legal Preview | 3.9 (.5) | .93 (.09) |
| High Context—Shape Intact | 3.7 (.7) | .97 (.05) |
| High Context—Shape Denied | 3.8 (.7) | .96 (.06) |

*Table 9. Confusability metric based on saccade latencies from dissertation pilot study (see Appendix D).*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Letter | | most dissimilar | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | most similar | |  | Letter |
| a | t | y | x | g | w | p | j | i | k | d | m | l | f | b | q | v | n | r | s | z | u | c | h | o | e | a |
| 360 | 360 | 363 | 365 | 367 | 368 | 368 | 370 | 370 | 371 | 374 | 375 | 375 | 375 | 377 | 378 | 381 | 382 | 387 | 390 | 394 | 396 | 401 | 406 | 458 |
| b | f | t | w | m | k | r | x | v | u | y | n | s | z | l | a | h | g | i | c | j | e | o | q | p | d | b |
| 347 | 350 | 351 | 360 | 361 | 362 | 363 | 364 | 365 | 372 | 372 | 374 | 375 | 375 | 375 | 377 | 380 | 383 | 385 | 388 | 394 | 406 | 419 | 424 | 445 |
| c | r | f | m | h | p | j | k | y | z | n | v | u | l | g | b | t | d | w | s | a | q | i | e | o | x | c |
| 349 | 349 | 360 | 361 | 361 | 364 | 364 | 374 | 375 | 375 | 376 | 377 | 378 | 378 | 385 | 386 | 387 | 395 | 396 | 396 | 397 | 402 | 404 | 406 | 431 |
| d | l | k | y | r | w | z | h | i | a | e | m | n | t | f | x | s | g | c | u | o | v | j | p | q | b | d |
| 351 | 357 | 358 | 362 | 362 | 364 | 365 | 367 | 371 | 372 | 374 | 379 | 381 | 382 | 383 | 384 | 386 | 387 | 387 | 388 | 392 | 402 | 404 | 422 | 445 |
| e | y | v | j | i | r | w | z | k | d | q | g | x | u | t | f | l | m | p | b | h | c | o | s | n | a | e |
| 337 | 350 | 365 | 366 | 369 | 370 | 371 | 372 | 372 | 373 | 373 | 376 | 382 | 384 | 384 | 386 | 387 | 391 | 394 | 403 | 404 | 406 | 410 | 416 | 458 |
| f | o | b | c | j | p | t | m | q | g | x | l | k | v | w | a | d | u | z | s | e | n | h | i | y | r | f |
| 336 | 347 | 349 | 358 | 359 | 363 | 364 | 365 | 365 | 366 | 371 | 372 | 375 | 375 | 375 | 382 | 382 | 382 | 383 | 384 | 384 | 391 | 394 | 400 | 414 |
| g | u | h | t | j | k | v | l | x | w | a | f | z | m | i | e | y | c | b | d | s | r | n | o | q | p | g |
| 342 | 353 | 354 | 354 | 356 | 356 | 359 | 362 | 365 | 365 | 365 | 366 | 367 | 371 | 373 | 376 | 378 | 380 | 386 | 405 | 407 | 411 | 415 | 427 | 430 |
| h | i | g | y | c | t | d | q | o | p | b | m | w | x | j | f | z | k | u | r | v | a | e | l | n | s | h |
| 344 | 353 | 358 | 361 | 362 | 365 | 369 | 371 | 376 | 377 | 384 | 387 | 388 | 389 | 391 | 394 | 396 | 397 | 398 | 399 | 401 | 403 | 408 | 410 | 418 |
| i | h | n | m | o | q | e | d | z | a | g | x | w | p | v | b | s | t | y | u | r | f | c | k | j | l | i |
| 344 | 357 | 364 | 365 | 366 | 366 | 367 | 367 | 370 | 371 | 372 | 375 | 375 | 378 | 383 | 386 | 387 | 388 | 393 | 393 | 394 | 402 | 402 | 419 | 428 |
| j | m | q | n | x | w | g | z | f | c | e | a | r | v | k | p | t | s | u | b | h | l | o | y | d | i | j |
| 340 | 351 | 351 | 353 | 354 | 354 | 356 | 358 | 364 | 365 | 368 | 370 | 375 | 377 | 378 | 383 | 385 | 385 | 388 | 389 | 392 | 395 | 400 | 402 | 419 |
| k | g | d | z | s | b | p | c | t | a | x | f | e | l | y | u | j | v | o | r | w | q | n | h | i | m | k |
| 356 | 357 | 359 | 361 | 361 | 363 | 364 | 367 | 370 | 371 | 372 | 372 | 372 | 376 | 377 | 377 | 379 | 381 | 382 | 382 | 383 | 387 | 396 | 402 | 411 |
| l | s | d | m | g | r | n | v | q | f | k | a | b | c | o | p | u | w | e | z | j | y | t | h | x | i | l |
| 346 | 351 | 358 | 359 | 361 | 368 | 369 | 370 | 371 | 372 | 375 | 375 | 378 | 378 | 379 | 382 | 383 | 386 | 386 | 392 | 394 | 396 | 408 | 414 | 428 |
| m | j | p | o | l | s | c | r | b | i | f | y | u | z | g | v | t | a | d | x | n | h | q | e | w | k | m |
| 340 | 349 | 349 | 358 | 359 | 360 | 360 | 360 | 364 | 364 | 364 | 365 | 366 | 367 | 368 | 372 | 374 | 374 | 376 | 383 | 384 | 385 | 387 | 407 | 411 |
| n | j | i | x | l | b | t | c | z | o | d | a | m | f | v | k | p | u | w | s | r | q | y | h | g | e | n |
| 351 | 357 | 365 | 368 | 372 | 374 | 375 | 375 | 378 | 379 | 381 | 383 | 384 | 385 | 387 | 387 | 389 | 393 | 395 | 399 | 401 | 405 | 410 | 411 | 416 |
| o | f | m | y | p | i | h | s | l | n | x | w | t | k | q | r | d | v | u | j | z | e | c | a | b | g | o |
| 336 | 349 | 350 | 363 | 365 | 371 | 376 | 378 | 378 | 379 | 379 | 380 | 381 | 386 | 388 | 388 | 388 | 392 | 395 | 403 | 406 | 406 | 406 | 406 | 415 |
| p | s | r | z | m | y | f | c | w | o | k | a | x | i | h | j | l | v | n | u | e | t | d | q | b | g | p |
| 339 | 346 | 347 | 349 | 357 | 359 | 361 | 362 | 363 | 363 | 368 | 369 | 375 | 376 | 378 | 379 | 383 | 387 | 390 | 391 | 394 | 404 | 411 | 424 | 430 |
|  | *Table 9 Continued.* | | | | | | | | | | | | | | | | | | | | | | | | |  |
| Letter | | most dissimilar | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | most similar | |  | Letter |
| q | y | j | f | i | h | s | l | x | e | z | v | a | r | k | t | u | m | o | c | n | w | p | b | d | g | q |
| 339 | 351 | 365 | 366 | 369 | 370 | 370 | 372 | 373 | 373 | 373 | 377 | 378 | 383 | 385 | 385 | 385 | 386 | 397 | 401 | 409 | 411 | 419 | 422 | 427 |
| r | p | c | m | l | d | b | u | y | e | s | j | w | q | a | k | o | t | i | v | h | n | z | g | x | f | r |
| 346 | 349 | 360 | 361 | 362 | 362 | 366 | 367 | 369 | 369 | 370 | 374 | 378 | 382 | 382 | 388 | 388 | 393 | 394 | 398 | 399 | 406 | 407 | 407 | 414 |
| s | y | p | l | v | m | t | k | r | q | z | b | o | f | d | j | w | i | a | u | n | c | x | g | e | h | s |
| 325 | 339 | 346 | 348 | 359 | 360 | 361 | 369 | 370 | 373 | 374 | 376 | 383 | 384 | 385 | 385 | 386 | 387 | 393 | 395 | 396 | 399 | 405 | 410 | 418 |
| t | b | g | s | a | h | f | k | m | n | o | d | v | j | e | q | u | c | i | r | z | p | w | l | x | y | t |
| 350 | 354 | 360 | 360 | 362 | 363 | 367 | 372 | 374 | 380 | 381 | 382 | 383 | 384 | 385 | 386 | 386 | 387 | 388 | 392 | 394 | 394 | 396 | 397 | 401 |
| u | g | w | m | b | r | z | x | c | k | f | l | e | q | j | t | d | n | p | o | s | i | y | a | h | v | u |
| 342 | 364 | 365 | 365 | 366 | 367 | 372 | 377 | 377 | 382 | 382 | 382 | 385 | 385 | 386 | 387 | 389 | 390 | 392 | 393 | 393 | 393 | 394 | 397 | 408 |
| v | s | e | g | b | m | l | w | q | f | j | c | a | i | k | t | p | n | o | d | r | x | h | u | z | y | v |
| 348 | 350 | 356 | 364 | 368 | 369 | 372 | 373 | 375 | 375 | 376 | 378 | 378 | 379 | 382 | 383 | 385 | 388 | 392 | 394 | 395 | 399 | 408 | 415 | 421 |
| w | b | j | p | d | u | g | z | a | e | v | r | i | f | o | x | k | l | s | h | y | n | t | c | m | q | w |
| 351 | 354 | 362 | 362 | 364 | 365 | 367 | 367 | 370 | 372 | 374 | 375 | 375 | 379 | 382 | 382 | 383 | 385 | 387 | 393 | 393 | 394 | 395 | 407 | 409 |
| x | j | g | a | b | y | n | f | p | k | q | i | u | e | m | o | w | z | d | h | v | t | s | r | l | c | x |
| 353 | 362 | 363 | 363 | 363 | 365 | 366 | 369 | 371 | 372 | 372 | 372 | 376 | 376 | 379 | 382 | 382 | 383 | 388 | 395 | 397 | 399 | 407 | 414 | 431 |
| y | s | e | q | o | p | d | h | a | x | m | r | b | c | k | g | z | i | w | u | l | f | j | t | n | v | y |
| 325 | 337 | 339 | 350 | 357 | 358 | 358 | 360 | 363 | 364 | 367 | 372 | 374 | 376 | 376 | 387 | 388 | 393 | 393 | 394 | 400 | 400 | 401 | 405 | 421 |
| z | p | j | k | d | m | g | w | u | i | e | q | s | c | n | b | x | f | l | y | a | t | h | o | r | v | z |
| 347 | 356 | 359 | 364 | 366 | 366 | 367 | 367 | 367 | 371 | 373 | 373 | 375 | 375 | 375 | 382 | 382 | 386 | 387 | 390 | 392 | 394 | 403 | 406 | 415 |

**6. APPENDIX B: FIGURES**

*Figure 1. A depiction of the boundary technique. The \* symbol represents the position of the eye whereas the | represents the invisible boundary.*

*Figure 2. A possible hierarchy of word processing stages during reading.*

*Figure 3. (A) The Glenmore model (Reilly & Radach, 2006). (B) The word decoding unit of the Glenmore model.*

*Figure 4. Stimulus presentation used in the letter confusability pilot study. A correct response would be to move the eyes to the letter ‘a’ on the left.*

****

*Figure 5. Mean values for first fixation duration by shape and similarity in Experiment 1. Error bars represent one standard deviation from the mean. Please note the scale of the y axis does not start at zero for ease of examining small differences.*

****

*Figure 6. Mean values for single fixation duration by shape and similarity in Experiment 1. Error bars represent one standard deviation from the mean. Please note the scale of the y axis does not start at zero for ease of examining small differences.*

****

*Figure 7. Mean values for gaze duration by shape and similarity in Experiment 1. Error bars represent one standard deviation from the mean. Please note the scale of the y axis does not start at zero for ease of examining small differences.*

****

*Figure 8. Mean values for landing position by shape and similarity in Experiment 1. Error bars represent one standard deviation from the mean. Please note the scale of the y axis does not start at zero for ease of examining small differences.*

****

*Figure 9. Mean values for incoming saccade size by shape and similarity in Experiment 1. Error bars represent one standard deviation from the mean. Please note the scale of the y axis does not start at zero for ease of examining small differences.*

****

*Figure 10. Mean values for first fixation duration by shape and context in Experiment 2. Error bars represent one standard deviation from the mean.*

****

*Figure 11. Mean values for single fixation duration by shape and context in Experiment 2. Error bars represent one standard deviation from the mean.*

****

*Figure 12. Mean values for gaze duration by shape and context in Experiment 2. Error bars represent one standard deviation from the mean.*

****

*Figure 13. Mean values for landing position by shape and context in Experiment 2. Error bars represent one standard deviation from the mean.*



*Figure 14. Mean values for fixation probability by shape and context in Experiment 2. Error bars represent one standard deviation from the mean.*

**7. APPENDIX C: Meta-Analysis Method and Results Report**

**7.1. Method**

**7.1.1. Search Strategy and Study Selection.** In order to locate potential studies for inclusion in this synthesis, electronic databases provided by the Florida State University (FSU) Library system were consulted. The following databases were accessed: Google Scholar, ISI Web of Science, PsychInfo, PsychArticles, and Dissertation Abstracts. Due to the nature of the experimental reading research literature, only a broad search was necessary. A search of electronic databases for both published and nonpublished documents, created between 1979 (the date the boundary technique was created) and 2009, included the following terms: eye movements, reading, and boundary. Searching for each item individually yielded results of over 200,000 articles. After combining the search items (i.e., eye movements AND reading AND boundary), only sixteen articles were returned (see Table 10). Of those sixteen articles, five were also included in the Dissertation Abstract database. In these cases, the published article was included as opposed to the dissertation document. No additional items were included from the Dissertation Abstracts database. Eighty-eight articles were identified from a search of the primary articles’ reference sections. This yielded 104 articles under consideration for this meta-analysis.

It is common, in meta-analysis, to gather a large amount of data (relevant articles) and then to prune away the majority of those to get to the center of a theoretical issue. This is equivalent to employing strict control in laboratory experiments; by controlling as many extraneous factors as possible, the inferences made based on laboratory data collection are taken with more weight. Equivalently, in meta-analysis, removing unrelated or only distantly related studies strengthens the inferences that can be made after the analysis. Thus, in order to answer the specific question posed for the meta-analysis, it was necessary to impose stringent exclusionary criteria. Inclusionary and exclusionary criteria for this synthesis are presented in Table 11.

Of the 104 articles identified, four studies were excluded because they did not provide enough information for effect size or pooled standard deviation calculations (see Effect Size Derivation below). Seven studies used single word reading instead of continuous single sentence

*Table 10. Summary of articles searched.*

|  |  |
| --- | --- |
| *Search Term* | *Number* |
| Eye Movements | 16,802 |
| Reading | > 100,000 |
| Boundary | > 100,000 |
| Eye Movements  + Reading | 5,796 |
| Eye Movements  + Boundary | 69 |
| Eye Movements  + Boundary  + Reading | 16 |
| Additional Articles from References | 88 |
| Total Articles Identified | 104 |

reading and were excluded. Of the remaining studies 97, a total of 74 were excluded because of considerations related to mask type. This decision was straightforward in most cases, where mask types other than random letters were used, including homophones and semantically related words, or the gaze contingent manipulation did not correspond to the original boundary technique (e.g., orthographic priming).

*Table 11. Inclusionary and exclusionary criteria for studies in the meta-analysis.*

|  |  |
| --- | --- |
| **Inclusionary Criteria** | **Exclusionary Criteria** |
| Eye movements as a dependent measure | Any other dependent measure |
| Boundary technique used | Any non-gaze-contingent displays |
| Single sentences | Single words, several words, partial sentences, or passages |
| Masking | Masks not used |
| Masks violate word shape (length masks, random letter masks) | Masks where word shape is maintained (semantic, phonological, or other mask) |
| Effect sizes able to be calculated via means and standard deviation/standard | No means or standard deviations/errors reported |

Furthermore, a decision was made to exclude twelve more studies, claiming to use “random” letter masks, because the first letter, or even the first few letters of the masking string were identical to the letters of the target word. In these cases, the possibility cannot be excluded that substantial useful information was obtained from the initial letters of the initial parafoveal string, thus they are not a true test random masking (including the possibility of word shape disruption). Seven studies consisting of 11 experiments and 47 effect sizes remained in the synthesis after all exclusionary criteria had been applied.

**7.1.2. Study Characteristics and Variables.** All studies included single sentece reading experiments. All studies used college students or community members, who are assumed to be skilled readers, as participants except Chance, Rayner, and Well (2005), who also included readers with low reading skill. In that case, only the data for skilled readers was extracted. No explicit age information could be extracted from the studies.

The majority of eye-tracking outcomes measured were generally the same from study to study and include: first fixation duration (fifteen effect sizes in: Chance et al., 2005; Inhoff, Eiter, Radach, & Juhasz, 2003; Juhasz, White, Liversedge, & Rayner, 2008; McDonald, 2006; Miller, Juhasz, & Rayner, 2006; Starr & Inhoff, 2004); gaze duration (fifteen effect sizes in: Chance et al., 2005; Inhoff et al., 2003; Juhasz et al., 2008; McDonald, 2006; Miller et al., 2006; Starr & Inhoff, 2004); and total reading time (nine effect sizes in: Inhoff et al., 2003; Juhasz et al., 2008; McDonald, 2006). Only a handful of studies reported other information such as landing postion (seven effect sizes in: Juhasz et al., 2008; Miller et al., 2006), rereading rate (one effect size in: Starr & Inhoff, 2004), and skipping probability (one effect size in: McDonald, 2006).

**7.1.3. Data Evaluation.** Several predictors were identified as having a potential impact on effect size including:

1. the year the study was published, as more recent technological and methodological innovations may impact findings;
2. the type of eye tracking system used can impact findings, as each has a different resolution and accuracy;
3. whether Keith Rayner appeared as an author was used as a theoretical position proxy (i.e., serial attention model vs. attentional gradient model);
4. the number of experimental sentences that were read;
5. the length of the target word, as longer words may show less of a disruptive effect from masking;
6. the type of mask used;
7. the percent of data excluded from the study; and
8. specific outcomes which may be differentially effected by masking (e.g., first fixation duration vs. gaze duration).

Outcomes of interest were dependent measures that can be extracted from eye tracking data. First fixation duration (FFD) is the length of time the word is fixated upon landing in the word for the first time. Any additional fixations within that word, but before leaving, are included in a measure called gaze duration (GD). Any additional fixations made after leaving the word for the first time are included in the total reading time (TRT). Most studies included measurements of FFD, GD, and TRT, however, some studies included additional measures such as the position the eyes landed in the word upon entering it, landing position (LP), or the likelihood that a word would be skipped during reading, skipping probability (SP).

**7.1.4. Effect Size Derivation.** Continuous measures are those most commonly reported for reading research. Thus, standardized mean differences (*d*) were calculated for each comparison. Effect sizes were calculated using both the mean and the standard deviation. Because of dependency issues associated with the variables recorded during reading research (e.g., first fixation duration and gaze duration are components of total reading time), Becker’s (1988) techniques for addressing dependency in the both effect size variance were used.

In the following equations, *Violated* is defined as a violation of word shape through masking whereas *Maintained* is defined as word shape being maintained (usually the control condition). The biased effect size was calculated as

where

All effect sizes (*d*) were then modified to include corrections for small sample sizes. The formula used for calculating the unbiased effect size was where

,

,

and

*.*

In Becker’s (1988) technique to accommodate dependency among outcome variables, a correlation coefficient is specified. Because no information was included about the relatedness of measures in any studies, a conservative estimate of 0.5 was used as the correlation coefficient for all studies. (In an examination of a set of our own data, correlations were between 0.4 and 0.6, thus making the value of 0.5 reasonable in this context). While this may seem an inadequate way to address the dependency seen in repeated measure designs, there is no technique, to date, that has been developed to address this issue. The Becker’s (1988) formula for estimating the sample variance is

.

Using these methods, a total of 47 effect sizes were calculated from the 7 studies included in the synthesis (see Table 3). Note that the rereading rate effect size in the Starr and Inhoff (2004) study was suspiciously large (d = 2.64), being more than three standard deviations over the mean (3SD = 1.68), thus it was removed from all further analyses.

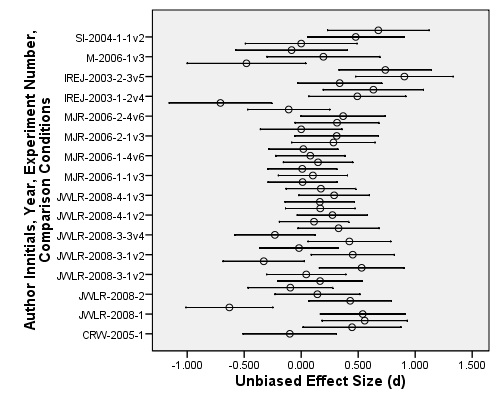
**7.1.5. Analytic Strategy.** Weighted tests of homogeneity were conducted. Such analyses are essential for determining if there is a consistent pattern of results, in which case a fixed effects model should be used, or if the results have different patterns, a random-effects model is the most appropriate to employ (Lipsey & Wilson, 2001). If there is significant variance left unaccounted after a random effects analysis, a mixed model and analyses was conducted. Predictor variables were evaluated using both categorical (ANOVA-like) and continuous (regression-like) analyses. All analyses were computed using PASW/SPSS.

**7.2. Results**

**7.2.1. Fixed-Effects Analyses.** In a fixed-effects analysis, the distribution of effect sizes around the mean is examined to see if the individual effect sizes vary only due to sampling error (Lipsey & Wilson, 2001). The null hypothesis for the test of homogeneity across all 46 effect sizes was , where is the population effect size in sample *i*. The chi-square test, Q (df = 45), was significant (p < 0.001). Because of the results of the test of homogeneity, the null hypothesis was rejected, denoting the effect sizes included in this

*Figure 15. 95% Fixed-Effects Confidence Intervals.*

synthesis do not stem from the same population and that the variations seen are not due purely to sampling error. In such circumstances, it is appropriate to use a random-effects model.

**7.2.2. Random-Effects Analyses.** Random-effects models are similar to fixed-effects models except that additional variability is included on top of sampling error alone. In order to compute a test of the null hypothesis where , or the average population mean effect size is zero, a Z test was conducted using the more conservative measure of sample variance (SVAR = 0.08) rather than the weighted sample variance (QVAR = 0.06). The Z test, Z = 3.73, was significant having surpassed the critical value of 1.96. The results of this test indicate that it is appropriate to reject the null hypothesis.

**7.2.3. Categorical Analyses.** Categorical analyses were conducted examining the type of eye tracking outcome measure, the type of mask employed, the type of tracker used, the number of conditions in the study and whether Rayner was present as an author (as a proxy for theoretical position; see Table 7.4). The test of the null hypothesis resulted in rejecting the null as there was significant between-group heterogeneity (). Results of the

*Figure 16. 95% Random-Effects Confidence Intervals.*

categorical analyses are presented below and in Table 5. represents the between-group heterogeneity test using a chi-square distribution with k-1 degrees of freedom where k is the number of effect sizes within the group. The value of the Birge Ratio, represented by , is the ratio between the *Q* value and the degrees of freedom. This allows direct comparison between categories composed of diffenent k values. Birge Ratios close to 1 indicate that there is little to no variability in scores.

Neither type of mask nor conditions accounted for significant between-groups variation. Eye tracking measure outcome accounted for the largest proportion of between-group variation ( = 43.61, p < 0.001, = 2.75). There was also significant within group variation seen for FFD and LP. The variation seen for LP may be due to the variation in the length of the mask. FFD variation is likely due to disruptions in early orthographic due to masking. This sort of as random letter masks interfere with orthographic processing, one of the earliest steps in lexical processing.

*Figure 17. 95% Random-Effects Confidence Intervals for Eye Tracking Outcome Measure. FFD = First Fixation Duration. GD = Gaze Duration. TRT = Total Reading Time. LP = Landing Position. SP = Skipping Probability.*

 The type of eye tracker used to record oculomotor data accounted for some of the between groups variation ( = 4.33, p < 0.05, = 2.63). The significant within-group variation was attributed to the Dual-Purkinje eye tracker. This is a surprising result as significant between-group variation was expected while within-group variation was not. Although the EyeLink system was not a significant source of within-group variation, it should be noted that only four effect sizes were associated with it. Including more studies using the EyeLink system may reveal equal heterogeneity in that group as well.

Rayner appearing as an author also accounted for a significant amount of between-group variation ( = 5.36, p < 0.05, = 2.63). There was also significant within-group heterogeneity when Rayner was present as an author as well as when he was not. While significant between-group variability was expected, significant within-group heterogeneity was not. This may be due to the presence of close Rayner collaborators and students (e.g, Barbara Juhasz), which were found on ten out of twelve of the Rayner absent effect sizes. Future synthesis in this area should use a specific theoretical position code to avoid confounding due to frequent collaboration issues. Despite the potential difficulties with this variable, Rayner’s presence as an author of an article remains a significant predictor of between-group heterogeneity.

**7.2.4. Continuous Analyses.** Weighted random-effects regression analyses were conducted for continuous variables and dummy coded variables, including: publication year (year); mask type (mask); target word length (length); type of tracker (tracker); amount of data excluded (excluded); Rayner present as an author (Rayner); the number of experimental

*Figure 18. 95% Random-Effects Confidence Intervals for Type of Tracker.*

*Table 12. Random-Effects Categorical Analysis.*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | 95% CI for *d* | |  | | | |  |
| Aggregation | k | d | Lower | Upper | | QW | QB | Birge Ratio | |
| *Eye Tracking Outcome Measure* |  |  |  |  | |  | *43.61\*\*\** | *2.75* | |
| First Fixation Duration (FFD) | 15 | 0.117 | 0.02 | 0.21 | | 31.90\*\* |  |  | |
| Gaze Duration (GD) | 15 | 0.33 | 0.24 | 0.43 | | 23.09 |  |  | |
| Total Reading Time (TRT) | 9 | 0.32 | 0.20 | 0.44 | | 12.037 |  |  | |
| Landing Position (LP) | 7 | -0.16 | -0.29 | -0.02 | | 15.87\* |  |  | |
| Skipping Probability (SP) | 1 | 0.00 | - | - | | 0.000 |  |  | |
|  |  |  |  |  | |  |  |  | |
| *Type of Mask* |  |  |  |  | |  | *2.09* | *2.63* | |
| Random Letter | 20 | 0.13 | 0.05 | 0.21 | | 26.31 |  |  | |
| Length | 27 | 0.21 | 0.14 | 0.28 | | 92.50\*\*\* |  |  | |
|  |  |  |  |  | |  |  |  | |
| *Type of Tracker* |  |  |  |  | |  | *4.33\** | *2.63* | |
| EyeLink | 4 | -0.08 | -0.34 | 0.17 | | 3.57 |  |  | |
| Dual Purkinje | 43 | 0.19 | 0.13 | 0.24 | | 113.00\*\*\* |  |  | |
|  |  |  |  |  | |  |  |  | |
| *Number of Conditions* |  |  |  |  | |  | *0.528* | *2.63* | |
| 4 | 18 | 0.17 | 0.16 | 0.27 | | 67.84\*\*\* |  |  | |
| 6 | 21 | 0.19 | 0.12 | 0.27 | | 30.01 |  |  | |
| 8 | 8 | 0.14 | 0.01 | 0.27 | | 22.53\*\* |  |  | |
|  |  |  |  |  | |  |  |  | |
| *Rayner an Author?* |  |  |  |  | |  | *5.36\** | *2.63* | |
| Yes | 35 | 0.15 | 0.09 | 0.21 | | 64.92\*\* |  |  | |
| No | 12 | 0.31 | 0.18 | 0.44 | | 50.63\*\*\* |  |  | |

\* = p < 0.05; \*\* = p < 0.01; \*\*\* = p < 0.001

sentences in the study (ExpSent); and the number of conditions in the experiment (conditions). For each analysis, the residuals were examined for outliers. As discussed by Pedhazur (1997), if any standardized residuals had an absolute value greater than 2, those cases were excluded from further analyses. Upon eliminating any outliers, the regression analyses were recomputed. Table 5 displays significant predictors found under the random-effects model. While all the predictors were significant, error was also significant indicating that the random-effects model is a poor fit for this data and that the unexplained variability is not due to sampling error alone. In order to better account for the significant unexplained variability, a mixed model analysis was then conducted.

*Table 13. Random-Effects Analyses.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Predictor** | **Outliers** | **QModel** | **df** | **p-value** | **QError** | **df** | **p-value** | **Birge Ratio** |
| Excluded | 3 | 15.20 | 1 | < 0.001 | 63.74 | 42 | < 0.025 | 1.88 |
| Length | 1 | 5.26 | 1 | < 0.025 | 97.30 | 44 | < 0.001 | 2.33 |
| Rayner | 2 | 17.33 | 1 | < 0.001 | 82.08 | 43 | < 0.001 | 2.31 |
| Tracker | 0 | 4.33 | 1 | < 0.05 | 116.58 | 45 | < 0.001 | 2.69 |
| Year | 1 | 19.72 | 1 | < 0.001 | 84.125 | 44 | < 0.001 | 2.36 |

*Table 14.* *Mixed-Model Analyses*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Predictor** | **Outliers** | **QModel** | **df** | **p-value** | **QError** | **df** | **p-value** | **Birge Ratio** |
| Excluded | 3 | 6.39 | 1 | < 0.025 | 45.32 | 42 | ns | 1.23 |
| Length | 3 | 1.96 | 1 | ns | 47.42 | 40 | ns | 1.23 |
| Rayner | 4 | 6.15 | 1 | < 0.025 | 40.33 | 41 | ns | 1.13 |
| Tracker | 3 | 3.53 | 1 | ns | 78.99 | 45 | < 0.01 | 1.83 |
| Year | 2 | 11.59 | 1 | < 0.001 | 47.51 | 43 | ns | 1.37 |

Under a mixed model, there is both a fixed and random variance component (please note that this term is not the same as the “mixed model” used in ANOVA) as represented by

. A new variance and weights were calculated by subtracting the mean variance from the mean square error then adding it to the variance for each effect size in order to calculate the component of the mixed-model. It is necessary to compute new variances and weights for a mixed model as those from the random model are often too large. By adding the component to the model, it represents the random component to the model and the component reflects sampling error. Following the same procedure as the above, new regression analyses were computed on the predictors that accounted for a significant amount of variance under the random-effects model.

As can be seen in Table X, the mixed model provides a much better fit for the data than either the fixed- or random-effects models. Four of the predictors from the random-effects model remained significant under the mixed model: amount of data excluded (Excluded); Rayner as an author (Rayner); type of tracker (Tracker); and year of publication (Year). For each of the significant predictors, none of the error terms reaches significance signifying that all the variation seen in this model can be accounted for by sampling error alone.

Because Rayner and Year were highly correlated (r = 0.78, p < 0.01), they could not be included in a model together due to assumptions related to multicolinearity, which was also true of Excluded and Tracker (r = -0.82, p < 0.01). In choosing between pairs of correlated predictors to include in the model, standardized betas were examined (see Table X).

*Table 15.* *Mixed Model Coefficient Information.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Predictor** | **β** | **SE** | **S β** | **Z** | **p-value** |
| Excluded | -1.27 | 0.50 | -0.35 | -2.53 | < 0.01 |
| Rayner | -0.24 | 0.10 | -0.36 | -2.48 | < 0.01 |
| Tracker | 0.28 | 0.15 | 0.21 | 1.88 | ns |
| Year | -0.07 | 0.02 | -0.44 | -3.41 | < 0.001 |

In general, the beta-values demonstrate the relationship between the effect sizes and the predictor variable included in the regression analysis; specifically, they describe the amount of change in effect size for a one unit change in the predictor variable. It is valuable to look at standardized betas as different scales can be directly compared via standard deviation units (Field, 2009). Between Year (Sβ = -0.44) and Rayner (Sβ = -0.36), Year was chosen as the stronger predictor to be included in the multiple regression model. Of Excluded (Sβ = -0.35) and Tracker (Sβ = -0.21), Excluded was chosen as the stronger predictor. It should be noted that the majority of effect sizes came from the DualPurkije tracker with only four effect sizes coming from the EyeLink system, making Tracker less useful of a predictor than Excluded. Thus, only Excluded and Year could be included in a regression model together. One important thing to note is that both Excluded and Year have a negative relationship with effect sizes.

The results of the mixed model multiple regression showed better fit than the random-effects multiple regression. The original model including both Year and Excluded was di = 134.072 – 0.067Yeari – 1.667Excludedi + ei, whereas the mixed model fitted equation was di = 148.185 – 0.74Yeari – 1.528Excludedi+ ui + ei. While the beta value for Excluded changed only slightly, there was a dramatic change in the contribution of Year as a predictor under the mixed model.

**7.3. Discussion**

The presented meta-analysis was conducted in order to investigate mask type as a proxy for word shape violations and to determine how much word shape violation interferes with lexical processing during continuous sentence reading, thus providing an essential foundation from conflicting results from the literature for an explicit experimental examination of word shape. In meta-analysis, it is appropriate to first test a fixed-effects model, which assumes that any variation seen in the data is due to sampling error. Upon testing for homogeneity, it was determined that the fixed-effects model was a poor fit. Next, a random-effects model was tested. Random-effects models assume that any variation seen is due solely to predictor variables. After conducting tests of the random-effects model, it was determined that it was also a poor fit for the data. Finally, a mixed model was used to evaluate the data. In a mixed model, there is both random and fixed component included, meaning that some variation was due to sampling error and some to predictor variables. The mixed model provided the best fit for the data seen. Under the mixed model, four predictor variables were significant (Excluded, Rayner, Tracker, and Year), however, due to issues of multicolinearity, only Excluded and Year were included as significant predictors in the final mixed model.

It was expected that the effect size data would be best accounted for by a random-effects model as it is highly unlikely that much sampling error would occur considering the high accuracy and precision of modern eye trackers. Because the mixed-model provided the best fit, a conclusion about the predictors being inadequate must be made. As this was the first stringent and comprehensive meta-analysis utilizing recent techniques to account for dependency, only intuition and deduction were used to identify possible predictors for the synthesis.

The amount of data excluded from a study was a significant predictor in the mixed model. Those studies with the highest amounts of excluded data (min = 7%, max = 38%, mean = 17%, SD = 8%) were also those with the lowest effect sizes (mean effect size for 38% excluded = -0.09; mean effect size for 7% excluded = 0.54). One explanation is that the trials being excluded were those showing an effect. Alternatively, there could have been a problem with the method (e.g., too difficult or challenging) that resulted in such a large amount of data being discarded. Whatever the cause, it is clear that a large amount of discarded trials is not advantageous in reading research.

Considering the amount of technological innovations associated with eye tracker development, there have been no large technological or methodological changes that can be attributed to the Year predictor. Interestingly, Year and Rayner are highly related. From examining scatterplots of the data, it seems as if a pattern of results from one lab was followed by responses from the Rayner lab, leaving the Rayner lab articles as the most recent ones, thus closely tying Year and Rayner. Furthermore, the majority of the effect sizes when Rayner is present as an author are also significantly lower than when Rayner is not present. While being a relatively poor theoretical proxy, this predictor significance provides evidence for the necessity of coding for theoretical viewpoint in any further meta-analysis of the reading literature, especially when parafoveal processing is being addressed.

The original question of mask type and word shape violations could not be addressed as expected. To our surprise, the studies synthesized in the meta-analysis never included a condition in which a similar random letter mask was employed in comparison to a dissimilar random letter mask. As there were no similar and dissimilar random letter masks to compare, it can be maintained as a reasonable hypothesis that length and random letter masks do interfere with lexical processing especially in the early stages of processing as reflected by increases in FFD (first fixation duration) when a mask is present. As length and random letter masks do change the shape of the word, the results imply that word shape violations are contributing to the effects seen. In conclusion, as it is impossible to conclude from existing data precisely in what way word shape violations are affecting continuous reading, empirical research on this question is clearly needed.

**8. Appendix D: Letter Similarity Method and Results Report**

**8.1. Method**

**8.1.1. Participants.** Nineteen Florida State University undergraduate students served as participants and received course credit for their participation. All participants had normal or corrected to normal vision.

**8.1.2. Materials and Design.** A repeated-measures design was used in this experiment such that each participant was exposed to 1300 trials during the 90-minute experimental session. Each letter of the alphabet twice served as both the central cue and as a peripheral distractor, including every possible letter combination (e.g., both “a” as the cue and “b” as the distractor as well as “b” as the cue and “a” as the distractor). The cue was situated in the center of the screen with the distractor letter either five letter spaces to the left or the right of the central cue. At a distance of 59 cm and with a 12-point font size, the distance between the cue and distractor was 1.33 degrees of visual angle, with each letter/character subtending 0.33 degrees.

**8.1.3. Apparatus.** All eye-tracking measurements were obtained using the EyeLink 2000 measurement system (SR Research Ltd., www.eyelinkinfo.com, Mississauga, Ontario, Canada). This system employs infrared video-based tracking technology to compute the center and size of the pupils in both eyes as well as the corneal reflection. It has a relative spatial resolution of 0.02° RMS at 2000 Hz, and an absolute accuracy of below 0.25, depending on calibration (see McConkie, 1981, for details on calibration; Inhoff & Radach, 1998, for measurement issues). The EyeLink 2000 allows accurate tracking using a desk-mounted feature with removes the requirement of attaching parts to a participant’s head thus allowing for maximally unobtrusive data collection over extended periods. All participants were tested on a Dell computer, used to control stimulus timing and record participant’s eye movements, saccadic accuracy, saccadic latency, saccadic trajectory and deviations, amplitude, and gain with a 22” color VGA monitor (1024 by 768 at 120Hz) on which the stimuli were presented. Eye movements were sampled twice per millisecond.

Although it is possible to record binocular eye movements with the EyeLink 2000, data was recorded from the right eye only. The on-line saccade detector of the eye tracking system was set to detect saccades with an amplitude of 0.15° or greater, using an acceleration threshold of 8000°/sec² and a velocity threshold of 30°/sec. All stimuli were presented in 12-point Courier New font face. The distance between the display screen and the reader’s eyes was set to 59 cm, thus causing each letter to subtended 0.33˚ of visual angle laterally. During the experiment, the head was stabilized by means of a chin rest.

**8.1.4. Procedure.** Participants were tested individually and were seated in a comfortable chair 59 cm from a 22-inch CRT monitor. A chin and forehead rest were used to stabilize the head, and participants were asked to avoid making any large head movements. The eye-tracking camera was then adjusted until a clear corneal reflection was present. After setting the threshold for detecting the pupil, the system was calibrated. Calibration was performed before each block of trials and a drift correction before each trial to ensure high data quality with mean average deviation in validation trials under 0.35°. Once the calibration procedure was successfully completed, the experimental session started.

*Figure 19. Stimulus presentation for similarity experiment.*

Each trial began with the presentation of masking stimuli composed of hash marks (see Figure 8.1). After 250 ms, the masks were removed and the stimuli were presented. Participants were given up to 1000 ms to discriminate between the parafoveal stimuli while maintaining fixation on the central cue. Successful completion of a trial required participants to execute a saccade to the parafoveal letter that was identical to the central cue with the exception of the hash marks that laterally masked the target stimulus.

**8.2. Results**

**D.2.1. Nonlinear Reduction—Trial Effects.** In order to evaluate the pattern of responding over time with a view to detecting either practice or fatigue effects, the saccade latencies for correct responses were analyzed using a linear mixed effects model (Baayen, 2006). Subject and items were treated as random effects. Items were assumed to be the unique target-distractor pairs (650 in total). The model giving the best to the data included random slopes for both subject and items.

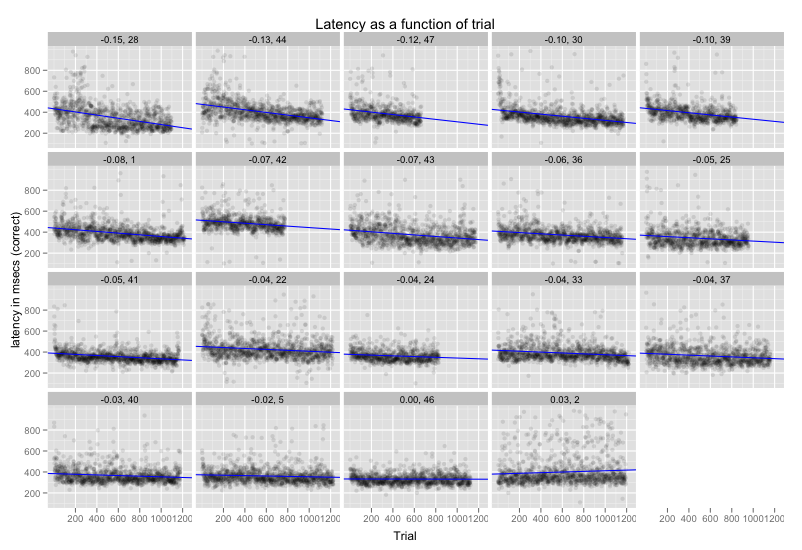
model3.tiff

*Table 16. Jack’s table of happy wonderfulness ;-)*

As can be seen from Table X above, there is a significant effect due to trial (|t|>2). The main source of the effect is a slight negative slope over time due to practice effects. There are, in addition, significant individual variations across both subjects and items. A model including

random slopes for both items and subjects has a significantly greater likelihood than ones that omit them (χ2 = 36.39; p < .001).

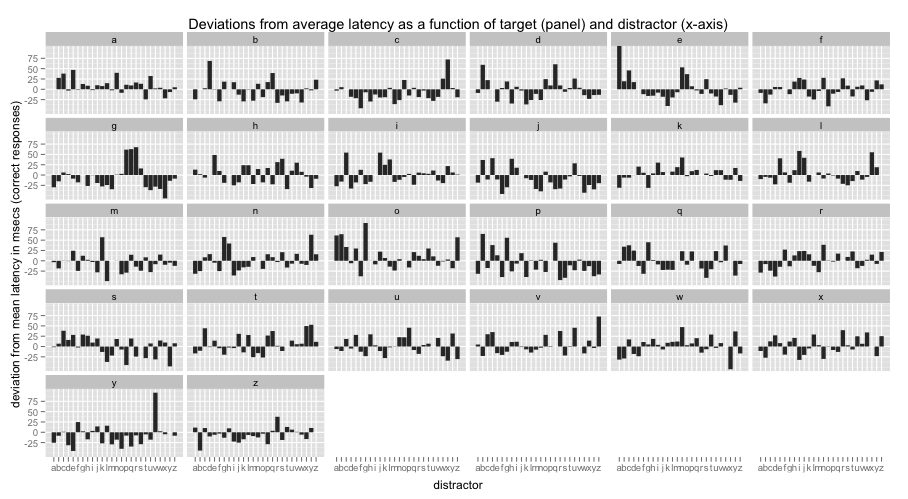
The source of the subject variation over trials can be seen in Figure 8.2, which graphs each data point for every subject and trial. The plots are ordered by size of the subject slopes (shown as a blue line). The earlier panels show a relatively steep negative slope which can be accounted for by stronger practice effects for a subset of the subjects. Note that the data-points are graphed semi-transparently so that over-plotting will show concentrations of data points as darker



*Figure 20. Saccade latencies as a function of trial reflecting individual differences.*

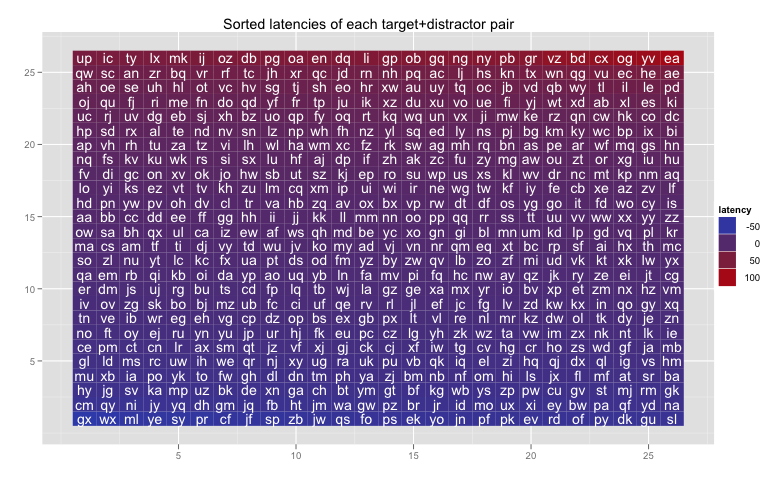
regions.

**8.2.2. Nonlinear Reduction—Item Effects.** An interesting feature revealed by the data here is the asymmetric nature of the response latencies. For example, if the target letter is o and the distractor is g the subject will have a significantly more delayed response than if it is the way around. This asymmetry holds particularly for o and the class of ascenders and descenders with an o-like body. On the other hand, a and e are close to symmetric in terms of their mutual confusability.



*Figure 21. Saccade latencies by letter pairs as a function of cue and target.*

Figure 8.3 gives the overall picture. Each of the panels represents a target word. The x-axis of the panel is the identity of a distractor ordered from a through z. Note that in the experimental paradigm, the target and distractor could never be the same. So there is never data for target a and distractor a, for example, as a would be the target. The y-axis in each panel represents the average deviation in milliseconds from the mean saccade latency for each subject. Bars above zero indicate a saccade that’s above average in latency for that specific target-distractor pair. Bars below zero, indicate faster than average saccades.

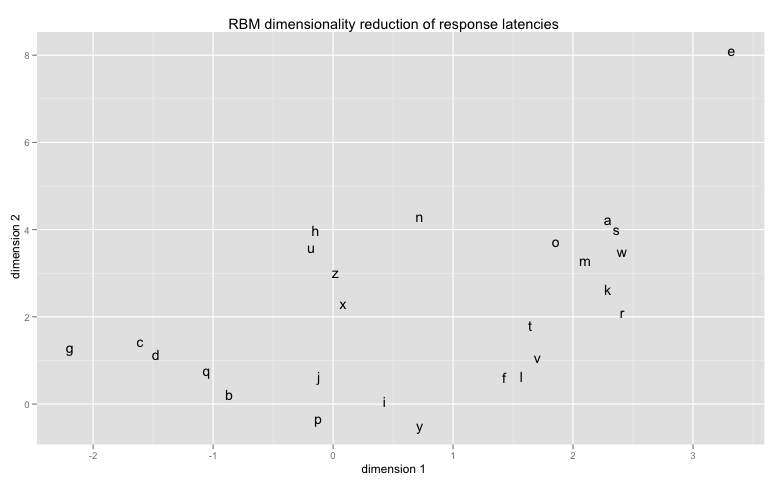


*Figure 22. Parewise saccade latency diagram.*

Figure 8.4 offers a complementary view of the latency deviation data. Each target-distractor pair is ordered from the bottom-left corner to the top right as a function of latency. So the target-distractor pair of gx generates the fastest responses, while ea generates the slowest. The gradation of colour from bottom to top is used to code for latency. Note that only the upper row of the table contains pairs with significant latency elevation. In order to provide a reference point, a line of identical pairs was included. These are set at zero latency deviation. As can be seen, slightly more pairs are below than above this line. These overall patterns are to be expected for a well-designed font, although there is clearly room for improvement. Indeed, the visualisations here could be used as a comparative behavioural metric of font quality as part of a font design process.

One can also see the asymmetry mentioned above evidenced in this graph. Note the two pairs ae and ea inhabit the upper right corner of the grid. In other words, they attract roughly equal latency penalties. In contrast, og slows responses, whereas go is more neutral (see row 16, column 21).

**8.2.3. Dimensionality Reduction.** In order to gain a better overview of the behaviour of each of the letters used in the experiment we used a non-linear dimensionality reduction algorithm referred to as the Restricted Boltzmann Machine (RBM; Hinton et al., 2005). The input to the RBM was a vector of 2x26 elements, where the first 26 elements were latency deviations for the letter as a target to each letter and the second 26 were for the letter as a distractor to each letter. Zero was used for the vector entries involving identical target and distractors. There were 26 of these vectors in all, one for each letter of the alphabet. The RBM mapped these 52D vectors to a 2D space. Because of the non-linear nature of RBM, the two dimensions should not be considered as either latent variables or anything like principle components. Their role is to cluster letters that behave similarly in similar regions of the 2D space.



*Figure 23. 2D representation of Euclidian space.*

An examination of the 2D map in Figure 7 shows some meaningful clustering of letters. we can see that letters of similar spatial frequency tend to cluster together (e.g., f, l, t, v, versus c, d, q, b versus a, s, m, w). Another trend appears to be for clustering of open versus closed letters (e.g., h, u, z, x versus g, d, q, b). Most interesting is how the more “problematic” letters are relatively isolated (e.g., g and e).

**8.3. Conclusion**

The analysis has demonstrated the viability of using saccade latency in a cued target-distractor paradigm as a measure of letter confusability. It has provided a rich picture of visual letter processing, suggesting features of the process that need to be focussed on in the construction of a more realistic model of the low-level visual aspects of reading. For example, openness versus closedness of a letter and their spatial frequency appear to play roles. Moreover, closedness of the type in the letter o seems to have a distinct advantage over other patterns closed when presented in the parafovea. This is evidenced by the og versus go asymmetry, for example.

**9. APPENDIX E: SAMPLE EXPERIMENTAL MATERIALS FROM EXPERIMENT 1**

Target Sentence:

Adam was having trouble meeting the tough **requests** made by his boss at work.

Similar Mask with Word Shape Maintained:

Adam was having trouble meeting the tough **nsgvaebe** made by his boss at work.

Dissimilar Mask with Word Shape Maintained:

Adam was having trouble meeting the tough **erjnrafa** made by his boss at work.

Similar Mask with Word Shape Maintained:

Adam was having trouble meeting the tough **zswvsuxu** made by his boss at work.

Dissimilar Mask with Word Shape Maintained:

Adam was having trouble meeting the tough **erxnrasa** made by his boss at work.

The small family was excited after the quick **adoption/cbaqblar/wlvyfhvu/coaexrar/wrvzsnvx** of the new baby girl.

The botanist gathered a large quantity of green **foliage/kailcpa/tvdhwjr/ncxrcos/ovsnwur** to study later in the lab.

Bells rang wildly after the pious **ritual/nlbvci/ehfnwd/zrxvca/ensnws** ended and the monks left the church.

Students crowded the instructor after the final **workshop/manhefaq/cvebatvy/evezacvz** of the semester had ended.

A sense of tense **hysteria/fjebanlc/tpafrehw/nvuxszrc/cavsrenw** filled the bank patrons as the man waved the gun.

The young boy kicked the squat **plough/qiavpf/ydvnjt/eaavon/zsvnuc** that had become stuck on the tree roots.

The groom experienced a vague **jealousy/yaciavej/qrwdvnap/uscvnvuv/wrwsvnva** as he watched his bride talk to an old pal.

The biologist showed pictures of a hairy **species/eqaslae/ayrmhra/uesorsu/azrmnra** of ape to the students.

Dillon was having trouble making his steep **mortgage/wanbpcpa/ovefjwjr/wazxocos/ovesuwur** payments every month.

The public listened attentively to the smart **rhetoric/nfabanlo/etrfvehm/znsxazro/ecrsvenm** of the politicians.

Leo cleaned the sandy **sponge/eqcrpa/ayvujr/uearos/azvxur** after washing his car at the beach.

The tailor bemoaned the total **shortage/efanbcpa/atvefwjr/unazxcos/acveswur** of silk cloth on the textiles market.

Anna honored the proud **religion/nsilplar/erdhjhvu/zsarorar/ersnunvx** of the settlers by lighting a candle.

Jeanne wished that the swift **daylight/bcjilpfb/lwpdhjtf/ocvaronx/rwasnucs** had not faded so quickly while she painted.

The roommates glared at each other as their toxic **dispute/blcqvba/lhaynfr/cruevxs/rnaznsr** erupted into a fistfight.

The police set up the bogus **exchange/avofcrpa/rnmtwujr/szoncros/rnmcwxur** to soothe the criminals they hoped to catch.

The grill had several enticing fresh **strips/ebnlqe/afehya/uxzreu/ascnza** of meat sizzling on it.

Wendy had difficulty interpreting the vivid **symbols/ejwdaie/apokvda/uvwecau/vaowvsa** on the ancient Mayan ruin.

Students of chess are advised to capture the swift **bishop/dlefaq/khatvy/ervnce/vnacvz** in order to win matches.

The elementary teacher was saddened by the hasty **printing/qnlrblrp/yehufhuj/ezrrxreo/zenxsnxu** on the messy homework.

The imprisoned teen indulged in a brief **fantasy/kcrbcej/twufwap/ncrxcuv/owxswea** from behind the bars of her prison.

The prisoners of war dreaded the cruel **patrols/qcbnaie/ywfevda/ecxzaau/zwsevsa** of the guards during their rounds.

The little girl infected others with her sunny **idealism/lbacilew/hlrxdhao/rcscaruw/nrvwsnao** while skipping around town.

The drill sergeant yelled stern **slogans/eiapcre/advjwua/uacocru/asvuwxa** at the new recruits to raise their spirits.

The old man prefers family visits on quiet **holidays/fcilbcje/tvdhlwpa/naarccvu/cvsnrwea** such as Easter or Christmas.

Jill was dismayed when I lost her thick **pencil/qaroli/yrumhd/esrsra/zrxmns** during art class last week.

The customer was confused by the short **receipt/naoalqb/ermrhyf/zsosrex/ermrnzs** he was given for his large order.

The audience was delighted by the witty **quality/gvcilbj/jnwdhfp/wvcarxv/xnwsnsa** of the comic's performance.

Bobby explained the plan to Jimmy in a husky **whisper/mfleqan/cthvyre/mnruesz/ecnvzre** that the teacher easily heard.

The grapes now had the dried **ridges/nlbpae/ehljra/zrcosu/enrvra** of raisins after being left out all night.

The squire played with his master's rusty **shield/eflnib/athrdl/unrsac/acnrsr** while waiting for him to return.

The crew cheerfully obeyed the orders of the stout **captain/ocqbclr/mwyfwhu/ocexcrn/mwzswnx** because he paid well.

The museum guide explained the brave **history/flcbanj/thvfvep/nruxazv/cnasvea** of the settlers to the young boy.

It was hard for Ted to see the faint **signals/elprcie/ahjuwda/urorcau/anuxwsa** on the radar screen.

The secretary was upset that there was no clear **category/ocbapanj/mwfrjvep/ocxsoazv/mwsruvea** for the file she held.

The squire hid behind the quiet **knight/hrlpfb/buhjtf/msronx/zxnucs** as the king inspected the new recruits.

The new museum curator made sure the dusty **portrait/qanbnclb/yvefewhf/uazxzcrx/zvesewns** was cleaned immediately.

Old Joe wondered if his fluid **capacity/ocqcolbj/mwywmhfp/ocecorxv/mwzwmnsa** would decline drastically with age.

The ninja crouched near the misty **vicinity/ulolrlbj/shmhuhfp/urorerxv/snmnxnsa** of the evil Lord's well defended castle.

Grandma made sure to save the cheap **recipes/nsolqae/ermhyra/zsoresu/ermnzra** for when she was eating all alone.

The scientist inspected the loose **proteins/qnabalre/yevfrhua/ezaxsreu/zevsrnxa** of the damaged viral structure.

**10. APPENDIX F: EXPERIMENTAL MATERIALS FROM EXPERIMENT 2**

Predictable Context Sentence:

Sir Lancelot was strong, muscular, and attractive.

Unpredictable Context Sentence:

The blacksmith was strong, muscular, and attractive.

Target Sentence:

All the young men looked up to the manly **knight/buhjtf/zxnucs** as a role model.

The cultists sacrificed a chicken to their god each morning in a terrible ceremony.

The cultists ate a chicken to each morning for the protein and energy it provided.

Betty had trouble watching their cruel **ritual/ehfnwd/ensnws** when she visited their compound.

The doctor visited Carlos after having his tonsils out this morning.

The doctor told Carlos to rest after a lot after his last visit.

Carlos felt much better after the brief **surgery/anejrep/aneurea** the doctor performed on his throat.

Most of the people riding in the Tour de France have very similar equipment.

The long journey will take many hours of difficult travel.

David got onto the wrong **bicycle/khmpmdr/vnmamsr** after his last rest stop.

Catholicism is a large sect while Voodoo is a minor one.

Hannah's grandmother made sure she went to her meetings every week.

Hannah liked belonging to a major **religion/erdhjhvu/ersnunvu** just like her grandmother did.

Jeff spent three days on the lake without hooking anything on his new pole.

Jeff spent the long weekend in a cabin for some needed woodland relaxation.

He called home and complained about the tough **fishing/thathuj/onacnxu** on his weekend getaway.

The doctor insisted on examining the person dying of cancer.

The reporter insisted on interviewing the irritated man.

He came to talk with the moody **patient/ywfhruf/zwsnrxs** even when it didn't seem like there was hope.

Fred arrived at the surprise party five minutes late and didn't get to yell "Surprise!"

Fred was looking forward to his date with Angelica at the nice restaurant.

He cursed his awful **timing/fhohuj/snonxu** at arriving late and made sure to apologize.

Zack broke his leg in a filthy little town near the jungle.

Zack forgot his suitcase in the taxi when he left on his trip to India.

He dreaded going to the dirty **hospital/tvayhfwd/cvaznsws** that was the only one the town had.

The new lawyer was amazed at his luck in finding cases.

The young businessman tried hard to find work.

He ended up getting forty **clients/mdhrufa/msnrxsa** during his first week in business.

The general hated eating eggs without having some links as well.

The general hated to forgo his cigar after his meals.

He expected to have smoky **sausages/awnawjra/awnawura** along with eggs each morning.

Sometimes Ted had to go without lunch so he could make his high house payment.

Ted has many things that he needs to take care of.

He had a lot of trouble meeting his steep **mortgage/ovefjwjrat/ovesuwurat** at the start of each month.

After the shelter fell down on his brand new truck, Bob parked it elsewhere.

After the shelter fell down on his old bed, Bob moved it elsewhere.

He moved it to the stout **garage/jwewjr/uwewur** a few blocks from his home.

On ranches in Texas, the cook rings the old chime for dinner time.

The cook hates the noisy beginning of dinner time.

He needs to bang on the rusty **triangle/fehwujdr/senwxusr** loudly so distant cowboys can hear it.

Paul regretted leaving his hat at home as he walked the desert at midday.

His water bottle was empty after the first four hours of trekking.

He protected his eyes against the harsh **sunlight/anudhjtf/anxsnucs** as he walked across the dunes.

John's Olympic shirt was torn during the gold medal soccer match.

John's outfit was torn after he fell out on the field.

He put on his finer **jersey/qrearp/wreara** for the awards ceremony that night.

John is tired of trying to get his old laptop to work.

John is tired of trying to get his old motorcycle to work.

He really wants to get a newer **computer/mvoynfre/mvoznsre** to help him solve the problem.

Roger dropped his watch far down into the muddy waters of the swamp while fishing.

Roger dropped his fishing hook in the bottom of his boat while fishing.

He searched for it in the murky **depths/lryfta/rrzsca** without being able to find it.

The knight enjoyed training squires to use weapons and armor.

The father loved playing soldier with his son on the weekends.

He showed the boy how to use the tiny sword and small **shield/athrdl/acnrsr** to defend himself.

Chef Marcus was famous for baking things like doughnuts, Danishes, and cakes.

Chef Marcus was famous for his tasty and delicious dishes.

He was disappointed at the student's lumpy **pastry/ywafep/zwasea** and frowned with frustration at the effort.

Mark waited in line a long time to get his seats for the concert.

Mark waited in line a long time before he reached the register.

He was glad that he was able to get great **tickets/fhmbrfa/snmzrsa** and sit near the stage.

Walter tended to make a huge income by stealing things then selling them to people.

Walter often got into trouble by repeatedly breaking the law.

He was quite excited when he made a large **profit/yevthf/zevons** off his illegal activities.

Little Bobby tended to punch, hit, and kick the other children at school.

Little Bobby tended to have quite a lot of interest in school.

He would often get into rough **fights/thjtfa/onucsa** with the other kids on the playground.

Harry likes to check out books far from the bad neighborhood.

Harry likes to read while he rides the bus.

He would rather go to the safer **library/dhkewep/snvewea** across town than the one near his home.

Charles always brought a family photo with him to competitions for good fortune.

Charles always brought a family heirloom with him to competitions for good fortune.

He would refuse to compete without the lucky **picture/yhmfner/znmsner** of his grandparents.

The army guard brought coffee when inspecting the perimeter on his route at night.

The doctor bought coffee when he got sleepy late at night.

He would take a break on his later **patrols/ywfevda/zwsevsa** to drink it and have a doughnut.

Cars rarely cross the river here anymore because of the danger involved.

Pedestrians rarely walk here anymore because of the danger involved.

It is unlikely the frail **bridge/kehljr/venrur** can hold the weight of the travelers.

The family waited hungrily for the bird to be done on Thanksgiving.

The family enjoyed getting together for breakfast now and then.

It was family tradition to enjoy some juicy **turkey/fnebrp/snezra** before sitting down to visit.

Optometrists like to shine bright lights in the eyes.

Doctors like to examine the body thoroughly at checkups.

Janet winced when the doctor looked at her black **pupils/ynyhda/znznsa** with his special instrument.

Many families have favorite meals that are handed down over generations.

Many families have favorite items that are handed down over generations.

Matilda cherished the older **recipes/ermhyra/ermnzra** that her family had kept throughout the years.

Having vegetable stew without some tasty spuds is a really bad idea.

Cooking a meal without some savory seasoning is a bad idea.

Patti dropped in some salty **potatoes/yvfwfvra/zvswsvra** to make sure the stew tasted really great.

The experts deciding the outcome of American Idol refused to see the truth.

The people involved in the decision refused to see the truth.

Paula sincerely wished those blind **judges/qnljra/wnrura** would have picked the correct contestant.

Henry gave Yvette a sparkling, clear gem for her engagement ring.

Henry gave Yvette a unique rock he found while travelling.

She delighted in showing off the large **diamond/lhwovul/rnwovxr** she was given to all her friends.

A wind blew in the open window and spread Shirley's manuscript across the floor.

A dog ran in the open door and spread Shirley's toys across the floor.

She had to quickly collect the loose **papers/ywyrea/zwzrea** before she lost any of them.

Cathy expects something large and interesting in the mail today.

Cathy expects something large and interesting for the garden today.

She is delighted when she sees the beige **package/ywmbwjr/zwmzwur** in the mailman's hands.

After she cleaned the kitchen sink, Gloria didn't put anything away.

After she cleaned the floor, Gloria didn't put anything away.

She later found the moist **sponge/ayvujr/azvxur** leaving a small puddle on the counter.

Lucinda always used the same writing utensil every time she took a test.

Lucinda always brought the same item with her when she needed confidence.

She made sure to bring her lucky pencil/yrumhd/zrxmns to every situation where she might need it.

Patty is awakened frequently, around 11:00 pm, by the crying baby.

A family of birds awakens Patty very early each morning.

She often misses the quiet **nights/uhjtfa/xnucsa** she had before the baby arrived.

Regina loves watching the youngsters play at the tree-lined park.

Regina likes walking by the bird-filled park on her way to work.

She really likes hearing the noises that the happy **children/mthdleru/mcnsrerx** make as they play together.

The queen surveyed her domain from atop a cold, high mountain.

The mother superior's abbey sat atop a cold, high mountain.

She ruled her chill **kingdom/bhujlvo/znxurvo** fairly and firmly from her seat on the peak.

Deborah loves selling homes with large counters, new microwaves, and fancy stoves.

Deborah loves selling homes with large yards, new fencing, and fancy landscaping.

She shows her clients the ample **kitchens/bhfmtrua/znsmcrxa** so that they will buy the home.

It was hard for Becca to hear while she was waiting for her flight.

It was hard for Becca to hear the person on the phone.

She tried to make her call at the noisy **airport/wheyvef/wnezves** before she departed on vacation.

Unhappy customers often demand to see a rude employee's supervisor.

You have to be "on your toes" when working at a locally owned grocery store.

Shoppers often ask to speak with civil **managers/owuwjrea/owxwurea** when unruly clerks refuse to help.

Susie's mom always told her not to talk to grownups she didn't know.

Susie's mom always told her not to play with dogs she didn't know.

Susie made sure to avoid the weird **stranger/afeaujre/asewxure** who tried to talk to her at the mall.

The tribal children were very poor and they only had dirty rags to wear.

The prisoners were unhappy and only had dirty socks to wear.

The activists were upset at the stark **poverty/yvsrefp/zvsresa** in which they had to live.

The circus had a large man that could lift very heavy things.

The circus had performers that were very entertaining.

The audience was particularly impressed by the sheer **strength/aferujft/aserxusc** of the strongman.

The drummer's cadence kept a steady beat for the dancers to follow.

The director helped the dancers through the performance.

The background that the solid **rhythm/etpfto/ecasco** provided the dancers made the performance great.

There were several large men adding a new coat to the faded wall.

There were several men who were repairing the wall together.

The children stopped by to watch the husky **painters/ywhufrea/zwnxsrea** finish the newly repaired wall.

Every year, the Catholic Church sent some clergy to hand out sugary holiday treats.

Every year, the town council sent community members to hand out sugary holiday treats.

The children were delighted by the happy **priests/yehrafa/zenrasa** who threw candy by the handful.

They only read two pages of the dialog she had written in her play before stopping.

She left the office with her head hung low in shame.

The director said it was the worst **script/amehyf/amenzs** he had read in twenty years.

Older inmates like to pick on newer convicts when they are first sent to jail.

The youth was on his own from the first moment he arrived.

The guards were unable to protect the young **prisoner/yehavure/zenavxre** from the hazing he endured.

Children often get a fever before they develop the swine flu.

Youngsters often suffer from a variety of childhood illnesses.

The high temperature was an early **symptom/apoyfvo/aaozsvo** of a more serious illness.

The mother preferred giving her children wooden toys.

The mother preferred giving her children cars and trucks to play with.

The little child refused to let go of the shiny **plastic/ydwafhm/zswasnm** plaything he found.

Because the people were forced to flee their country, they were frightened.

Because the people had to leave before they were ready, they were uncertain.

The men at the border watched the timid **refugees/ertnjrra/eronurra** with a mix of admiration and pity.

On quiet nights, one can hear melodic lullabies from the valley below.

People in the valley loved to celebrate loudly.

The mountain folk enjoyed the faint **singing/ahujhuj/anxunxu** they heard before they went to sleep.

It was hard to see the old portrait through the layers of dirt.

It was hard to see the old sculpture through the layers of dirt.

The museum curator had the dusty **painting/ywhufhuj/zwnxsnxu** cleaned as soon as it arrived.

The other workers had left the scaffolding up for the painters to use as a floor.

The construction workers had left the boards up for the painters.

The painters appreciated the handy **platform/ydwftveo/zswsoveo** that allowed them to reach the roof.

Anna liked to go the chapel to pray late at night when few people were there.

Anna liked to sing her heart out late at night when few people were there.

The peaceful quiet of the still **church/mtnemt/mcnemc** helped her feel her petitions were heard.

The aristocracy of other countries often visit our nation's capital.

The general public of other countries often visit our nation's capital.

The press loves to take pictures of such minor royalty/evpwdfp/evawssa while they're visiting.

They realized their expedition would take many miles over many days to return home.

They realized their repairs would take many hours over many days to complete.

The sailors prayed for a swift **journey/qvneurp/wvnexra** as they set out for home at last.

The principal hated to spend money on graduation.

The principal hated to spend money on things for the school.

The students anticipated a cheap **ceremony/mrerovup/mrerovxa** but were pleasantly surprised.

The student spent very little time working on the science fair assignment he handed in.

The student spent very little time on the writing assignment he turned in.

The teacher grimaced as he graded the hasty **project/yevqrmf/zevwrms** the lazy student had submitted.

The country received sanctions due to the open buying and selling of people.

The country received sanctions due to the open buying and selling of illicit goods.

The world media was disgusted when reporting the overt **slavery/adwsrep/aswsrea** taking place there each day.

Young burglars often need advice from crooks who are older and more experienced.

Young pirates often need advice from those who are older and more experienced.

The youngsters consulted the elder **thieves/fthrsra/scnrsra** because of trouble finding the good loot.

Those at the party enjoyed the textured Ruffles potato chips.

Those at the party enjoyed the texture of the lovely salad.

They found that the crisp **ridges/ehljra/enrura** only enhanced the flavor of the dish.

The roommates always sent only one check to the power company.

The roommates always went to the post office once a week.

They mailed their joint **payment/ywporuf/zwaorxs** for the utility bill on the first of the month.

The manual instructed the operator to type in particular numbers in order.

The manual instructed operators to type in their password first.

They must type the exact **sequence/arjnrumr/arxnrxmr** necessary or the machinery in the plant won't work.

The students were forced to listen to many boring presentations

The students were forced to listen to many boring commercials.

They rejoiced as the final **speaker/ayrwbre/azrwzre** took the stage.

Because they were infertile, the child-hungry couple contacted the agency.

The poor family was having trouble feeding all of their children.

They were excited after the quick **adoption/wlvyfhvu/wrvzsnvx** of their new baby girl.

There were thin, ribbon-like slices of steak cooking over the coals for the BBQ.

There were many different types of meat cooking over the coals for the BBQ.

Tom the dog looked longingly at the fresh **strips/afehya/asenza** of meat sizzling on the grill.

My textbook is full of illustrations depicting structures inside the body.

My textbook is full of historical knowledge people have gathered over the years.

We spend hours studying human **anatomy/wuwfvop/wxwsvoa** in our first semester of medical school.

The man washed and combed the fine wool on his livestock at the fair.

The man washed and combed the fine fur on his livestock at the fair.

Winning the grand prize made the proud **shepherd/atrytrel/acrzcrer** a very rich and famous man indeed.

**11. REFERENCES**

Ashby, J., Rayner, K., & Clifton, C. (2004). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability*. The Quarterly Journal of Experimental Psychology Section A, 58: 6*, 1065-1086.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Balota, D. A. Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364-390.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal fo Mathematical and Statistical Psychology, 41,* 257-278.

Benschop, R. (1998). What Is a Tachistoscope? Historical Explorations of an Instrument. *Science in Context, 11,* 23-50.

Besner, D. (1989). On the role of outline shape and word-specific visual pattern in the identification of function words: None. *The Quarterly Journal of Experimental Psychology Section A, 41 (1)*, 91-105.

Bouma, H. (1971). Visual recognition of isolated lower-case letters. *Vision Research, 11,* 459-474.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), Cognitive Processes in Eye Guidance. Oxford, England: Oxford University Press.

Buswell, G. (1935). How people look at pictures. Chicago: University of Chicago Press.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59 (3)*, 209-217.

Chance, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59 (3),* 209-217.

Cooper, H., & Hedges, L. V. (2009). Research Synthesis as a Scientific Process. In H. Cooper, L. V. Hedges, & J. C. Valentine (eds.). The Handbook of Research Synthesis and Meta-Analysis (2nd Ed.). New York, NY: Russel Sage Foundation.

Courrieu, P., Farioli, F., & Grainger, J. (2004). Inverse discrimination time as a perceptual distance for alphabetic characters. *Visual Cognition, 11 (7),* 901-919.

Dambacher, M., & Kliegl, R. (2007). Synchronizing timelines: Relations between fixation durations and N400 amplitudes during sentence reading. Brain Research, 1155, 147-162.

Dodge, R. & Cline, T. S. (1901). The angle velocity of eye-movements. *Psychological Review,* 8, 145-157.

Dunn-Rankin, P. (1968). The similarity of lowercase letters of the English alphabet. *Journal of Verbal Learning and Verbal Behavior, 7,* 990-995.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112 (4),* 777-813.

Faul F., Erdfelder E., Lang A.-G., & Buchner A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Faust, M., Babkoff, H., & Avidor-Reiss, I. (2000). Sentence and word outline shape as co-primes for target words presented to the two visual hemifields. *Brain and Language, 73,* 50-61.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31 (2)*, 119-122.

Field, A. (2009). Discovering Statistics Using SPSS. London, UK: Sage.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31 (5)*, 862-877.

Grainger, J. (2007). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23 (1*), 1-35.

Grainger, J., & Jacobs, A. M. (1999). Temporal integration of information in orthographic priming. *Visual Cognition , 6 (3/4),* 461-492.

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Science, 12 (10),* 381-387.

Huey, E. B. (1908). The Psychology and Pedagogy of Reading. MacMillan Press. 469 pages.

Hyona, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition, 32 (4),* 523-532.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 29-54). Oxford, England: Elsevier.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431-439.

Inhoff, A., Eiter, B., Radach, R., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *The Quarterly Journal of Experimental Psychology Section A, 56 (5)*, 803 – 827.

Jacobs, A. M., Nazir, T. A., Heller, O. (1989). Perception of lowercase letters in peripheral vision: A discrimination matrix based on saccade latencies. *Perception & Psychophysics, 46 (1),* 95-102.

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Perforance, 34 (6),* 1560-1579.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16 (1/2)*, 262-284.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Lete, B., & Pynte, J. (2003). Word-shape and word-lexical-frequency effects in lexical-decision and naming tasks. *Visual Cognition, 10 (8),* 913-948.

Lipsey, M. W., & Wilson, D. B. (2001). Practical Meta-Analysis. Thousand Oaks, CA: Sage.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research, 46,* 4416-4424.

McDonald, S. A., & Shillcock, R. C. (2003a). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research, 43,* 1735-1751.

McDonald, S. A., & Shillcock, R. C. (2003b). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14 (6),* 648-652.

Miller, B., Juhasz, B. J., Rayner, K. (2006). The orthographic uniqueness point and eye movements during reading. *British Journal of Psychology, 97,* 191-216.

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Science, 7 (3),* 141-145.

Paap, K. R., Newsome, S. L., & Noel, R. W. (1984). Word shape’s in poor shape for the race to the lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 10 (3),* 413-428.

Pedhazur, E. J. (1997). Multiple Regression in Behavioral Research: Explanation and Prediction. USA: Thompson.

Pelli, D. G., Burns, C. W., Farell, B., Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research, 46,* 4646-4674.

Perea, M., & Rosa, E. (2002). Does “whole word shape” play a role in visual word recognition? *Perception and Psychophysics, 65 (5)*, p. 785.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 148-162.

Pollatsek, A., Reichle, E. D., Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1-56.

Radach, R. R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology , 16 (1/2),* 3-26.

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in words during reading. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 77–100). Oxford, England: Elsevier.

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology, 16(1),* 27- 51.

Radach, R., Reilly, R., & Inhoff, A. (2007). Models of oculomotor control in reading: Towards a theoretical foundation of current debates. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), Eye movements: A window on mind and brain. Oxford, England: Elsevier.

Radach, R., Inhoff, A. W., & Heller, D. (2004). [Orthographic regularity gradually modulates saccade amplitudes in reading .](http://www.fcrr.org/vip_lab/articles/radach,%20inhoff,%20heller%20orthographic%20regularity.pdf) *European Journal of Cognitive Psychology, 16*, 27–51.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85 (3),* 618-660.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8 (1),* 21-30.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65–81.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7,* 4-22.

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7,* 34-55.

Reilly, R., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J. Hyona, R. Radach, & H. Deubel (Eds.). The mind’s eye: Cognitive and applied aspects of eye movements. Elsevier.

Rudnicky, A. I., & Kolers, P. A. (1984). Size and case of type as stimuli in reading. *Journal of Experimental Psychology: Human Perception and Performance, 10 (2),* 231-249.

Schrank, F. A., Becker, K. A., & Decker, S. (2001). *Calculating ability/achievement discrepancies between the Wechsler Intelligence Scale for Children–Third Edition and the Woodcock-Johnson III Tests of Achievement (Woodcock-Johnson III Assessment Service Bulletin No. 4).* Itasca, IL: Riverside Publishing.

Smith, J. D. (1992). The Remarkable Ibn al-Haytham. *The Mathematical Gazette, 76 (475)*, p. 189-198.

Starr, M. S., & Inhoff, A. W. (2004). Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology, 16 (1*), 203-225.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Science, 5 (4),* 156-163.

Vainio, S., Hyona, J., & Pajunen, A. (2009). Lexical predictability exerts robust effects on fixation duration, but not on initial landing position during reading. *Experimental Psychology, 56 (1),* 66-74.

Vitu, F., O’Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics, 47 (6),* 583-600.

Wade, N. J. (1998). A Natural History of Vision. MIT Press.

Walker, P. (1987). Word shape as a cue to the identity of a word: An analysis of the Kuera and Francis (1967) word list. *The Quarterly Journal of Experimental Psychology Section A, 39 (4*), 675-700.

Woodcock, R. W. (1998). Woodcock Reading Mastery Tests-Revised (WRMT-R/NU). Circle Pines, Minnesota: American Guidance Service.

Young, L. R., & Sheena, D. (1975). Eye-movement measurement techniques. *American Psychologist, 30*, 315-330.

**12. BIOGRAPHICAL SKETCH**

In the summer of 1998, Jessica C. Hill graduated with a Bachelor’s of Science Degree in Psychology with Departmental Honors after completing an honors thesis with Dr. L. Mark Carrier at Florida State University. Under the direction of Dr. Art Kramer, Jessi completed a Master’s Degree in Visual Cognition and Human Performance at the University of Illinois at Urbana-Champaign. After a brief leave of absence from school, Jessi returned to her studies in the fall of 2006 at Florida State University under the direction of Dr. Ralph Radach in the Doctoral program in Developmental Psychology.

Jessi’s research interests include: attention and human performance, specifically, developmental attentional differences; research and development of more objective diagnostic criteria for individuals with Attention Deficit/Hyperactivity Disorder (ADHD); fundamental measurement and methodological issues in eye tracking research; and the differential contributions of various types of masks (e.g., random letters, pseudowords) to the preview benefit seen in reading research.

1. Although it is common to refer to this process as a “decision,” this is not a conscious process. A better conceptualization is that once a threshold of activity is reached within the oculomotor system, a certain process is triggered (e.g., Reilly & Radach, 2006). [↑](#footnote-ref-1)