THE INFLUENCE OF UNIT SIZE ON ADULT READERS WITH A MODIFIED STROOP AND NONWORD READING TASK

by

JENNIFER MARIE SIECZKO

(Under the Direction of Paula J. Schwanenflugel)

ABSTRACT

The purpose of this experiment was to examine which levels of spelling-sound relations are most influential with proficient readers. Some researchers believe that small units or grapheme-to-phoneme correspondences (GPC) are most influential, others believe that large units or rime units are most influential, while others argue that both are important and that the use of each is task dependent (Brown & Deavers, 1999). The current study extends on research done by Strauss (2001), which aimed to determine which size unit was most influential in early word reading with a modified Stroop task. The modified picture-word Stroop task was also utilized for this study, but with adult readers and was designed to determine whether GPC or rime units were most influential in adult automatic reading. In addition, a nonword reading task adapted from Brown & Deavers (1999) was used to determine which size unit was most influential in strategic reading.

INDEX WORDS: Stroop, Nonword, Automatic Reading, Strategic Reading, GPC, Rime Units, Spelling-sound relations

THE INFLUENCE OF UNIT SIZE ON ADULT READERS WITH A MODIFIED STROOP AND NONWORD READING TASK

by

JENNIFER MARIE SIECZKO

B.S. Ed., The University of Georgia, 1995

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2004

THE INFLUENCE OF UNIT SIZE ON ADULT READERS WITH A MODIFIED STROOP AND NONWORD READING TASK

by

JENNIFER MARIE SIECZKO

Major Professor: Paula J. Schwanenflugel

Committee: Nancy F. Knapp
Linda Labbo

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

ACKNOWLEDGEMENTS

My greatest appreciation goes to my major professor, Dr. Paula J. Schwanenflugel for her support and guidance. I would also like to thank Beth Meisinger for her kind help.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Few reading researchers dispute the relevance of spelling-sound relations in learning to read. However, opinions are divided on exactly how important they are and the extent to which they are utilized by children and adults in word decoding. Spelling-sound relations are the relationships that exist between the phonemes heard in pronounced speech and various spelling (or orthographic) patterns. These spelling patterns may be contrasted according to their size (large or small units of the word) that readers may employ when decoding a word. Some believe that small units are key and that grapheme-phoneme correspondences (GPC) are most helpful in reading (Coltheart & Leahy, 1992), while others argue that larger units such as rime units play an important role (Treiman, 1983). In the middle are those who believe that both are important and that what use readers make of these different levels of spelling-sound relations is dependent on the task at hand (Brown & Deavers, 1999). The unit size issue is much debated and there is at least some evidence that supports all sides. Different types of experiments and tasks have been used to determine how readers best decode words or read them automatically and the implications of these differing tasks and findings will be explored. Reading is a wide-ranging and complex area to explore; so while this study does focus on unit size and the influence of it on automatic and strategic reading, it does not dispute that there are many other aspects important to reading instruction.

Terminology

Small-units are spelling-to-sound relations at the most basic level of decoding and are sometimes called *grapheme-phoneme correspondences* or GPC. It is said that readers utilize individual graphemes, single letters or clusters, as they relate to an individual phoneme (Venezky, 1970). A grapheme is a single letter or letter cluster, while a phoneme is the sound that represents that letter or letter cluster (Coltheart & Leahy, 1992). For example, in the word *sheet*, the graphemes *sh*, *ee*, and *t* each represent separate phonemes. There is a direct one-to-one correspondence from each letter or letter cluster to each sound (Berndt, Reggia, & Mitchum, 1987). Supporters of the view that readers engage small-units believe that GPC strategies aid readers to more successful attempts when they come across unfamiliar words (Brown & Deavers, 1999).

Large units involve a higher level of spelling-to-sound relations, taking advantage of larger pieces of words in the decoding process. The term *rime unit* will be used in this paper synonymously with large unit. It is believed that readers utilize recurrent rime of spoken syllables that include a vowel followed by a consonant (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). For single syllable words, this view organizes word parts into *onset* and *rime* units. For a consonant-vowel-consonant word, the onset is the initial consonant or cluster of the word, the coda is the final consonant or cluster of the word and the rime is the *peak* and *coda* of the word (Treiman, 1983). For example, with the word *truck*, *tr-* is the onset, *-u* is the peak, *-ck* is the coda and together the peak and coda *-uck* is the rime. Advocates of the large-units model believe that readers can produce more consistent and accurate pronunciations when they recognize a familiar rime in an unfamiliar word and utilize it for pronunciation (Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995).

The terms used to describe units of various sizes vary from study to study while carrying similar meanings. Some researchers use the word *analogy* to describe the pronunciations given to nonwords or pseudowords when a reader is using the large units found in known words related phonologically to the nonword (Glushko, 1979). In some studies, ambiguous nonwords are used to assess what size unit rules are being used by readers to tackle unfamiliar words. Ambiguous nonwords are nonwords that can be pronounced several different ways. Two possible pronunciations for ambiguous nonwords are those that use a GPC rule or those that use an analogy strategy. For example, when reading the nonword "dalk", the reader might pronounce it to rime with the word "talk" or "chalk" if they use an analogy response but with an /a/ sound as in "apple" if a GPC strategy were being used (Brown & Deavers, 1999).

Researchers may also describe words or nonwords as *regular* or *irregular,* and *consistent* or *inconsistent.* The regularity of a word is determined by whether or not the pronunciation of the word follows GPC rules (Coltheart & Leahy, 1992). A nonword is considered regular when the GPC pronunciation given to it would be the same as the rime pronunciation, and would be considered irregular when the nonword would have a different pronunciation dependent on whether GPC rules or rime rules are being used (Brown & Deavers, 1999). Consistency is determined by the amount of words with the same rime ending that are vocalized the same way. Words that are inconsistent include rimes that have two or more pronunciations, one regular pronunciation and one or more irregular pronunciations (Coltheart & Leahy, 1992). Some researchers have adopted the term "exception" to describe a word in the same capacity as "irregular" (Glushko, 1979).

Unit Size and Reading Skill

There are important educational implications of how influential units of differing size are for strategic and automatic skilled reading. If skilled decoding and automaticity is achieved by taking advantage of GPCs, then beginning readers may benefit more from a small unit phonics-based type of instruction. Conversely, if skilled decoding and automatic reading is achieved while utilizing larger units, then it may best serve educators to adopt a reading program that includes onset and rime instruction. If the use of these units is task-dependent, then it would seem to imply that both of the models be incorporated into reading instruction.

It is also imperative to consider the different skills that proficient readers must show knowledge of. Two skills focused on for the purposes of this study are: strategic reading and automatic reading. To become an expert reader, it is important that the individual be successful in decoding unfamiliar words in strategic reading as well as successful in automatically reading words. It is ideal that young readers attain automaticity in decoding so that their focus can be directed toward comprehension of text (Laberge & Samuels, 1974). Automatic reading means that decoding is effortless, quick, accurate, and occurs without intention (Logan, 1997). Since reading is a complex skill that takes years to develop, it is necessary to track the development of reading and observe the ways that skilled automatic readers decode. This can lead to the formation and implementation of better resources to aid teachers and parents in helping their students/children to become automatic readers more easily and quickly. As even the most fluent reader is likely to come across a new word at some time or another, it is also important to observe how advanced readers read strategically and successfully to tackle new words. It is possible that small-units could be most influential on strategic reading, while larger units are more influential on automatic reading, or vice-versa, which would support a more blended type of instruction. The second possibility is that both are influenced by the same size unit, requiring a more focused instruction.

The evidence regarding the size of units used by skilled readers or in the development of skilled reading comes from several different kinds of experiments. Many researchers have used nonword and word tasks to measure the influence of size units on the pronunciation of unfamiliar words (Coltheart & Leahy, 1992; Brown & Deavers, 1999; Treiman, 1983; Glushko, 1979); others have used priming tasks to note if context makes a difference in pronunciations (Goswami, 1993; Brown & Deavers, 1999); others have used Stroop-type tasks to learn more about the automaticity of size units in reading (Strauss, 2001).

Nonword tasks are used to measure strategies readers use to decode unfamiliar words. Glushko (1979) had adults read nonwords aloud in a mixed list consisting of exception, regular, regular consistent and regular inconsistent words and nonwords to determine what kinds of pronunciation rules were used. Glushko recorded only GPC responses as correct. Regular pseudowords were named more quickly and accurately than exception pseudowords. This suggests that nonwords are not strictly pronounced with basic GPC rules and that analogy with existing words (rime) were used to produce the “incorrect” pronunciations, lending support to rime units being influential in adult strategic reading.

Similarly, Coltheart and Leahy (1992) conducted several nonword reading experiments designed to examine pronunciation of and speed with which children and adults read nonwords with different consistencies and regularities. Brown and Deavers (1999) also presented children and adults with lists of ambiguous and nonambiguous nonwords where the rime or GPC pronunciation given would depend on the strategy used. Both studies support that rime unit responses increase as reading skill increases, but that all skill levels use more GPC pronunciations overall. However, “overall” includes both regular and irregular nonwords. As analogy pronunciations is not an option for regular nonwords, these conclusions seemed skewed. When the results are separated by regular and irregular words, GPC is only the majority with regular words for all ages. This is not necessarily true for irregular nonwords. There are more rime pronunciations than GPC pronunciations for the irregular words when read by advanced readers.

Treiman (1983) used nonwords to understand the size of the units involved in adult speech processes. The adult participants in the study were taught word and nonword games, requiring participants to divide stimuli words and nonwords according to the rules they were taught. The rules taught required participants to separate/combine words or nonwords according to the coda, rime and onset units of words. The participants’ preferences toward and ease of learning certain rules over others were of interest. The vast majority of the participants found it easier to and acted upon dividing the words between the onset and rime as opposed to dividing it more phonetically, favoring a large units strategy.

Timed real word tasks have also been used to test GPC/rime unit issue. Coltheart and Leahy (1992) found that regular real words were named more quickly and accurately overall by both children and adults (see also Glushko, 1979) than words with irregular spellings; however, perhaps, in a more sophisticated treatment of this issue, Treiman, Mullennix, Bijeljac-Babic, and Richmond-Welty (1995) conducted an analysis of the consistency and frequency of consonant-vowel-consonant words and parts of those words (beginning consonant, vowel, ending consonant, onset, and rime). Through this word analysis, the researchers discovered that the pronunciation of vowels as single vowel small units is less consistent than vowels presented in the context of rime units. Because of this, they reasoned that rime units were more likely to be utilized during decoding than GPC units. In a follow-up experiment, as the consistency of the initial consonant and of the rime increased, the reaction times of the adult participants decreased, which was viewed as evidence for the primacy of rime units in word reading.

Priming tasks (or clue word techniques) have also been seen by some experts as useful to determine the size unit strategy used by a reader, specifically to note if pronunciations are task-dependent (Goswami, 1993). The clue word presented as a prime is chosen because it shares some common characteristic with the stimulus word. Readers are told that they can use the clue word to help them with the stimulus word, and researchers look to see whether the common characteristic is carried over to the pronunciation of the stimulus word. Transfer would be shown if the participant pronounced the test word with the same rime, vowel(s), or onset as the clue word presented. For example, if rime units theories are supported, the word *bug* would show transfer to the word *rug* that shares the same rime unit, but not to the word *cup* that only shares a vowel. Goswami (1993) looked at size unit across grade level and overall, it was observed that early readers only carried over pronunciations according to rimes in words (e.g. bug-rug); but as reading developed, articulations for vowel graphemes (e.g. beak-heap) and onset-vowel (e.g. beak-bean) units were also transferred from clue word to test word. The author attributed this to an increase in complexity of reading as the participants become more proficient. The clue word technique was also adopted by Brown and Deavers (1999) in order to look at task context on the use of analogy responses to lists of nonwords. Adults' analogy responses were recorded and compared in different contexts: a regular nonword list, an irregular nonword list, and a list with clue words. The regular nonword list had more GPC pronunciations; the irregular nonword list and the prime/clue word list both had more analogy pronunciations. The context in which the words were presented had an effect on the size unit used. Considering that the regular nonwords could only be pronounced correctly using a GPC strategy, results of the pronunciations for that list are not surprising. However, it is interesting that for irregular words, that could be pronounced correctly with either GPC or rime strategies, that a rime unit strategy was chosen a majority of the time.

Clue-word task experiments tend to produce more rime-unit (analogy) pronunciations than those of the nonword task experiments (Brown & Deavers, 1999). This could be attributed to the priming caused by presenting the clue word before the target word; that is, if a clue word that requires a rime unit pronunciation (as opposed to a GPC pronunciation) is presented before an irregular nonword, the reader is likely to read the nonword with the rime pronunciation because the rime "seed" has been planted.

The Stroop Task

The experiments presented thus far have shown that readers appear to have both rime unit and GPC unit strategies available to them to use in context-specific manner. However, it is unclear how automatic these strategies are. Automaticity of any given skills is usually signified by quick, accurate, and obligatory responding. Even though word naming is rather quick to carry out in skilled adult readers, findings using the task seem contradictory and effects supporting the utility of rime units are quite small. Consequently, it may be necessary to develop tasks which by-pass the strategic element of human cognition and look at the more automatic aspects of word reading. One task that may do this is the Stroop task.

In a landmark experiment, J.R. Stroop (1935) investigated the effect of contradictory ink colors on reading color words aloud to measure interference in his color-word task. Interference was defined as the difference between the reaction time of the participant to the stimuli (contradictory ink colors) and the control condition (color patches). There was a highly significant difference between the two conditions, with the stimulus condition having the longest reaction times. The interference caused by attempting to name color ink of a word with a conflicting color term compared against a control of some type has been termed a ‘Stroop Effect”.

Since the original study, many studies have been conducted utilizing the Stroop task (MacLeod, 1991) and it has been used extensively to study the development of automaticity in word reading skill. Comalli, Wapner, and Werner (1962) investigated age differences and concluded that the largest interference occurred with 7 year olds then decreased with age until 19, remaining constant until 45 and increasing again at 65-80. The interference was said to have come from the word being automatically read by the participant, therefore causing them to be slower in naming the color. In studies that followed, more interference is equated with more automatic processing, at least early in the development of reading skill. Schadler and Thissen (1981) and Stanovich, Cunningham, and West (1981) have modified the task even further by contrasting naming color words printed in congruent print color, the color words in incongruent print color, naming the print color of unpronounceable nonwords created with the same letters as the color words in each color, and naming strings of Xs in each color (Schadler & Thissen, 1981). The Xs and nonwords caused more interference for children just beginning to learn to read compared to the more skilled readers. The incongruent stimuli interference was highest at 4th grade decreasing as skill level increased. The congruent stimuli showed a linear decrease from beginning to more skilled readers, the authors attributed this to the general improvement in response times due to age and the aiding effect of the congruent word to naming the color.

Stanovich, Cunningham, and West (1981) carried out a longitudinal study in first grade evaluating the development of the interference from letters, high frequency words, and low frequency words on color naming (from less-skilled to skilled throughout the school year). Letter interference was higher than the interference of high-frequency words and there was non-significant increase in interference of high-frequency words over low-frequency words throughout the year; more importantly, there was an increase in interference across all stimuli from the beginning to the middle of the year and skilled readers showed more interference and earlier interference than the less-skilled readers. This study shows an important shift in the type of stimulus conditions that are being used in Stroop tasks to look at different types of words.

Research has also been conducted using a picture-word Stroop task to look more closely at automatic word processing (Ehri, 1976; Guttentag & Haith, 1978; Ehri & Wilce, 1979). The picture-word Stroop task requires the participant to name a picture that has a conflicting label printed on it. Ehri (1976) studied beginning and adult readers in a picture-word task with congruent and discrepant labels. Although the beginning readers showed greater interference compared to the adult readers, early and proficient readers showed larger interference for conditions with conflicting labels compared to late and less proficient readers.

Word-specific training can also influence the size of the Stroop Effect. Ehri and Wilce (1979) conducted a study with first and second grade participants. These participants were divided in to two groups based on the pretest given: old-word/speed learners and new-word/accuracy learners. The old-word/speed learners were participants who could read most of the stimulus words correctly and the new-word/accuracy learners could not read at least 16 of the 20 words correctly. The training and practice trials were designed to insure that the children learned the pronunciations and the meanings of the printed distractor words. The old-word participants showed a decline in interference after the word training, while the new-word participant showed an increase in interference after the training. The authors note that training participants who could already read the words to read them more quickly apparently reduces the interference cause in a picture-naming task; while training participants new to the words to recognize more of these words will increase the interference in the picture-naming task. The new-word participants' results support interference as automaticity.

Using a picture-word Stroop task, Guttentag and Haith (1978) studied the development of Stroop interference to nonwords compared to words with first graders at early and late points in the school year, "poor and good" third grader readers (according to the classroom teachers' evaluation), and adults. The study compared control picture, nonletter character strings, nonpronounceable nonwords, pronounceable nonwords, unrelated word (intracategory), and related word (extracategory; e.g. a picture of a bed with the word chair). There was interference due to word category for all but the early first grade readers and letter interference existed for most participants. A handful of the early first graders did not show letter interference. Whether the letterstrings were pronounceable or not only affected the adults and third grade "good" readers, suggesting that lower level units were only automatized by the older children and adults. The fact that these different types of nonwords were differentially interfering in picture naming suggests that it may be a task suitable for examining automaticity of subword reading units.

More recently, Strauss (2001) conducted an investigation that examined the influence of unit size on the development of Stroop interference as a way of determining the development of word reading unit automaticity. He reasoned that by selecting words carefully, one could determine the size unit most influential in word reading by examining the relative interference across different types of words. For example, if rime units are the automatized reading unit, words containing high probability rime units should interfere more than words containing low probability rime units. In contrast, if GPC units are the automatized reading units, words containing high probability GPC units should interfere more than words containing low probability GPC units.

Using first, second and third grade participants, Strauss displayed line drawings with conflicting labels in four different experimental conditions and two controls (a picture control and random letterstring control). Nonwords were chosen to follow high probability GPC rules. Some words were chosen such that they followed a low probability GPC rule but had high rime predictability word; other words were chosen such that they followed a high probability GPC rule and also high rime unit predictability. Other words were chosen for having high probability GPC units but low rime predictability units. Presumably, if rime units were automatized then both the both word conditions having high probability rime units should show the greatest interference. If GPC units were automatized, then both word conditions containing high probability GPC units should show the greatest interference. There was a main effect of grade level, such that interference was greater for first and second grade than for third grade, children of all ages behaved similarly across word conditions. Overall, the small-units hypothesis was supported in that words with high GPC predictability showed the greatest interference.

The Purpose of the Current Study

Following Strauss (2001), the current study focuses on fluent readers with the goal of discerning whether adults, like the children in his study, also display a tendency towards greater automaticity of small units over larger ones. The current study carries out an exact replication of the Strauss study, but using college adults. This adult replication is important because it is possible that adults may show a different pattern of word interference than that found by Strauss. For example, it may be that the automaticity of GPC units over larger rime units is something more typical of less mature readers. One possibility is that, as children gain more experience with word reading, the units engaged become larger. Because the pronunciation of rime units is more predictable than GPC units, these units become increasingly automatized such that they become the operative unit in word reading. Alternatively, it may be that GPC units are the operative automatized unit in word reading for both children and adults. In that case, one might expect that the pattern of interference in children and adults are identical.

The current study also contrasts automatic word reading, such as found in the Stroop task, against strategic word reading, such as found in nonword pronunciation. In order to determine which size unit is most influential in strategic reading, a nonword reading task will also be used with the adult participants. A slightly revised nonword list from Brown and Deavers (1999) will be used and the pronunciations given to "more irregular" words from the list will be analyzed. These words are considered ambiguous, due to the fact that they can be pronounced differently depending on whether the participants utilize a GPC rule strategy or rime unit strategy or some "other" rule.

As mentioned previously, there are some that believe that the size of the unit employed in reading is task dependent. The purpose of this study is to determine what size units proficient readers use when reading written words automatically and also when reading strategically out loud. It is important to trace the steps to becoming a proficient reader both to automatic reading and to successful strategic reading to help our young readers to become accomplished readers in the least painful or difficult way. So, the questions that have shown themselves are: Do proficient readers use small units or large units most when reading automatically? Do proficient readers use small units or large units most when reading strategically?

CHAPTER 2

METHOD

Participants

Thirty-four adults from the undergraduate introductory psychology research pool at the University of Georgia participated in this study. The students participated for course credit. Their ages ranged from 17 years to 22 years (M=19.59, SD=1.26). There were 17 male subjects (50%) and 17 female subjects (50%). All subjects were native English speakers. The WRAT3 Blue Reading test was administered to be certain that all participants were proficient readers. All participants read at or above a high school level (stand score 93 for ages 17-19 and 90 for ages 20-24) on the test to be included in the study. The participants all met this requirement and most exceeded it to score at a post-high school level (standard score M=112.41, SD=6.30).

Apparatus

I used the modified Stroop task developed by Gregory Paul Strauss (2001). The stimuli were displayed on a 15" computer screen of a Dell Inspiron (8000) laptop computer. The participants could adjust their viewing distance to comfort, but it was typically at about 2 feet. The verbal responses were measured with a voice-operated microphone (Audio-Technia Cardioid low Impedance model # ATR 20C) that was connected to a response box. This serial response box (Psychology Software Tools, Inc., Model # 200A) relayed the response to E-Prime Data-Aid software (Psychology Software tools, Inc., Version 5.0) on the computer that timed the responses.

Stimulus Materials

For the modified Stroop task, 66 line drawings obtained from Cycowicz, Friedman, Rothstein, and Snodgrass (1997) picture norms were selected. These pictures were digitized black line drawings of various shapes, animal and objects on a white background. The pictures were considered to be highly identifiable, with namability ratings of 90% and higher. The words that were superimposed on the pictures were written in lowercase, 18 point Arial Black regular font. The words or letter strings that were printed on the pictures were three to five letters long. The stimuli appeared on the screen for five seconds, or until a sound was detected by the microphone. There were 66 pictures used, 60 were experimental, and 6 were practice stimuli (one for each condition). The picture control condition stimuli were not varied from the original format designed by Cycowicz et al. The researchers intended to include 10 items per condition, though, it was necessary to rearrange the format after it was discovered that the word "rock" was included in the Low GPC + High Consistency Rime condition when it actually followed the rule of High GPC + High Consistency Rime condition. Therefore, it was added to that condition which resulted in 9 items in the Low GPC + High Consistency Rime condition and 11 items in the High GPC + High Consistency Rime condition. Appendix A presents all stimuli used for the Stroop task and Appendix B presents examples of picture stimuli used (Strauss, 2001).

The following details of how the word conditions for the experiment were established are taken directly from Strauss (2001). Normative word frequency was examined at first, second, and third grade levels to control for word frequency across conditions, using the Zeno, Millard, and Duvvuri (1995) child language corpus. The word conditions did not differ in the normative frequency of the interfering words, F (2, 27) = .434, p = .653 (High GPC + Low Rime Unit M = 91.33, SD= 193.16; Low GPC + High Rime Unit M =166.55, SD= 171.42; High GPC + High

Rime Unit M =141.10, SD= 159.06). GPC predictability was also varied across conditions according to the Berndt, Reggia, and Mitchum (1995) norms. High and low rime unit predictability conditions, used to control rime unit interference, did not differ in terms of GPC, F (2, 18) = 1.13, p = .275 (Low Rime Unit M = 2.18, SD= .44, High Rime Unit M = 2.36, SD= .22). Predictability in terms of rime units was taken directly from Treiman, Mullennix, Bijeljac-Babic, and Richmond-Welty (1995), who determined rime unit predictability by using a computerized version of the *Merriam-Webster Pocket Dictionary*, while controlling for consonant-vowel-consonant (CVC) pronunciation and the orthographic neighbors of each CVC word (Treiman et al., 1995).

For the nonword-naming task, one nonword list from Brown & Deavers (1999) was used with the participants to determine if they would use GPC or rime-units as a strategy to read unfamiliar words. The list was revised slightly to fit the difference in dialect between the original study and the current study. The original study was done in the United Kingdom, and the pronunciation of the nonwords *paft, jance, hanch,* and *nast* were considered irregular rimes because in most Southern British dialects the *a* in these words are pronounced the same as in *car* and *party* (Brown & Deavers, 1999). Because in the United States the pronunciation for these words is quite different, these four words were removed from the "more irregular" group of nonwords. The "more irregular" group of nonwords were nonwords that could be pronounced differently depending on whether participants used a GPC rule strategy or a rime unit strategy. In order to make the "more regular" group the same length, four words were randomly removed from that list as well for the purposes of this study. The "more regular" group of nonwords were those for which both the GPC and rime unit strategies would have produced the same pronunciation. The more regular group was used as fillers, as they provided no interesting contrast for pronunciations. These groups were combined to form one list in random order in which the participants were to read from and their responses were recorded as either using GPC, rime, or other strategies (see Appendix C). The list was presented to the participants in 12 point black font on a white piece of 8 X 11 piece of paper attached to poster board for sturdiness.

Procedure

The participants were tested in quiet laboratory room. Each participant was seated at a desk and the revised Brown and Deavers nonword list was presented to him or her. Participants were told, "Please read out loud from this list of nonwords. They're not real words, I just want you to tell me how they sound. You'll start at the top and go down the first column before starting on the next column. It is not a speed test, so take your time." The responses of the participants were marked by the experimenter as using GPC, Rime, or other on a separate piece of paper. After the nonword list task was completed, a computer screen was placed in from of the participant and he or she was asked to hold the voice-operated microphone in hand while directions were read from the computer screen by the experimenter. These direction said, "Welcome to the picture naming task, in this task, you will first see a little cross in the middle of the computer screen like this: (+). Then, you will see a picture with some writing over it. For instance, you might see a picture of a baby with 'desk' written over it. Tell us what the picture is and pretend the writing isn't even there. Just name the picture. Can you tell me what you are supposed to do?" The experimenter explained to the participant that they must speak loudly and clearly in to the microphone and that they must hold it close to their mouth constantly so the computer will record each spoken response. The subjects were asked if they had any questions. If there were any misconceptions, they were clarified before beginning the task. When the participant was ready, the practice stimuli were administered.

Before the practice stimuli were presented, a sentence stating, “Now we’re ready to do some practice. Get ready to name the picture” was presented and read aloud by the experimenter from the computer screen. The six practice stimuli were then shown in random order, which included one stimulus from each condition. After the subject responded to all practice stimuli, another sentence was presented and read from the computer screen stating, “Now we’re going to do the real thing. Just do your best and that’s good enough for me.” The 60 experimental trials were then presented to the participant.

For each trial, the initial stimulus was initiated when a button on the response box was pressed by the experimenter, which caused the fixation point (+) to appear for 500 ms, ensued soon after by the picture stimulus. Participants named each picture stimulus by speaking loudly and clearly into the microphone. The experimenter then scored the accuracy of the participant’s response by pressing the appropriate button on the response box. Answers that differed from a suitable name for the object were coded as incorrect. When the microphone did not pick up the participant’s response, when the microphone picked up a sound other than the participant’s response, or when the participant made a sound other than their verbal response, the experimenter coded this as a microphone error. All other responses were scored as correct.

Design

A within-subjects design was used with six levels of the Stroop factor (High GPC + High Predictability Rime words, High GPC + Low Predictability Rime words, Low GPC + High Predictability Rime words, Nonwords, Letter Strings, and Picture Controls). Picture naming time measured in milliseconds and error rate measured as percent incorrect both served as dependent variables.

The Brown and Deavers nonwords task used a within-subjects design having two levels of strategy (GPC, rime unit). The dependent variable was the percentage of irregular words that were read either with GPC pronunciation or rime unit pronunciation.

CHAPTER 3

RESULTS

To determine the influence of unit size in automatized reading in comparison to strategic reading, it was necessary to analyze the two different tests in the study separately. The nonword reading task was used to determine unit size most often used in strategic reading and the modified Stroop task was used to indicate the unit size most often used in automatic reading.

Nonword Reading Task

The responses that the participants gave to the irregular nonwords were scored as GPC, Rime, or other. A GPC strategy response would pronounce the given vowel grapheme with the most common pronunciation (according to Venezky, 1970) with all consonants accurately pronounced. If the rime strategy is used, the pronunciation should rhyme with the analogous word (e.g. pronouncing *dalk* to rhyme with *talk*) according to Brown & Deavers (1999). A response was scored as “other” if it conformed to neither of these pronunciation rules.

All 34 participants used Rime units the majority of the time when reading the 11 irregular nonwords. GPC pronunciations were only in a minority of pronunciations, while strategies that fell under “other” were used even less. The means and standard deviations for each condition can be found in Table 1. These results indicated that adults tended to use rime units as a strategy in reading unfamiliar words (see Table 1). A paired sample t-test was conducted to find the difference in means between rime and GPC pronunciations for the nonword reading task. The “other” category was dropped for this analysis to preserve the independence assumption of the t-test. There was sufficient evidence of a statistical difference between the two means, t (33) =19.26, p<.01, favoring a Rime pronunciation.

Table 1

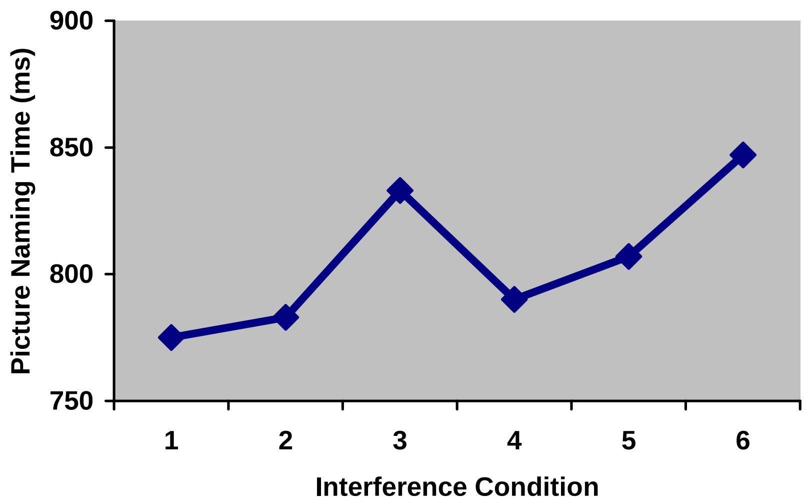
Number of Items for Which Participants Used Rime or Grapheme-Phoneme Correspondence Rule Strategy with the Nonword Reading Task

Strategy	*Mean*	*Standard Deviation*
Rime	8.38	.99
GPC	2.24	.96
Other	.38	.55

Modified Stroop Task

Prior to analyzing reaction times, it was necessary to eliminate problematic reaction times. Reaction time trials that were from incorrect responses were eliminated from the data (.34% of the data). Microphone errors and their corresponding reaction times were also eliminated from the data (3.73% of the data). Next, outliers that were over 2 standard deviations above the mean for each participant were removed from the data. 4.61% of the data were considered outliers. With these data points eliminated, mean reaction times and error rates were calculated for each participant in each condition. Further, mean reaction times and error rates were calculated for each item in the study. These error rates were uniformly low (Control M = 0, SD = 0; Nonpronounceable Letter strings M = 0, SD = 0, Nonword M = .01, SD = .029; Low Predictability GPC with Highly Consistent Rime Units M = .01, SD = .027; High Predictability

GPC with Highly Consistent Rime Units M = 0, SD = 0, High Predictability GPC with Inconsistent Rime Units M = .01, SD = .024) Figure 1 displays the mean reaction times for each condition in the study. Reaction time was considered the main dependent variable because error rates were so small.



Note: 1= Control, 2= Nonpronounceable Letter strings, 3=Nonword, 4= Low Predictability GPC with Highly Consistent Rime Units, 5= High Predictability GPC with Highly Consistent Rime Units, 6= High Predictability GPC with Inconsistent Rime Units

Figure 1. *Picture naming time as a function of Stroop interference condition*

To determine the effect of word units on picture naming time (ms), a repeated measures analysis of variance was conducted comparing the time taken to name pictures in each of the stimulus condition's six levels (control, nonpronounceable letter strings, nonword, high predictability GPC with highly consistent rime units, high predictability GPC with inconsistent rime units, low predictability GPC with highly consistent rime units) with each other. The main effect for the Stroop condition was significant using participants as the random variable, F (5, 165) = 14.12, $p < .01$, partial $eta^2 = .299$. Using items as the random variable (treating each item as a case or subject), however, the main effect of condition was not significant, F (5, 54) =1.47, p= .214, partial $eta^2 = .120$. However, this lack of significance could be attributed to lack of statistical power in this item analysis (.48) because there were only 9-11 items in each condition. Consequently, I view the analysis using subjects as the random variable as the more reliable indicator of the effect of condition. Looking at Figure 1 suggests that there was more interference related to the two high GPC conditions and the nonword condition than with the control, letterstring, and low GPC conditions.

One goal of the experiment was to compare the influence of GPC versus rime unit in automatic reading with adults. If large units (rime units) are operative in automatic reading, then only conditions using high predictability rime units (Low GPC + High Rime, High GPC + High Rime) should show interference in picture naming over the letterstring control alone. If GPC or small units were most influential in adult automatic reading, then the only conditions using high predictability GPC units (High GPC + Low Rime, High GPC + High Rime, Nonwords) should show interference in picture naming over letterstrings alone. (In each case, I considered the letterstring control to be a better control since it contains letter information but no reading unit information.) Alternatively, it may be that both unit types have been automatized. In that case, all word types should show interference. To test these hypotheses directly, partial ANOVAs comparing each condition against the letter string control were carried out. Comparisons were made against the letterstring control rather than the picture only control, in order to be able to attribute the interference to word units per se and not the mere presence of letters alone (see Table 2).

The results show that the most interference was caused by the high GPC with low rime condition, followed by the nonword condition. Overall the findings point to greater interference with stimuli composed of highly consistent GPC units. Both the High GPC + High Rime and High GPC + Low rime (as well as nonwords) showed significant interference compared to letterstring controls.

Table 2

Picture-Word Interference Follow-Up Analysis

Condition compared to control	F(1,33)	p-value	partial eta²
Nonword	32.094	.000	.493
Low GPC + High Rime	1.587	.217	.046
High GPC + High Rime	8.301	.007	.201
High GPC + Low Rime	46.921	.000	.587

CHAPTER 4

DISCUSSION

The two tasks in this study were designed to determine if adults utilize GPC or rime units when reading automatically, without intention, and when reading strategically out loud. It was the hoped that comparison of the results of this study to similar studies done with early readers could provide insight into whether proficient readers and early readers utilize related or opposing strategies. The purpose was to follow the steps from early to proficient readers to note any differences or similarities that may aid instruction for beginning readers. The study centered around two questions: (a) Do proficient readers use small units or large units most when reading automatically? and (b) Do proficient readers use small units or large units most when reading strategically?

Through the nonword reading task, the pronunciation that adult readers gave ambiguous nonwords would determine whether GPC or rime units were the operative units in adult strategic reading. Through the modified Stroop task, differential interference across conditions would determine whether GPC or rime units were encoded automatically.

If it were found that both tasks are influenced most greatly by GPC units, then this might imply indicate that instruction for reading should focus primarily on small unit phonics instruction in order to lead children to becoming more proficient readers. If it were found that rime units heavily influence both tasks, then an instruction focusing more on analogy or rime units might be appropriate. If the two tasks lead to contrasting results, then both types of instruction might be called for in different contexts.

The current study indicated that proficient readers use small units most when reading automatically and large units most when reading strategically. It was found that, with proficient readers, there was more interference with consistent GPC words in the modified picture-word Stroop task. It was also found that, when these same proficient readers were given a nonword-reading task, pronunciations indicated a definite preference toward rime units.

The results from the nonword reading task and the modified Stroop task show opposite patterns. The nonword reading task supports the importance of rime units in reading unfamiliar words, while the modified Stroop task supports that GPC rules are more likely to be automatized than rime units. Both of these results, however, do collectively support the idea that what size units are most influential is task dependent (Brown & Deavers, 1999). This stresses the important role of both small and large size units in successful reading.

Throughout, I have interpreted greater interference as indicating greater automaticity of particular word units. However, a new theory of Stroop effects needs to be considered in light of our findings and other studies that utilize the Stroop task. Tectonic theory by Melara and Algom (2003) is different from previous theories of Stroop interference. The theory deals with selective attention and structures. According to the theory, information that makes up a structural relation can grow from either recent experiences (externally) or from memories of previous experiences (internally). The purpose of each structure is to moderate continuous cognitive processing by determining how quickly the different stored memories are retrieved from one instance to another. They propose that there are three types of information from which specific structural relations emerge: salient, surprising, and correlated information. According to the theory, attention is prompted by the more surprising, correlated and salient features of the surroundings in defiance of a clear goal for the task or a usually successful strategy to focus on one certain feature.

Melara and Algom contrast their theory to other Stroop theories, including the theory pertaining to automaticity. According to the tectonic theory, the effects noticeable in a Stroop task are due to attention and not automaticity.

I doubt that tectonic theory can explain the current findings. Indeed, if tectonic theory were correct, it would suggest that more attention is given to the high predictability GPC conditions than the consistent rime unit conditions. It may also indicate that the GPC condition words were somehow more salient, surprising, or correlated for the readers. This is to say that they may stand out more to the readers, take them by surprise, or relate more to the pictures. For the current study, I would not think that these factors would come in to play, as steps were taken to be sure that the words were somewhat equal in length, therefore in terms of size should not have made these conditions more salient. The words that were used in my study were also selected on the basis that they were very common and recognizable, so should not have been surprising to the participants. These words were randomly assigned to different pictures across a variety of categories, so should not have been consistently correlated. If indeed the high GPC condition words were more attention demanding (or *less* automatized) and produced greater interference, this could potentially turn the interpretation of my results on its ear, implying that it is rime units that are more automatized. If the rime units were less attention demanding, then one would expect to see both Low GPC + High Rime and High GPC + High Rime conditions to show reduced interference, but that is simply not the case. Thus, I feel the current pattern of findings might be best interpreted within existing automaticity theory.

I feel that a number of other Stroop-related findings related to early reading also provide conflicting implications for tectonic theory. Schadler and Thissen (1981) found that congruent stimuli showed less interference than incongruent stimuli, which is in direct opposition to the idea that correlated features would call for more attention and interference. On the other hand, the Ehri and Wilce study discussed could be looked at differently in view of the tectonic theory. The old-word/speed learners showed less interference through word training, while the new-word/accuracy learners showed more interference. If the tectonic theory holds true, because of the word training, the words became more attention grabbing for the new-word learners; while the words became somehow less attention grabbing for the old-word learners. Thus, even though the old-word learners have shown that they become quicker and accurate in their pronunciations of these words following training. This does not indicate that they are no longer reading the words automatically.

Melara and Algom summarize their finding to say that in the Stroop task attention is induced by information, in that way it is mimicking life. If the participants in the current study were affected by selective attention rather than automaticity, then the question that is answered by the picture-word Stroop findings would change from, “which words are read more automatically?” to “which words demand the most attention?”. More research must be done to determine which question is the right one to ask.

The purpose of running both experiments with adults was to examine the influence of word units on the reading of proficient readers. By comparing results of this experiment using adult readers to studies using younger children, we can learn how word units are processed as a function of task as reading becomes efficient. The modified Stroop results can be compared directly with those of Strauss (2001) who studied early readers. Both his and my study suggest that words with high predictability GPC units are more likely to interfere with picture naming. Although the younger readers from Strauss' study had longer reaction times and made more errors than the adults in the current study, the mean reaction time for the different word conditions paralleled for these two studies. The nonword reading task results corresponds with some of the previous research done in the subject area (Treiman, 1983; Glushko, 1979) that indicates rime units are most influential in nonword reading with adults. On the other hand, it conflicts with the results of other research in the area that was done with early and proficient readers (Brown & Deavers, 1999; Coltheart & Leahy, 1992) that says argues that although more advanced readers utilize rime units to a higher degree than beginning readers, it is still GPC that dictates pronunciation with unfamiliar or nonwords.

A more recent study supports the notion that English readers make use of both small-unit and large-unit strategies when reading (Goswami, Ziegler, Dalton, & Schneider, 2003). Goswami et al compared English readers to German readers on a nonword task in this study, as English words are said to have less consistent pronunciations overall than German words. The researchers used three kinds of nonword lists: small-unit nonwords (regular GPC as defined previously), large-unit nonwords (regular rime), and a mixed list that combined both types of words. It was found that English readers performed more accurately and quickly on small-unit lists and large-unit lists than on the mixed list, showing what the authors called a "switching cost". (The German readers did not show this same effect.) This "switching cost" is the price paid for going back and forth between small and large-unit processing. These results support that in order to read successfully, English readers must use both small-unit and large-unit strategies.

The major issue for interpreting the current findings was the lack of generalizability over items. It is unclear whether the pattern of results could be attributed to the operation of just a few items or whether this lack of generalizability occurred because of low statistical power of the analysis. Treating the items as the subjects in the analysis did not show a significant main effect of condition, possible because of the low number of items on the Stroop task. It would be necessary to increase the number of items, not only in the modified Stroop task, but also in the nonword reading list. Given that the same pattern of results has been demonstrated (and an additional unpublished study examining individual differences in first graders, Schwanenflugel, Straus, Sieczko, Kuhn, Morris, & Stahl, 2003) I believe that the same hypotheses would be supported, only more sufficiently.

Although this study was conducted with adults, these results carry implications for instruction. The ultimate goal is to give children the tools they need to become skilled readers more easily and quickly. It was important to examine which unit sizes were most influential in successful adult reading to determine what should be introduced to our early readers so they may also be successful. The modified Stroop findings support the use of phonics-based teaching strategies in the early grades, to better assist the students in becoming more automatic readers. The nonword reading task supports analogy or rime units based instruction to improve strategic reading. The need for instruction including both small units and large units is supported by the experiments collectively. While GPC appears to have more influence in automatic reading, rime units are most helpful in strategic reading. Both skills are important to the process of becoming a fluent reader.

REFERENCES

Berndt, R.S., Reggia, J.A., & Mitchum, C.C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers, 19(1),* 1-9.

Brown, G.D.A. & Deavers, R.P. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Child Psychology, 73,* 208-242.

Brown, T.L., Gore, C.L., & Carr, T.H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic"?. *Journal of Experimental Psychology: General, 131(2),* 220-240.

Coltheart, V. & Leahy, J. (1992). Children's and adults' reading of nonwords: Effects of regularity and consistency. *Journal of Experimental Psychology, 18(4),* 718-729.

Comalli, P.E., Jr., Wapner, S., & Werner, Heinz (1962). Interference effects of Stroop color-word test in childhood, adulthood, and aging. *Journal of Genetic Psychology, 100,* 47-53.

Cycowicz, Y.M., Friedman, D., Rothstein, M, & Snodgrass, J.G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65,* 171-237.

Ehri, L.C. (1976). Do words really interfere in naming pictures? *Child Development, 47,* 502-505.

Ehri, L.C. & Wilce, L.S. (1979). Does word training increase or decrease interference in a Stroop task? *Journal of Experimental Child Psychology, 27,* 352-364.

Glushko, R.J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5(4),* 674-691.

Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology, 56*, 443-475.

Goswami, U., Ziegler, J.C., Dalton, L, & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics, 24,* 235-247.

Guttentag, R.E. & Haith, M.M. (1978). Automatic processing as a function of age and reading ability. *Child Development, 49,* 707-716.

LaBerge, D. & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Logan, G.D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly, 13(2)*, 123-146.

Macleod, C.M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109(2),* 163-203.

Melara, R.D & Algom, D.(2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110(3),* 422-471.

Schadler, M, & Thissen, D.M. (1981). The development of automatic word recognition and reading skill. *Memory and Cognition, 9(2),* 132-141.

Stanovich, K.E., Cunningham, A.E., & West, R.F. (1981). A longitudinal study of the development of automatic recognition skills in first graders. *Journal of Reading Behavior, 13(1),* 57-74.

Strauss, G.P. (2001). The influence of unit size on the development of Stroop interference in early word decoding. Unpublished undergraduate honors' thesis, University of Georgia, Athens.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Child Psychology, 18,* 643-661.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15,* 49-74.

Treiman, R., Mullennix, J., Bijeljac-Babic, R. & Richmond-Welty, E.D. (1995). The Special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124,* 107-136.

Venezky, R.L. (1970). *The structure of English orthography*. The Hague, The Netherlands: Mouton.

Zeno, S., Millard, and Duvvuri (1995). The educator's word frequency guide. Touchstone Applied Science Associates.

APPENDIX A

INDIVIDUAL WORDS AND CORRESPONDING PICTURES FOR EACH CONDITION

HI GPC + LOW RIME
Doll (apple)
Dose (book)
Gas (window)
Mood (fish)
Shall (suitcase)
Yes (penguin)
Fog (lion)
Bead (kite)
Tooth (arrow)
Booth (camel)

HI GPC + HI RIME
Deck (sun)
Pet (necklace)
Sick (clock)
Deer (corn)
Name (door)
Job (pants)
Dear (piano)
Lock (snowman)
Thin (flag)
Coin (ladder)
Rock (cat)

LOW GPC + HI RIME
Ball (frog)
Bar (whistle)
Jar (hanger)
Loss (chair)
Tall (heart)
Turn (candle)
Third (owl)
Bore (carrot)
Car (house)

NON-WORDS
Keck (bell)
Zet (bread)
Gick (elephant)
Weer (broom)
Bame (ear)
Pob (monkey)
Cear (boot)
Qock (brush)
Dhin (cheese)
Noin (truck)

LETTER STRING
Gseoq (foot)
Sviw (hand)
Qlri (hat)
Dsbx (key)
Rpce (pear)
Wfom (pencil)
Udsep (scissors)
Ixr (duck)
Bocpe (toothbrush)
Amrtn (hammer)

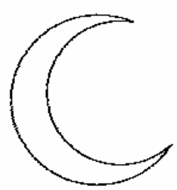
PICTURE CONTROL
(bowl)
(car)
(dog)
(fork)
(horse)
(moon)
(pumpkin)
(table)
(zebra)
(yoyo)

Note:
BOLD= word ( )= picture

APPENDIX B

SAMPLES OF PICTURE STIMULI

1.

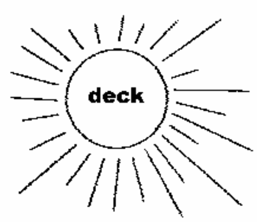


2.

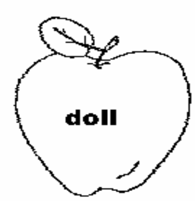
3.

4.

5.



6.



* <u>Note.</u> 1= Control, 2 = Letterstring, 3 = Nonword, 4 = Low GPC + High Rime Unit Consistency, 5 = High GPC + High Rime Unit Consistency, 6 = High GPC + Low Rime Unit Consistency

APPENDIX C

NONWORD READING TASK

nime (crime)

dalf (half)*

durn (turn)

nalm (palm)*

jiss (miss)

jild (child)*

nisk (disk)

dalk (talk)*

klarch (march)

polt (bolt)*

pobe (globe)

yalt (salt)*

heech (speech)

jask (mask)*

pirt (skirt)

deld (held)

nold (cold)*

dace (space)

kealth (health)*

dook (book)*

yoan (loan)

boup (group)*

Note:

Bold=nonword, ( )= corresponding word,

* = irregular