GENERALIZATION AND MAINTENANCE IN APHASIA REHABILITATION

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Neuroscience

By

Elizabeth H. Lacey, B.A.

Washington, D.C.
May 26th, 2010
Cognitive rehabilitation after stroke or brain injury is often desired by patients but dismissed by doctors and insurance companies. However, it has been shown to be effective, even for patients whose deficits persist for more than one year post-stroke. The following studies investigate two important factors in the rehabilitation of language disorders: generalization and maintenance. Chapter 2 examines the reasons for the success of Multiple Oral Re-reading, a non-invasive, easily administered treatment that has been reported to generalize and is currently in clinical use for two acquired reading disorders (pure alexia and phonological alexia). The treatment consists of reading text aloud multiple times a day. We hypothesized that MOR generalizes because of the repetition of high frequency words in text. We designed text passages to test the hypothesis that participants would show generalization only with untrained text that included a critical mass of the words contained in the passages they re-read in treatment. We further hypothesized that the improvement patterns would differ in the two types of alexia. Contrary to the conclusions of previous studies, our results indicate that generalization effects in MOR are due to the repetition of specific words in text. However, most patients also showed improvement when specific phrases were re-used, indicating that practice of difficult words in context may be crucial to reading improvement. Chapter 3 examines the effect of error-production during treatment on initial learning and long-term maintenance of a behavioral intervention for word-finding (anomia). We compared two versions of the same treatment: an errorful (EF) version in which
guessing is encouraged and errors occur frequently and an errorless (EL) version in which supportive cues greatly reduce the chances of making an error. We hypothesized that EL treatment would be associated with better maintenance, especially in patients with memory deficits, but our comparison revealed no advantage to EL. The results suggest that attentional engagement may be more relevant than error-prevention in the rehabilitation of people with aphasia. These two studies contribute to the literature on evidence-based practice in cognitive rehabilitation and have implications for cognitive models of reading and the interaction of memory and language.
Dedication

This dissertation is dedicated to Dr. Stephen Henderson Lacey, Kathy Lacey, Susie Lacey and Maggie Lacey. I have learned the secret of life by your example: Be smart. Be funny.
Acknowledgements

Thesis Adviser: Rhonda B. Friedman, Ph.D.

Thesis Committee: Darlene V. Howard, Ph.D., David P. Roeltgen, M.D., John VanMeter, Ph.D., Jeffrey R. Binder, M.D.

Co-authors of related manuscripts: Susan Nitzberg Lott, M.A., CCC-SLP, Sarah F. Snider, M.A., CCC-SLP, Aimee Syms Carney, M.A., CCC-SLP

Directors of the Interdisciplinary Program in Neuroscience, past and present: Karen Gale, Ph.D., Barry Wolfe, Ph.D, G. William Rebeck, Ph.D.

Collaborators: Malle Tagamets, Ph.D., Carlos Cortes, M.D.

Members of the Cognitive Neuropsychology Lab, past and present: Robyn Oliver, Ph.D., Christine Watson, Ph.D, Kim Wittenberg, Elizabeth Christy, Nora Watson, Anne Sperling, Ph.D., Lea Pilgrim, Ph.D., Jacquie Kurland, Ph.D., Leah Orchinik, Tejas Patil, Tamar Gefen, Molly Schaefer, Nicole Tomlinson, Monika Mellem, Lauren Ullrich

Dissertation work funded by NRSA award to Elizabeth H. Lacey: F31 DC008926, RO1 HD036019 and RO1 DC006747 to Rhonda B. Friedman
## TABLE OF CONTENTS

Chapter 1. Introduction ........................................................................................................... 1

Chapter 2. Generalization in Multiple Oral Re-reading treatment for alexia:

The parts may be greater than the whole ................................................................. 10

Method ......................................................................................................................... 16

Results .......................................................................................................................... 23

Discussion ..................................................................................................................... 26

Figures and Tables ........................................................................................................ 34-39

Chapter 3. To err is human: The potential benefit of errors to initial learning and maintenance in anomia rehabilitation ............................................................... 40

Experimental Design and Methods – Experiment 1 .............................................. 46

Results – Experiment 1 ............................................................................................. 49

Experimental Design and Methods – Experiment 2 ............................................ 49

Results – Experiment 2 ............................................................................................. 57

Discussion ..................................................................................................................... 60

Figures and Tables ........................................................................................................ 68-74

Conclusions and Future Directions ........................................................................... 75

Appendix A. List of Publications and Copyright Permissions ............................ 79-80

Appendix B. Examples of untrained reading passages from Chapter 2 .................. 81

References ..................................................................................................................... 90
Chapter 1.

Introduction

Given the extreme variability of the aphasia patient population both between patients, even those with seemingly similar lesions, and within patients in terms of daily fluctuations in ability and energy, it is important for the field of aphasia treatment research to focus on dependent variables that can be measured across a wide variety of patients with different aphasia types and levels of severity. Research that focuses on developing efficient treatments that are the most likely to result in permanent changes in language use will benefit the largest number of people with aphasia. The following chapters examine 3 different patient types: persons with anomia, phonological alexia, or pure alexia. Anomia is found in most people with aphasia and refers to their difficulty with finding the word they wish to produce or to name a picture of a common object upon confrontation. Phonological alexia is an acquired reading disorder characterized by inability to read pseudowords and a relative difficulty reading real words that are low in semantic value. Pure alexia is an acquired disorder of reading characterized by intact writing and other language skills in the presence of whole word reading deficits as measured by a monotonic increase in reading time as length of a word increases. We examine the treatment outcomes of participants with these disorders in two studies designed to investigate generalization of treatment effects (which can make treatments more efficient) and maintenance of treatment effects (which indicates a permanent change in language use if it can be shown to be long-term). Unlike treatment conducted in a clinic setting, our studies were carried out with a focus on experimental control and hypothesis testing. The studies described in the following chapters have implications for how to design more effective aphasia treatments.
Language disorders after stroke are common (21-38% of acute patients; Berthier, 2005). Though many people recover language function in the days and weeks following a stroke, many do not. In the acute stage, patients often steadily regain language function over time due to restored blood flow to key areas (Alexander & Hillis, 2008). It is difficult to determine the effect that speech therapy has at the acute stage because so many functions are returning spontaneously (i.e., without intervention). However, a 1998 meta-analysis showed that effect sizes of recovery in patients who receive speech therapy at the acute and sub-acute (3-12 months post-onset) stage are significantly larger than effect sizes for recovery in patients who do not receive therapy (acute stage: 1.15 vs .63; subacute stage: .57 vs .34; Robey, 1998).

However, many chronic patients (more than a year post-onset) continue to live with significant deficits. Though it is unclear how long spontaneous recovery can continue to occur after a stroke, chronic patients are defined as one year or more post-onset (Robey, 1998). These are the patients with whom the following chapters are concerned. By using multiple baseline designs, studies of experimental treatment interventions can show that chronic patients are stable before a treatment begins. Testing of behaviors not expected to improve before and after the treatment can show specificity of the treatment. Because the variable of spontaneous recovery is less likely to affect these data, they can more clearly delineate some of the mechanisms underlying the success or failure of aphasia treatments. Robey’s meta-analysis (1998) showed that treatment administered more than 1 year post onset had an effect size similar to that found during the subacute stage, and much larger than the effect found for untreated patients (.66 vs .05).
Group studies in aphasia rehabilitation

Multiple case studies and case series with chronic patients also reiterate the finding that chronic patients can still re-gain lost functions (e.g., Kendall et al., 2006). However, the field of aphasia rehabilitation still lacks studies performed using the gold standard for evaluation of treatment efficacy: doubleblind, placebo-controlled clinical trials. Review papers and multiple case studies have examined the potential of specific treatments to be applied to large groups of similar patients (e.g., Nickels, 2002; Leff & Behrmann, 2008), but there are several reasons why it is difficult to conduct clinical trials for specific deficits in aphasia. Aphasia patients often have multiple deficits in different combinations and at different levels of severity within the language domain as well as within the domains of vision, audition, attention and executive function. While there is some overlap of lesion location between stroke patients, which correlates with some overlap of deficits, there is still a great deal of variability.

Though clinical trials with the aphasic population are problematic, there are some aspects of rehabilitation that apply to all treatments and these can be studied in a controlled manner. Recently, the University of Florida convened a Language Work Group to propose a framework specifying important variables in conducting aphasia treatment research (Raymer et al., 2005). Several variables were identified, two of which are of interest here: generalization of treatment effects and maintenance of treatment effects. The Work Group defined these two variables as follows:

**Generalization:** “The influence of the intervention on other untrained behaviors…that may or may not have some type of systematic relation to the trained behavior”.

**Maintenance:** “The stability of an acquired behavioral change over time in the absence of continued intervention”.

3
In the following chapters, we examine generalization in a treatment for alexia and maintenance in a treatment for anomia. Though we had specific reasons for choosing to study these variables with the patients we did to target the deficits we did, the advantage of these variables is that any treatment study could be designed to investigate them and better understanding of their mechanisms could lead to improvements in several different treatments for different patients.

**Generalization**

Most treatment studies that have found generalization report that the untrained material is similar in some way to the trained material. One type of generalization involves the transfer of a trained strategy for circumvention of a deficit to items not specifically trained. Lott et al. (1994) found that a patient with severe pure alexia trained on the entire alphabet to use a tactile-kinesthetic strategy for letter naming was able to generalize the strategy to name untreated letterstrings and read words. Another type of generalization, which is much more difficult to achieve, is the transfer of treatment effects that improve the actual processes of retrieval, production, decoding and comprehension of language to untrained stimuli. For example, Kendall et al. (2008) investigated generalization in a naming treatment focused on practicing phonemes and phoneme sequencing rather than directly treating naming deficits. These authors predicted that phoneme training would generalize to the production of any word and therefore improve naming. Their results were mixed, but some patients did improve. Some also showed signs of continued improvement 3 months after training had ended, which the authors cite as evidence that patients could be integrating improved access to phonology in their daily lives, thereby continuing to improve their language without being actively involved in speech therapy. Other naming studies have shown that the effect of training items from one semantic category
can generalize to untrained items from a similar semantic category (Nickels, 2002). There is some indication that training with items that are more difficult or complex in some way generalizes to items that are less complex, such as naming treatment using atypical items (Kiran & Thompson, 2003) or sentence production treatment using more syntactically complex sentences (Thompson et al., 2003).

Because it is not feasible to develop cost-effective ways of treating patients on every word they have difficulty with, it is important for generalization of aphasia treatments to be improved. Treatments that generalize to untrained words/text are far more useful and efficient. Chapter 2 examines generalization using a common treatment for acquired reading disorders, Multiple Oral Re-reading (MOR), that has been reported to generalize to untreated passages of text. The two goals of the study were: to systematically test a prevailing hypothesis purported to explain generalization in MOR and to determine the mechanism and meaning of differences in generalization between two different alexia types.

**Generalization in Multiple Oral Rereading for Alexia**

The study described in Chapter 2 examines the reasons for the success of MOR (Moyer, 1979), a non-invasive, easily administered alexia treatment that has been reported in the literature and is currently in clinical use (Webb & Love, 1986; Kim & Russo, 2010). The treatment consists of reading text passages aloud multiple times a day. Findings that MOR improves reading speed on practiced as well as novel text have been inconsistent, making MOR’s role in the rehabilitation of alexia unclear (Tuomainen & Laine, 1991; Beeson & Insalaco, 1998). We hypothesized that MOR’s treatment mechanism works through the repetition of high frequency words in the text passages that the patient re-reads (i.e., bottom-up
processing), arguing against a hypothesis presented in other studies that MOR works by top-down processing. We designed and controlled our text passages to test the hypothesis that participants would not improve on all novel text but would improve on text that includes a critical mass of the words contained in the passages they were re-reading, thus supporting the idea that MOR works from the bottom up. We further hypothesized that we would see improvement on different types of words for the two different types of alexia patients we tested (phonological and pure). By controlling variables within the practiced and unpracticed text passages, we examine why this treatment seems to work well and generalize. Understanding the mechanisms underlying generalization in this treatment can inform its use in a clinical setting and provide information about how this treatment might be most effectively applied to patients with different deficits and different levels of severity. Further, if we understand how generalization works in this treatment, it might generate hypotheses for how other treatments could be changed to make them more likely to generalize.

**Maintenance**

The issue of maintenance is similar to the issue of generalization in that it also has implications for a wide variety of treatments. Many aphasia treatments have been shown to be successful, but decline from post-treatment levels at 3 months (Fridriksson et al 2005), 1 month (Fillingham et al., 2006), and even 1 week (Fridriksson et al., 2006). Maintenance is rarely tested at longer intervals. However, insurance covers only a certain number of sessions and it would certainly be in the best interest of patients if treatment effects maintained long after insurance coverage has ended. Follow-up testing is sometimes reported in studies of aphasia rehabilitation, but maintenance of the treatment is rarely the focus of the study (e.g., McNeil et
al., 1998; Leger et al., 2002; Kiran, 2008). Chapter 3 describes a study that was specifically designed to test the effect of error-production on initial learning and long-term maintenance. There has been much discussion in the aphasia literature about applying errorless (EL) learning paradigms, which have been successful in people with memory impairments, to the treatment of acquired aphasia (Fillingham et al., 2006; McKissock, & Ward, 2007). Despite the fact that there is still no evidence showing that EL learning is superior to more traditional, errorful (EF) approaches, an assumption seems to prevail that EL treatment paradigms are more beneficial for people with aphasia. The unfounded excitement about EL techniques carries over into the clinic. It is necessary to conduct studies comparing EF and EL designed to test the superiority of EL as a hypothesis before it becomes standard clinical practice.

The Effect of Errors on Initial Learning and Maintenance in a Treatment for Anomia

In Chapter 3, we test the hypothesis that EL learning is associated with better initial treatment outcomes and long-term maintenance than EF learning in a design that sacrifices some ecological validity for experimental control of the EF and EL variables. Our study compares EL to EF learning in a common treatment for anomia but, unlike previous studies, we treated patients in both conditions simultaneously until they reached a pre-set criterion or plateau. This ensured that participants were given enough time to show a difference in performance between the two conditions if one exists. EL learning theories are based on the idea that errors could become paired with a test stimulus and that this pairing becomes particularly detrimental in people whose memory impairment prevents them from distinguishing error responses from correct responses. In Chapter 3, we argue that the memory deficits observed in most aphasia patients are not severe enough to warrant the use of EL learning paradigms. Our data indicate
that attentional engagement, which is likely to be reduced during EL treatments, may be more relevant than error-prevention to initial learning and maintenance. The study in Chapter 3 begins to uncover some of the fundamental reasons that treatments might succeed or fail initially and at maintenance testing and challenges the notion that EL treatments are necessarily superior to EF treatments.

Conclusions

There are examples in the literature of studies that measure both generalization and maintenance. For example, Greenwald et al. (1995) reported two single-subject studies of naming treatment using phonologic and semantic cues. The phonologic cueing treatment showed some signs of generalization to untreated items. When tested 10 days after treatment, performance had declined in both treatments. Conversely, McNeil et al. (1998) found that after 8 weeks of a different type of cueing hierarchy treatment, a patient showed no generalization to untrained items but maintained the treated items at follow-up after 3 months. These studies demonstrate that testing for generalization and maintenance in any treatment gives us important, if sometimes conflicting, information. However, it would advance the field considerably if instead of simply testing and reporting the outcomes on these measures, we also conduct experiments to determine why they occur in some studies and not others. The following chapters attempt to do just that.

By studying behavioral interventions under rigorously controlled conditions, we can learn more about how they work, why they work and why they don’t work. The following chapters examine non-specific treatments for patients with a wide array of deficits. However, our detailed assessments of the patients allow us to examine if these non-specific treatments have specific
effects for certain types of patients. If we can uncover some of the factors affecting generalization and maintenance, perhaps we can design treatments that are more likely to work on both of these levels. As of now, The American Academy of Neurology recognizes only one aphasia therapy as effective: Melodic Intonation Therapy (MIT) for severe nonfluent aphasia (Alexander & Hillis, 2008). These patients represent only a fraction of people with aphasia. Even within this group, some do not respond to MIT (Næser & Helm-Estabrooks, 1985). Also, there are many patients whose deficits are not severe enough to require MIT who can still benefit from therapy. Some of the less severe patients might even be able to return to work if a therapy program is successful enough. The onus is on the research community to demonstrate experimentally why a treatment is or is not effective. Some neurologists have estimated that 60-70% of their patients are denied coverage for cognitive rehabilitation. However, they also report that their involvement in advocating for coverage can be effective (Cajigal, 2007). Hypothesis-driven research examining generalization, maintenance and other measurable variables may improve aphasia treatment and eventually improve acceptance of its efficacy among medical doctors as well as insurance companies.
Chapter 2. Generalization in Multiple Oral Re-reading treatment for alexia: The parts may be greater than the whole

Introduction

The Multiple Oral Re-reading (MOR) technique is a treatment for acquired alexia that requires patients to read the same passages of text aloud several times a day. It has been shown to increase the reading speed of the practiced passage and, more importantly, to generalize to novel text (e.g., Moody, 1988; Beeson & Insalaco, 1998; Beeson, Magloire & Robey, 2005). In the original 1979 study using the MOR technique (Moyer, 1979) the patient read a passage aloud for 30 minutes a day for one week. A new selection was introduced each week and speed of reading the practiced and novel selections improved over the three months of treatment. In the 30 years since, the success of the MOR technique has been replicated several times for patients with different types of alexia. However, no clear explanation has yet been established for why MOR works.

In the first paper to describe MOR, Moyer suggested that the treatment may work because “all components of written language structure are simultaneously maintained over practice”. More specifically, the “structure provided by the whole facilitates processing of the parts”, and vice versa, in an interactive fashion (Moyer, 1979, p. 143). In other words, she proposed that an interaction occurs between bottom-up and top-down processing. In the context of the literature on MOR, bottom-up processing refers to recognition of single word forms and/or the visual, orthographic, and phonological processes that support single word reading. Top-down processing refers to the use of the context within which words are read: the syntax and semantics of the text passages. Though Moyer originally proposed an interaction of these two processes, subsequent studies have concluded that, since the treatment effect generalizes to novel
text passages made up of different words, it follows that improvements in top-down reading processes drive the generalization (Tuomainen & Laine, 1991; Beeson & Insalaco, 1998). The first goal of the current study is to test the hypothesis that MOR’s generalization effects can be attributed to improved top-down processing.

The second goal is to determine the source of MOR’s generalization effects in two different types of alexia: pure alexia and phonological alexia. The demonstrated success of the MOR treatment is of great significance for people with alexia, and motivates further investigation of the mechanisms underlying its success. Understanding the specific causes of its success should lead to the construction of MOR training passages that are optimal for the particular alexic deficit that is targeted. It is to this aim that the second goal is directed. In pure alexia, the reading deficit occurs in the context of intact spelling, indicating that orthographic representations remain intact. Thus, pure alexia is theorized to be due to degradation of the connections between visual input and the orthographic lexicon. The result is a reading impairment that is more severe for long words compared with short words, but that does not differ according to the syntactic class of the words. In phonological alexia, length is not a factor in reading success. However, people with phonological alexia have poor pseudoword reading as compared to reading of real words, and they typically have difficulty reading functor words and affixed words in isolation (Friedman, 1995) and/or in text (Friedman, 1996). According to the model of reading presented in Figure 1, pure alexia arises from damage within the visual system, prior to accessing the orthographic lexicon (Friedman & Alexander, 1984). Phonological alexia could arise from damage to the connections between orthography and phonology or to the phonological lexicon itself (Friedman, 1995). Based on this or similar cognitive models of reading, it is possible to predict different ways in which MOR treatment might work for these
two alexia types.

Previous studies have tested top-down vs. bottom-up hypotheses to explain the effect of MOR for both types of alexia by simultaneously measuring improvements in text reading and single word and/or pseudoword reading (Beeson & Insalaco, 1998; Tuomainen & Laine, 1991). However, no study has yet examined how individual words within the practiced passages are improving, nor has any study tested phonological and pure alexia patients as part of the same experiment. In a study of MOR in 3 pure alexia patients, Tuomainen & Laine (1991) sought to determine whether MOR has its effect by acting directly on the word form system (i.e., through bottom-up processing), which would be reflected by improvement on single words as well as text, or whether MOR works through “semantic and syntactic constraints” (top-down processing), which would be reflected by improvement to text alone. The authors favor a top-down processing account to explain their data, though one of their pure alexia patients improved on both text and single words, one improved only on text, and one did not show improvement.

Beeson and Insalaco (1998) evaluated the MOR technique with phonological alexia patients. Following treatment, both patients’ text reading speed improved for novel text as well as for single words. One patient was only five months post-stroke, making interpretation of those data difficult. In their interpretation of the patient who began treatment one year after her stroke, the authors note that the patient improved on reading of functors in isolation more so than she improved on nouns, adjectives or verbs after MOR treatment. The authors theorize that this could be due to repetition of high frequency functor words during oral re-reading (through bottom-up processes) or to improved access to functor words in isolation through practicing of the “syntactic frames” provided by the text (through top-down processing). The current study is designed to examine these two possibilities further.
Individual words in the passages used for re-reading, in addition to those used to assess generalization to novel text, need to be tightly controlled in order to tease apart the top-down vs. bottom-up explanations of the MOR treatment effect. By definition, high frequency words can be expected to appear in any text used in the practiced as well as the novel passages. Previous studies have not been able to determine to what degree repetition of high frequency words affected their results, because they did not control the individual words used in the training and novel text passages. We designed our study to address this methodological issue.

Our study focuses on people with two types of alexia: mild pure alexia and “phonological text alexia”, a term for a mild phonological alexia first described by Friedman (1996). Our participants with pure alexia read faster than the mildest patient reported by Behrmann (1998) and about twice as fast as the two successful patients reported by Tuomainen and Laine (1991; patients HT and TT). Our participants with “phonological text alexia” are comparable to Beeson and Insalaco’s (1998) patients and other mild phonological alexia patients reported in the literature (Crisp & Lambon Ralph, 2006; patients DB and TH) in terms of reading error patterns, but our participants had to begin treatment reading at a higher rate of speed than those previously reported in order to be included in our study. Friedman theorized that phonological text alexia exists on a continuum with phonological alexia (Friedman, 1996). Patients with phonological text alexia have impaired pseudoword reading, but relatively intact single word oral reading, including function word (functor) reading. However, oral text reading, particularly of functionals and affixes in text, is impaired (Friedman & Lott, 1995; Friedman, 1996). In terms of behavioral presentation, then, the restriction of functor reading errors to text is what distinguishes patients who are at the level of phonological text alexia on the continuum from those at the level of phonological alexia. Since text reading is where these patients’ deficits are most apparent, MOR
is a particularly suitable treatment for them. The treatment is also suitable for high-level pure alexics who are able to access wordforms but can benefit from improved reading speed.

The current study is unique in that we controlled the functors, content words, and phrases used in the untrained passages and assessed generalization by timing participants’ reading of these passages before and after each week of MOR therapy. Our untrained passages contain the same number of words as the trained passages, but differ in terms of which words from the trained passages are re-used (functors, content words, phrases, or a minimal overlap of words). If, as others have hypothesized, generalization effects can be attributed to improved top-down processing, we would expect equal improvement across all untrained passages after treatment, including those containing minimal trained words. An alternative hypothesis is that MOR’s generalization effects could be due to bottom-up improvements in specific words or groups of words. Our tightly controlled test passages allow us to track what types of words are improving as text reading improves in order to differentiate the two hypotheses.

We administered MOR therapy to two participants with mild pure alexia and four participants with phonological text alexia. Our design required that, in a single two-hour session, the participants read 8 text passages 300 words in length as well as practicing that week’s training passage with the Speech Pathologist. Most of the patients reported in MOR studies previously would likely be unable to complete all of this reading without fatigue, but it was required in the current study in order to test our hypotheses. Therefore, we included only mildly impaired patients in this study.

It should be noted that even within diagnostic category (pure or phonological text alexia) the participants were dissimilar in their lesion sites and cognitive profiles prior to therapy. In order to take into account the similarities and differences when evaluating our hypotheses, we
use a “case-series methodology” here (Lambon Ralph, Moriarty & Sage, 2002). This method allows detailed discussion of each participant (as in a single-case study) as well as comparison across participants (as in group studies).

Our predictions for a bottom-up processing account of MOR generalization effects for the two alexia types studied here are as follows: we predicted that the phonological text alexics would improve on the phrase passages, the functor passages or both, but not on the content passages (performance on these words should be at ceiling pre-treatment). Repetition of the functors within the training passages should improve access to them, and this should transfer to untrained passages containing those same functors. As originally suggested by Moyer, difficult words may need to be practiced in context in order for access to them to improve. Our untrained Phrase passages reflect the benefit of practicing not just single, difficult words in context, but the specific phrases that contain some of these words.

Unlike patients with phonological alexia, patients with pure alexia were not expected to be affected by the part-of-speech of the trained words re-used in the untrained passages. If the reinforcement of visual-orthographic connections to the specific words in the practiced paragraphs is responsible for the reduction in reading times for pure alexia patients after MOR, then the pure alexia patients would improve on all three types of untrained passages (Functor, Phrase and Content) because all three contain trained words. However, they would not improve on passages with minimal overlap of words from the trained passages. Each time a word is seen and read aloud during training, visual-orthographic links are reinforced. Therefore, if that word is seen again in a novel context, time to access the word would be expected to decrease. Improved access would be the same whether the repeated items are content words, functor words or phrases.
Finally, it is possible that, as others have suggested, MOR improves the use of top-down processing (Tuomainen and Laine, 1991) rather than affecting the underlying deficit and improving reading from the bottom up. If this is the case, then the benefits of MOR therapy should be the same for pure alexia and phonological text alexia. Participants should improve on any untrained passage, including those with no practiced words, as all passages should benefit equally from improved top-down processing.

Method

Participants (See Tables 1 and 2)

**Phonological Text Alexia Participant 1 (PhTA1).** PhTA1, a 60 year-old woman, was diagnosed with cerebral amyloid angiopathy, which was the likely cause of the three successive intracranial hemorrhages she suffered in left temporo-parietal and occipital areas. Although she had a right visual field cut, she did not have a length effect. Rather, her text reading was slow and characterized by errors on functors and affixes.

**Phonological Text Alexia Participant 2 (PhTA2).** PhTA2, a 61 year-old man, suffered an infarct in the left middle cerebral artery, resulting in damage to the left posterior frontal region, extending caudally to post-central anterior parietal cortex, and ventrally to the posterior frontal opercular aspects of the Sylvian fissure. His text reading was slow and he made errors primarily on functors and affixes.

**Phonological Text Alexia Participant with Low Reading Accuracy 1 (PhTA3).** This participant, a 69 year-old woman, suffered an infarct in the left middle cerebral artery, affecting nearly all of the left frontal lobe, as well as posterior temporo-parietal regions, including part of Wernicke’s area. In the acute stage, she presented with non-fluent aphasia. When she entered our study 9 years later, her speech was fluent but included frequent phonemic paraphasias and
some word-finding difficulties. Prior to entering the current study, PhTA3 completed a separate experimental treatment protocol in our lab, as described elsewhere (Lott et al., 2009). Her reading was slow and included many functor and affix errors and substitutions. This participant was unable to achieve our cut-off reading accuracy score of 90% on the experimental passages.

**Phonological Text Alexia Participant with Low Reading Accuracy 2 (PhTA4).** This participant, a 39 year-old woman, suffered a left hemisphere stroke resulting in a lesion encompassing most of the left insula and affecting the temporal lobe from the temporal pole to Wernicke’s area. The parietal and frontal opercula were also damaged. Her reading was slow with errors primarily on functors and affixes. Her oral single word reading included errors on multi-syllabic words. This participant was also unable to achieve our cut-off reading accuracy score of 90% on the experimental passages.

**Pure Alexia Participant 1 (PA1).** This participant, a 53 year-old man, suffered a left hemisphere stroke affecting the occipital and medial temporal lobe. He has a small area of blurred vision in the upper right visual field. He had returned to work at the time of the study and presented as highly functional on all language tests, though oral text reading was somewhat halting with occasional errors on word-endings. He showed no measurable length effect, but was considered pure alexic because his reading was impaired in comparison to his pre-morbid abilities and to his writing skills, which were intact (perfect scores on all writing sections of the Boston Diagnostic Aphasia Examination (BDAE; Goodglass, Kaplan & Barresi, 2001).

**Pure Alexia Participant 2 (PA2).** This participant, a 62 year-old man, suffered a closed head injury when he was hit by a car. Primary impact caused a fracture of the right frontal bone/orbital roof/medial orbit causing a lesion in the right frontal lobe. He sustained left occipital and temporal hemorrhagic contusions as a result of contrecoup injury. He has a field
cut in the upper right quadrant of the visual field. He performed well on most language tests. He showed a slight length effect (average speed for reading 3 letter words was 870 msec, 5-6 letter words was 1253 msec and 7 letter words 1525 msec) and made occasional, self-corrected errors, usually on word-endings in text reading (for example, he read “provide” as “project”), but his writing was intact (perfect scores on all writing sections of BDAE (Goodglass, Kaplan & Barresi, 2001), except for writing “tomatoe” for “tomato”).

(TABLE 1 about here)

Participant Testing

Table 2 shows participants’ scores on sections of the BDAE (Goodglass, Kaplan & Barresi, 2001), the RCPM (Raven’s Colored Progressive Matrices; Raven, 1976) and TONI (Test Of Nonverbal Intelligence; Brown, Sherbenou & Johnson, 1997); the latter two tests are standardized tests of intelligence that use visuospatial rather than linguistic stimuli. Table 3 shows participants’ scores for a pseudoword reading list developed in our lab. The pseudowords are 3-4 letters in length and were created by changing one letter in each of a list of matched real words. Participants read these two lists on separate testing days.

Oral and silent reading speed and accuracy were assessed using standardized reading passages 1, 3, and 5, and 2, 4, and 6, respectively, from the Gray Oral Reading Test (GORT III, Wiederholt & Bryant, 1992). Passages from Form A were used for pre-testing, while passages from Form B were used for post-testing (see Table 3). As in Moyer (1979), reading speed, not comprehension, was the focus of the treatment. Therefore, comprehension of our experimental training and generalization passages was not assessed, but general reading comprehension was
evaluated pre- and post-treatment with the GORT. Silent reading comprehension was also assessed using passages from the Gates-MacGinitie reading test (1965).

The four participants with phonological text alexia show deficits on pseudoword reading and make functor and affix errors in text. The two participants with pure alexia show no pseudoword reading deficits and their errors on word endings are not exclusive to affixed words. Based on these patterns of deficit, the alexia diagnosis categories in which we placed these participants are appropriate (see Tables 2 and 3). (TABLES 2 and 3 about here).

Stimuli

Stimuli consisted of trained and untrained text passages (See Appendix B for examples). There were five types of text passages: the training (practiced) passages; three types of untrained passages used to assess the source of generalization to novel text; and neutral control passages, which contained minimal overlap of words with the trained passages, as described below.

Training passages came from an educational workbook at grades 6-8 reading levels (Instructional Fair, Inc., 1990). These passages had minimal dialogue, few words or names with challenging pronunciations, and came from the earliest parts of the workbook, which progressed gradually in difficulty. A total of eight passages were edited and abridged to 300 words in length, and each was printed on a single page in Times New Roman 15-point font. The words in each training passage were coded as functors or content words according to their syntactic role in the training passage. Some words were coded as ‘other’ and were not included in analyses or word counts when creating the passages used to assess generalization. ‘Other’ words included adverbs and inflectionally affixed words (in which the affix does not change the part of speech of the root word). Including these words would make interpretation of the results of the content passages difficult, as we would be unable to determine if errors were due to the affix or to the
content word itself. The articles ‘a’, ‘an’, and ‘the’ were also coded as ‘other’. Phrases of 3-5 words in length were also coded and marked in the training passages.

Using the coded phrases, functors and content words from each training passage, we created the following three types of new passages to be used as untrained passages:

1. A Phrase passage sharing 60-80% of the phrases contained in the training passage, while differing in overall narrative structure and content.

2. A Functor passage sharing 60-80% of the derivationally affixed words and functors contained in the training passage, with minimal overlap of phrases or content words and differing in overall narrative structure and content.

3. A Content passage sharing 60-80% of the content words contained in the training passage, with minimal overlap of phrases, a minimal number of shared affixed words, and, to the extent possible, a minimal number of overlapping functors as well. These passages also differed in overall narrative structure and content from the training passage.

The Phrase, Functor and Content passages were never trained. They were used as testing material only, before and after their respective Training passage was trained. Each passage contained 295-300 total words, no dialogue, no words with challenging pronunciations, no repetition of proper nouns or of words coded as ‘other’ (with the obvious exception of the articles ‘a’, ‘an’ and ‘the’). The passages differed completely in subject matter from each other and from the Training passages. Words re-used to create the untrained Phrase, Functor and Content passages had the same meaning and/or part of speech when used in the Training passage. All untrained passages contained an equivalent number of practiced words. All 32
passages were matched in terms of the total number of content, functor, and “other” words per passage.

In addition, four Control passages were created to control for the possible effects (on reading speed and accuracy) of simply re-reading the same passage after a week’s delay, regardless of any specific training or practice. Control passages were constructed in a similar fashion to the other untrained passages, but contained no trained content words and minimal trained functors from that week’s, or any previous weeks’ Training passages.

Total time to read each passage aloud was assessed. We analyzed the reading time results of participants who read at or above 90% accuracy and used accuracy as the measure of improvement for those participants who read below 90% accuracy. The inaccurate reading of a significant number of words would distort measures of speed and lead to uninterpretable reading time data for the passages.

**Procedure**

**Overall.** Treatment consisted of one two-hour session per week for 8 weeks. Each week a different Training passage was trained. The Control, Phrase, Functor and Content passages for that Training passage were never trained, but were tested before and after training.

**Experimental Testing.** Each session began with post-tests of the prior week’s Control passage (if applicable), training passage, and modified Phrase, Functor and Content passages, followed by pre-tests of the next week’s passages. The participant was instructed to read each passage out loud from beginning to end without stopping, as quickly and accurately as possible. Reading speed was recorded with a stopwatch and errors were recorded during testing by the experimenter. No feedback regarding accuracy or response time was provided during pre- or post-testing.
Post-tests were interleaved with pre-tests such that post-testing of the previous week’s stimuli always preceded pre-testing of the upcoming week’s test passages. For example, the order of testing for Week Two was as follows:

1. Week One passage POST-tests – Reading aloud: Control passage one, Training passage one, untrained Phrase passage one, untrained Functor passage one and untrained Content passage one.
2. Ten minute break.
4. Training with feedback for Training passage two.

Most sessions included both pre- and post-testing except the first session (no pre-testing) and the last session (no post-testing)

Training. After the week’s pre- and post-testing were completed, treatment for the week’s Training passage began. Though training with feedback was not always the method used in previous studies, we felt that correcting errors during training was better for the participants. Also, higher accuracy results in more precise measurements of reading times. During training, the participant was told that should s/he make a mistake, the experimenter would stop him/her and point to the incorrect word for him/her to re-read. The participant was instructed to re-read only the specific word, not to go back to the beginning of the sentence. If the participant could not read the word correctly after this cueing, the experimenter said it for him/her. The participant was only instructed to repeat whole sentences if s/he made enough errors to completely lose the flow of the sentence. The participant read the training passage three times in this manner with cueing from the experimenter. This concluded the training session for the
passage, and the participant was then given a copy of the passage for home practice. In order to increase the likelihood that the participant was reading the passage correctly at home, the participant was instructed to call the experimenter’s office once every day for the next week and read the passage over the phone with feedback provided. The participant also practiced the passage 5 more times each day, unassisted. To encourage participants to carry out home practice as instructed, they were provided with written log sheets on which to record each home practice session, which they then reviewed with the experimenter at each treatment session.

**Post-Testing.** After the final week of MOR therapy, reading (including Form B of the GORT) and language tests were re-administered to assess any changes in profile. Those tests in which a participant had achieved a score within normal range prior to therapy were not re-administered (See Tables 2 and 3).

**Results**

(Figure 2 shows reading speed for participants PhTA1, PhTA2, PA1 and PA2. Figure 3 shows reading accuracy scores for PhTA3 and PhTA4. A repeated measures ANOVA was used to compare pre- to post-treatment performance for each passage type within subjects. Subjects were not compared to each other statistically)

**Participants with Phonological text alexia (PhTA1, PhTA2)**

Figure 2 displays the mean number of syllables read per second for each of the five passage types (Control, Training, Phrase, Functor, Content) pre- and post-training, averaged across the 8 weeks of treatment (except for Control passages (n = 4)), for PhTA1 and PhTA2. The interaction of testing time by passage type was significant for both participants, \( F(4, 31) = 16.02, p < .0001 \) and \( F(4, 31) = 9.10, p < .0001 \), respectively. Planned pair-wise comparisons
between pre and post-training for the five passage types also revealed significant increases in syllables/second for reading the Training passages for both PhTA1 (mean change in syllables/s = 0.43, $t = 8.27, p < .0001$) and PhTA2 (mean change = 0.65, $t = 4.02, p < .01$). For the untrained, generalization passages, the Phrase passages differed significantly in speed between pre and post treatment for PhTA1 (mean change = 0.10, $t = 3.74, p < .01$), and the Functor passages differed significantly for PhTA2 (mean change = 0.19, $t = 5.05, p < .01$). The Content passages did not improve significantly for PhTA1 ($t = 1.13, p > .05$) or PhTA2 ($t = 1.30, p > .05$), nor did the Control passages (PhTA1: ($t = 1.52, p > .05$) or PhTA2: ($t = .20, p > .05$)). ANOVAs of speed of reading revealed no difference for passage type pre-treatment for either PhTA1 or PhTA2 ($F(4, 31) = 1.08, p > .05$ and $F(4, 31) = 2.11, p > .05$, respectively), but highly significant differences post-treatment (PhTA1: $F(4, 31) = 22.21, p < .0001$; PhTA2: $F(4, 31) = 20.37, p < .0001$).

Overall pre- and post-treatment accuracy was high for PhTA1 (97.7% and 98.1%) and PhTA2 (95.5% and 96.4%). (FIGURE 2 about here)

For both participants, average syllables read per second on the GORT form B, both orally and silently, increased after treatment, while comprehension remained relatively high (see Table 3). In addition, PhTA2 showed a significant improvement in pseudo-word reading from 5/20 pre-treatment to 12/20 post-treatment (McNemar $p=0.019$, one-tailed).

**Participants PhTA3 and PhTA4 (Phonological text alexia, low reading accuracy)**

PhTA3 and PhTA4 were unable to read the passages at 90% accuracy pre-treatment, making measurements of their speed of reading invalid. Instead, we report their accuracy results here and in Figure 3. PhTA3’s accuracy on the Training passages improved significantly from pre- to post-treatment (74.1% to 90.1%; mean change in percent correct = 16.5 $t = 9.89$, $p<.0001$) as did PhTA4’s (89% to 94%; mean change in percent correct = .054, $t = 5.29, p<$
PhTA3’s accuracy in reading the generalization passages improved after treatment, but was still quite low (range= 66% - 76%). Figure 3 shows the changes in her accuracy from pre- to post-treatment for the different generalization conditions. Planned pairwise comparisons revealed significant increases in accuracy for the Functor and Phrase passages for PhTA3. The mean change in percent correct for the Phrase passages was 8.6%, \( t = 6.81, p < .0001 \) and for the Functor passages 16.5%, \( t = 9.89, p < .0001 \). PhTA4 showed significant improvement only on the Phrase passages (from 87% to 90% correct; mean change = .027, \( t = 2.69, p < .05 \)). Neither PhTA3 nor PhTA4 showed significant changes in accuracy on the Control passages (PhTA3: \( t = .64, p > .05 \) PhTA4: \( t = .25, p > .05 \)) (FIGURE 3 about here).

After MOR therapy was completed, PhTA3 and PhTA4 were re-tested with GORT Form B and showed an increase in syllables per second on oral and silent reading. Comprehension remained stable or improved for PhTA4 and decreased for oral reading for PhTA4 (see Table 3). PhTA3 and PhTA4 also showed an increase in pseudoword reading accuracy, though this improvement did not reach statistical significance. PhTA3 improved from 1/20 to 5/20 (McNemar \( p = 0.063 \), one-tailed) and PhTA4 improved from 1/20 to 6/20 (McNemar \( p = 0.109 \), one-tailed) after MOR therapy.

PhTA4 also showed significant improvements on the Boston Naming Test (BNT) (McNemar, \( p = .019 \), two-tailed) and the oral word reading section of the BDAE (Goodglass, Kaplan & Barresi, 2001) (Wilcoxon, \( p = .039 \), 2-tailed; see Table 2).

**PA1 and PA2 (Pure alexia)**

Figure 2 displays the mean number of syllables read per second for each of the five passage types pre- and post-training, averaged across the 8 weeks of treatment (except for Control passages \( n = 4 \)), for participants PA1 and PA2. The interaction of testing time by
passage type was significant (PA1: $F(4,31) = 26.08, p < .0001$; PA2: $F(4, 31) = 25.72, p < .0001$). Planned pair-wise comparisons of the five passage types revealed significant increases in syllables/second for reading the Training passages for PA1 (mean change in syllables/s = 0.88, $t = 8.87, p < .001$) and for PA2 (mean change in syllables/s = 1.00, $t = 10.23, p < .0001$). As predicted, all generalization passages that overlapped with Training passages improved significantly for PA1 (Phrase: $t = 4.79, p < .01$; Functor: $t = 5.22, p < .01$ Content: $t = 3.84, p < .01$). For PA2 all generalization passages improved (Phrase: $t = 3.48, p < .05$; Content: $t = 4.27, p < .01$) except the Functor passages ($t = 1.00, p > .05$). Also as predicted, the Control passages did not improve significantly for PA1 ($t = 3.058, p > .05$) or PA2 ($t = 2.55, p > .05$). ANOVAs of speed of reading (syllables/s) revealed no difference for passage type pre-treatment (PA1: $F(4, 31) = .85, p > .05$; PA2: $F(4, 31) = 2.57, p > .05$), but highly significant differences post-treatment (PA1: $F(4, 31) = 28.32, p < .0001$; PA2: $F(4, 31) = 39.53, p < .0001$). Pre- and post-treatment accuracies for all passages for both participants were close to 100%.

Average syllables read per second both orally and silently on the GORT form B increased after treatment while comprehension remained high in both participants (see Table 3).

**Discussion**

Our main hypothesis regarding the effects of MOR was borne out by both the phonological text alexia patients and the pure alexia patients. That is, the MOR technique improves the reading of untrained passages that overlap with trained passages at the level of the specific alexic deficit. The treatment is simple to administer and, though it is repetitive, the participants reported here enjoyed it. They liked having homework outside of therapy, and they
found the feedback during training helped them to understand which words were making reading difficult for them.

Reading rate for novel passages containing practiced functors and/or phrases increased after MOR treatment for participants with phonological text alexia (difficulty reading functors in text). Reading rate for novel passages containing a significant number of practiced words of any kind increased after MOR treatment for participants with pure alexia (slowed identification of all types of words). Reading rate for novel passages containing a minimal number of words from the training passages (Control condition) did not increase after MOR treatment for any of the participants. The lack of a generalization effect for truly novel text stands in contrast to previous studies hypothesizing that generalization to novel text is due to top-down processing and indicates a bottom-up mechanism for MOR’s treatment effect.

Our results for the Training passages are consistent with previous literature (e.g., Moyer, 1979; Beeson & Insalaco, 1998) in that all patients improved on these passages through practice. Although we did not measure comprehension for the experimental passages, the results of the GORT indicate that comprehension did not suffer when reading speed increased (see Table 3). In the four patients for whom we could reliably measure speed, we also replicated previous studies that found generalization of the treatment effect to novel reading material differing in content and narrative structure. None of our patients showed a significant difference in speed or accuracy between pre and post-treatment time-points on the Control passages designed to test generalization to a truly novel passage. These results are inconsistent with a top-down processing explanation for MOR’s effect, which predicts improvement on all passages, regardless of how they overlap with the training passages. To reliably test the top-down processing hypothesis, it was necessary to work with patients who could read at a certain speed.
Though this exclusionary criteria made our treatment effect sizes smaller than what has previously been reported, it also gave us the controls we needed in the design to test the top-down processing hypothesis properly.

Our results suggest that bottom-up improvements specific to the deficits of each patient play an important role in MOR’s treatment and generalization effects. The results from the untrained Content passages are the first line of evidence to support our bottom-up hypothesis. The Content passages were designed to reflect the outcome of practicing content words during re-reading of the Training passages. The results for the phonological text alexia patients tell us that practicing these content words does not significantly contribute to generalization to novel text. Also supporting the bottom-up hypothesis is the finding that these patients did improve on the Functor and/or Phrase passages indicating that a critical mass of functors or phrases containing functors from the practiced passage must be included in an untrained passage in order for generalization effects to be measurable for patients with phonological text alexia. Previous studies have investigated this idea. Beeson and Insalaco (1998), whose patients were also mild phonological alexics, found a decrease in reading time for functors presented in isolation after MOR treatment. Though their ultimate conclusion is that “context effects” are responsible for their patients’ improvement on novel text, they also discuss the possibility that the reason their patients improved on single functor reading was that the high frequency of functors in the training passages allowed more repetition of functors than of words of any other part-of-speech. Our results favor this latter interpretation. Our study controlled for the number of functors and which functors were re-used in our untrained passages. Therefore, we were able to determine that our phonological alexia patients benefited from repetition of the specific functors and/or phrases that they practiced. Without sufficient overlap of functors and/or phrases, as in the
Content passages, no improvement was seen. Lack of improvement on the Content passages in the phonological alexia patients argues against the top-down processing hypothesis to explain generalization in MOR. Again, top-down processing would be expected to affect all passage types equally.

Further supportive evidence comes from the four Control passages, which represent a negative control condition, in that they were written using as few words that overlap with words in the Training passages as was feasible. Neither the phonological nor the pure alexia patients improved on the Control passages. If top-down processing were the driving factor behind MOR, we should have seen improvement on these passages. This condition is particularly important in the case of the pure alexia patients. They were expected to improve on all untrained passages that overlapped with the Training passages; therefore, only the Control condition can show their level of generalization to a passage that is truly novel text (i.e. with minimal overlap of any practiced words, including functors). Neither pure alexia patient showed generalization effects on these passages.

We argue against a top-down processing or “context effects” account for MOR because our results indicate that a significant part of the mechanism behind MOR is bottom-up. However, both theories, due to their grounding in single word reading models, are probably too simplistic to explain the phenomenon of generalization after MOR training. That five of the six participants presented here showed improvement on the untrained passages containing trained phrases indicates that some combination of top-down and bottom-up processing may explain the treatment effect. The Phrase passages measure the benefit of practicing specific sequences of words as well as the benefit of practicing difficult words in context (i.e., within the phrases). Moyer’s original hypothesis, that bottom-up and top-down processes interact during practice in
MOR and that words (parts) benefit from practice within the structure of text (whole), is supported by our data on Phrase generalization. Treatments like MOR that include the practicing in context of specific words appropriate to a patient’s deficit may be more likely to generalize.

In order to gain more information about the source of the improvements on the Phrase passages, we analyzed the errors for a subset of the untrained Phrase passages. Using accuracy as the measurement of improvement for PhTA3 and PhTA4 allowed us to analyze what was driving generalization in the Phrase passages in a way that would not be possible with speed data. Comparing accuracy changes on the re-used phrases, PhTA3 showed the greatest improvement on functor words within phrases (25%) as opposed to only 14% on content and unaffixed “other” words. Affixed words coded as “other” showed a slight (5%) decrease in accuracy. Analysis of PhTA4’s errors showed a different pattern. She showed a slight (2-4%) improvement on content, functor and unaffixed “other” words within the phrases, but an 18% increase in accuracy for the affixed words coded as “other”. This analysis suggests that, even within the syntactic context of a phrase, treatment effects are based on improvements to specific word types that are unique to the patient’s particular deficit.

PhTA4 and PhTA1 improved, albeit by different measures, only on the phrase passages. PhTA3 improved on the Phrase as well as the Functor passages and further analysis revealed that improvement on the Phrase passages was due to increased accuracy for the functors within the phrases. For PhTA4, it was the affixed words, a group of words our design did not measure, which appeared to be responsible for the improvement on the Phrase passages. Though we were unable to determine the specific word types that were the source of improvements in speed, it is possible that PhTA1’s improvement on the Phrase passages was due to improved speed of reading affixed words. In future studies, we will be able to include an untrained generalization
passage made up of practiced, affixed words to look into this issue further. We could also include a passage that re-uses not just functors but other words that phonological alexics might be expected to improve on. It is possible that all phonological text alexics would have shown consistent generalization on the Functor passages if we had included inflectionally affixed words in our definition of “functor”.

Three of the four phonological text alexia patients showed improvement on pseudoword reading after treatment, though only PhTA2’s improvement reached significance. These data indicate that MOR may actually improve bottom-up connections between orthography and phonology, and that this improvement can be detected in a task that measures these connections. This is consistent with Tuomainen and Laine’s (1991) hypothesis that if MOR targets the deficit underlying a reading problem and works from the bottom up, improvement in single word reading can be predicted. Pseudoword reading is a reasonable measure of changes to orthography-to-phonology connections in our patients, as most were at or near ceiling on reading of single, real words (see Table 3). PhTA2’s improved orthography-to-phonology connections did not translate to measurable improvements in speed in the more complex task of text reading, though it may have been part of the reason he improved on the functor passages.

Obviously, there were no pseudowords in the Training passages, so PhTA2’s improvement on pseudoword reading cannot be explained by repeated practice of specific words. It seems to be an example of MOR treatment improving the connections themselves. While the other patients seemed to benefit from practicing difficult words in context (Phrase passages), PhTA2 seemed to benefit from practicing the specific words he had difficulty with (reflected by his improvement only on the Functor passages), perhaps improving on functor word reading as his orthography-to-phonology connections improved. Though PhTA2 did not show any other
evidence of improved integrity of these connections, it is possible that he would have with
further weeks of MOR treatment. It may be that MOR can act directly on the deficit itself in
some people and work on specific words that are difficult because of the deficit in others.

The patients in the current study began treatment reading more quickly than most other
patients reported in previous MOR studies. Enrolling these types of participants was the most
effective way for us to examine the two hypotheses to explain MOR’s treatment effects. The
information our data have provided can be applied to future studies investigating how to
optimize MOR treatment for patients of all levels. Our phonological alexia participants did not
demonstrate a part-of-speech effect in single word reading and one of our pure alexia participants
did not demonstrate a length effect. Previous studies have shown that after MOR treatment,
participants beginning at a lower reading rate than ours can demonstrate abolition of length
effects (Beeson et al., 2005) and part-of-speech effects (Beeson & Insalaco, 1998). Our data
may reflect what would have happened if these patients continued to be treated with MOR:
reading speed and/or accuracy gains would get smaller, but reading would continue to improve.

Taken together, the data reported here indicate that the treatment mechanism of MOR for
people with alexia is at least partially due to bottom-up processes. In phonological alexia,
repetition priming of functors and phrasal units containing difficult words led to improved speed
and/or accuracy in untrained text. The exact pattern of improvement varied across patients, but
no phonological alexia patient improved on the Content passages. Pure alexia patients, who do
not have phonological problems leading to difficulty with functors, did show improvement on
the Content passages. However, they did not improve on passages containing few practiced
words (Control passages), indicating that, for pure alexia, strengthening of visual-orthographic
connections improves speed of reading practiced, but not unpracticed, words. Therefore, the
generalization effects associated with MOR appear to be largely driven by improved bottom-up processes that are specific to the type of deficit of the patient.

By evaluating these patients as a group as well as individually, we were able to elucidate patterns of breakdown that cause alexic reading deficits. Moyer reported success with MOR 30 years ago and Speech Pathologists continue to rely on it clinically. There is certainly some evidence that MOR works, and, as a result of the current study, we now have a better understanding of why it works. Our design required all participants to be mildly impaired. However, our participants whose accuracy was too low to measure speed showed improvements in accuracy that were consistent with our hypotheses, suggesting that MOR can benefit more impaired patients. Future case series studies should investigate the use of MOR with other and more severe alexia patients, as the treatment is simple, tends to generalize, and can be useful in informing cognitive models of alexia.
Figure 2.1. The figure depicts a simple cognitive model of reading.
Figure 2.2. Reading speed in syllables per second for phonologic text alexia patients PhTA1 and PhTA2 and pure alexia patients PA1 and PA2. Significant improvements in reading speed were found for the training passages (T) for all patients. Various patterns of improvement were found on the novel generalization passages containing specific words from the T passages (P=Phrase; F=Functor; C=Content words). No significant improvements were found for the control (Ctrl) passages.
Figure 2.3. Reading accuracy for phonologic text alexia patients PhTA3 and PhTA4. Significant improvements in reading accuracy were found for the training passages (T) as well as the generalization passages that contained trained functors (F) and/or phrases (P). Neither patient showed generalization in the content (C) passages or the Control (Ctrl) passages.
Table 2.1. Study Participants

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Education (Years)</th>
<th>Lesion location</th>
<th>Time post-onset</th>
<th>Aphasia diagnosis</th>
<th>Alexia diagnosis</th>
<th>Etiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhTA1</td>
<td>60</td>
<td>18</td>
<td>left temporo-parietal and occipital, right frontal</td>
<td>13 mo.</td>
<td>Anomic</td>
<td></td>
<td>Phonologic Text</td>
</tr>
<tr>
<td>PhTA2</td>
<td>61</td>
<td>18</td>
<td>left posterior frontal, anterior parietal</td>
<td>12 mo.</td>
<td>Resolved Non-fluent</td>
<td>Phonologic Text</td>
<td>Stroke</td>
</tr>
<tr>
<td>PhTA3</td>
<td>69</td>
<td>18</td>
<td>left frontal, posterior temporo-parietal</td>
<td>9.5 yrs.</td>
<td>Anomic</td>
<td></td>
<td>Phonologic Text</td>
</tr>
<tr>
<td>PhTA4</td>
<td>39</td>
<td>12</td>
<td>left insula, temporo-parietal</td>
<td>3 yrs.</td>
<td>Anomic</td>
<td></td>
<td>Phonologic Text</td>
</tr>
<tr>
<td>PA1</td>
<td>53</td>
<td>18</td>
<td>occipital, medial temporal</td>
<td>2 yrs.</td>
<td>None</td>
<td></td>
<td>Pure</td>
</tr>
<tr>
<td>PA2</td>
<td>62</td>
<td>16</td>
<td>right frontal, left temporo-occipital</td>
<td>2 yrs.</td>
<td>None</td>
<td></td>
<td>Pure</td>
</tr>
</tbody>
</table>
Table 2.2. Standardized Language and Cognition Tests

<table>
<thead>
<tr>
<th>TESTS</th>
<th>PhTA1 pre</th>
<th>PhTA2 pre</th>
<th>PhTA3 pre</th>
<th>PhTA4 pre</th>
<th>PA1 pre</th>
<th>PA2 pre</th>
<th>PhTA1 post</th>
<th>PhTA2 post</th>
<th>PhTA3 post</th>
<th>PhTA4 post</th>
<th>PA1 post</th>
<th>PA2 post</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCPM</td>
<td>36</td>
<td>36</td>
<td>21</td>
<td>36</td>
<td>36</td>
<td>35</td>
<td>34</td>
<td>36</td>
<td>33</td>
<td>40</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>TONI (score/percentile)</td>
<td>38/94%</td>
<td>36/79%</td>
<td>26/66%</td>
<td>33/61%</td>
<td>40/90%</td>
<td>35/74%</td>
<td>36/94%</td>
<td>36/79%</td>
<td>36/66%</td>
<td>40/90%</td>
<td>35/74%</td>
<td>35/74%</td>
</tr>
<tr>
<td>BDAE</td>
<td>Basic word discrimination</td>
<td>37</td>
<td>34</td>
<td>34</td>
<td>37</td>
<td>35</td>
<td>35</td>
<td>37</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commands</td>
<td>15</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Ideational Material</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Repetition: Words</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition: Sentences</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>*</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsive naming</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td>18</td>
<td>20</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td>60</td>
<td>26</td>
<td>32</td>
<td>56</td>
<td>20</td>
<td>19</td>
<td>31</td>
<td>58</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special categories</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral word reading</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>26</td>
<td>25</td>
<td>21</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral sentence reading</td>
<td>10</td>
<td>9</td>
<td>2</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral Sentence comp.</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence/paragraph</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RCPM: Raven's Coloured Progressive Matrices; TONI: Test of Nonverbal Intelligence; (Cognitive tests were used as exclusionary criteria and were not tested post-treatment)

*Patient became frustrated and refused to complete more than the first two sentences (which she repeated correctly).

** Due to experimenter error, the BNT results were not valid. However, PA2’s responsive naming, special categories and spontaneous speech showed no indication of anomia.
Table 2.3. Reading test scores, pre- and post-MOR treatment

<table>
<thead>
<tr>
<th>TESTS</th>
<th>Max</th>
<th>PhTA1 pre</th>
<th>PhTA1 post</th>
<th>PhTA2 pre</th>
<th>PhTA2 post</th>
<th>PhTA3 pre</th>
<th>PhTA3 post</th>
<th>PhTA4 pre</th>
<th>PhTA4 post</th>
<th>PA1 pre</th>
<th>PA1 post</th>
<th>PA2 pre</th>
<th>PA2 post</th>
</tr>
</thead>
<tbody>
<tr>
<td>GORT (Pre: Form A; Post: Form B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral (passages 1,3,5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>15</td>
<td>12</td>
<td>11</td>
<td>14</td>
<td>12</td>
<td>7</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>12</td>
<td>13</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Average syllables/s</td>
<td></td>
<td>0.62</td>
<td>0.8</td>
<td>0.94</td>
<td>1.1</td>
<td>0.89</td>
<td>1.01</td>
<td>0.66</td>
<td>1.11</td>
<td>2.87</td>
<td>4.25</td>
<td>1.81</td>
<td>1.97</td>
</tr>
<tr>
<td>Silent (passages 2,4,6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>15</td>
<td>10</td>
<td>12</td>
<td>11</td>
<td>13</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>14</td>
<td>15</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Average syllables/s</td>
<td></td>
<td>0.6</td>
<td>0.95</td>
<td>0.98</td>
<td>1.2</td>
<td>1.38</td>
<td>1.50</td>
<td>2.35</td>
<td>4.01</td>
<td>1.90</td>
<td>4.48</td>
<td>1.83</td>
<td>2.24</td>
</tr>
<tr>
<td>Gates-MacGinitie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silent Reading Comprehension</td>
<td>36</td>
<td>28</td>
<td>27</td>
<td>34</td>
<td>32</td>
<td>18</td>
<td>21</td>
<td>21</td>
<td>18</td>
<td>36</td>
<td>35</td>
<td>36</td>
<td>35</td>
</tr>
<tr>
<td>Pseudoword reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudowords</td>
<td>20</td>
<td>13</td>
<td>13</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Matched real words</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>17</td>
<td>18</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

*Italics indicate a significant difference between pre and post-testing*
Chapter 3. The potential benefit of errors to initial learning and maintenance in anomia rehabilitation

Introduction

Research on the testing and treatment of memory impairment can inform research on aphasia. However, the direct application to aphasia rehabilitation of techniques developed for amnesia has been problematic. Errorless (EL) learning has shown promise in the rehabilitation of memory-impaired populations (Evans et al., 2000) and has recently been investigated as a possible method for the treatment of anomia in people with aphasia (Fillingham et al., 2005, 2006; McKissock et al., 2007). In trial-and-error or errorful (EF) approaches, participants are encouraged to provide a response to stimuli and any errors they make are corrected. In EL learning, a supported environment prevents subjects from producing errors. The theory behind EL learning is that, in patients who lack intact explicit, episodic memory to encode a response as an error, EL learning prevents the association between the error and the target stimulus from being learned implicitly (Baddeley & Wilson, 1994).

The first studies showing an advantage to EL over EF learning involved teaching pigeons to discriminate between red and green discs using the two different methods (Terrace, 1963). Glisky et al. (1986) later applied the concept to learning in people with memory impairments and found that descending cues (or the method of vanishing cues, an EL approach) were more effective than trial-and-error methods (an EF approach). Baddeley & Wilson (1994) renewed interest in the topic when they compared young, elderly and memory-impaired subjects on a test of word-stem completion and found that EL showed an advantage over EF, particularly in the memory-impaired group. Other studies in amnesic populations support the theory that error prevention improves learning (Wilson et al., 1994) and still others have shown a similar effect of
EL techniques with schizophrenic patients (O’Carroll et al., 1999), amnesic patients of varying etiologies including stroke (Evans et al., 2000) and Korsakoff’s Syndrome patients (Komatsu et al., 2000). The advantage of EL learning for stroke-induced anomia has been less apparent.

The Use of EL Learning in Aphasia Treatment

A review of EF and EL treatment studies in anomia rehabilitation since 1985 concluded that both techniques could produce positive results when tested immediately post-treatment and several months later (Fillingham et al., 2003). A study by Fillingham et al. (2006) compared the two treatments using descending (EL) then ascending (EF) cues in 11 patients, treating each for 5 weeks. No clear advantage to EL was found in this study. McKissock et al. (2007) compared repetition (EL) to guessing (EF), dividing EF into two conditions: one with feedback after errors and one without. Results from the EF condition that included feedback did not differ from the EL condition. Patients maintained the effects of both treatments at 12-14 weeks post-treatment. Even in a study in which feedback was not provided at all, there was no clear difference between the two techniques (Fillingham et al., 2005). These authors found no difference between EF and EL in the 4 of their 7 patients who showed a treatment effect, even though they were not given any feedback. In fact, one patient showed a treatment effect only in EF, a condition in which he was never given the correct answer during treatment. Therefore, despite great interest in applying the use of EL learning techniques to the aphasic population, there is little evidence to suggest it is a superior treatment strategy.

Memory Impairment in Aphasics vs. Amnesics
Baddeley and Wilson’s (1994) theory for the success of EL in amnesic populations is based on the assumption that explicit memory is impaired to the degree that patients must rely on the only intact memory system available to them: implicit memory. These patients’ particular problem with error production, then, is their inability to use explicit, episodic memory to distinguish between their own production of an error and a therapist or experimenter’s feedback providing the correct response. The authors suggest that, if errors are prevented during rehabilitation sessions, the Hebbian learning capacity of implicit memory can best be utilized.

One explanation for the equivocal findings on EL learning for stroke-induced anomia is that the majority of the aphasic population does not suffer from the same degree of memory impairment as the amnesic population and can rely on explicit memory to some extent. Moreover, the early literature showing an advantage to EL over EF learning for memory-impaired populations included mostly short-term studies of new learning (i.e. arbitrary paired associates for experimental purposes or learning how to use a new device). We know of only two studies in the amnesic population showing re-learning (previously intact face-name associations). The studies did not compare EF to EL methods, but did successfully treat AD patients using an EL paradigm (Clare et al., 1999; 2002). Most other studies were completed in a single testing session with a short delay period to test retention of new learning in populations whose chief complaint was memory deficits (e.g., Baddeley & Wilson, 1994; Hunkin et al., 1998). In aphasia, the chief complaint is language deficits and the goal is to re-learn previously intact associations (e.g., picture names) that were once automatic. Treatments are designed to be administered over several weeks, not in a single session, and ideally, the re-learned associations are retained. Experimental findings for people with aphasia vs. amnesia are not easily
interchangeable due to these fundamental differences between the goals and circumstances of treatment for the two populations.

The Role of Effort in the EF/EL Question

A specific type of EL paradigm that has been successful with amnesic populations is the method known as spaced retrieval (SR) (Camp et al., 1996). In SR, patients are given the answer to an item and are then asked to produce it again without assistance, first at a short interval and then at increasingly longer intervals. If the patient makes an error, the therapist goes back to the shorter interval and gradually tries to increase it again. Intervals can be measured by how much time passes between attempts or by how many other items are practiced in between (Balota et al., 2007). One reason that SR EL paradigms may be effective is that they take into account the variable of effort, which has been largely ignored in studies comparing EF and EL treatments for aphasia. Tailby & Haslam (2003) studied EL learning in memory-impaired patients using a condition designed to require minimal effort (writing a word provided by the experimenter), and one in which more effort was required (writing a word after being provided with the initial phoneme and a definition of the word). The authors found that the more effortful EL condition resulted in better learning than the less effortful EL condition. Evans et al. (2000) found that when memory-impaired patients used an imagery strategy designed to provide deeper encoding of stimuli and better engagement in the task, their free recall of arbitrary face-name associations improved using an EL method that had not been successful for free recall without the imagery strategy. As Clare & Jones (2008) point out, even the original studies of EF and EL learning in pigeons (Terrace, 1963), were conducted with highly motivated animals (they were food-deprived) and it was still quite difficult to train them to make discriminations in an EL paradigm.
The SR treatment technique has also been borrowed for aphasia treatment from the memory impairment literature. In a single case study conducted in our lab, we compared EF to EL in an anomic patient using ascending and descending phonological cues for picture naming (Lacey et al., 2004). The patient was able to learn items to a pre-set criterion in the EF condition, but not in the EL condition. We hypothesized that the patient was not encoding items in the EL condition because the design required only that he repeat the word multiple times; there was no incentive to listen carefully to the descending cues. We therefore created a new design, based on the SR technique, which would make the EL treatment more effortful. Instead of administering all cues for an item in succession, we alternated between two treatment items so that a cue for a given item was not provided immediately after the preceding cue. This design required more effort from the patient in the EL condition because, after receiving a cue and producing item 1, he had to focus on item 2, then return to item 1 and produce the complete name after being given a diminished cue. When the SR manipulation was added to the treatment, the patient reached criterion in the EL and EF conditions. Fridriksson et al. (2005) compared the errorless SR technique to ascending cues (EF) in three people with aphasia. The study reports an advantage to the SR paradigm, though the results were collapsed across the three patients and over multiple maintenance test sessions in order to show this advantage. The study is clinically relevant because the patients chose treatment items that they thought would be useful to them, but these stimuli pose a problem for experimental control. Our own results for learning items that were matched across sets for frequency and number of syllables indicate that an SR EL paradigm is more effective than simple descending cues without any manipulation to control for effort (Lacey et al., 2004). However, the question of whether an SR EL paradigm is more effective than EF has not yet been addressed.
A Note on EL terminology

It has been suggested that different terms should be used for techniques that are truly errorless and those that are error-reducing. We do not distinguish between these two terms. We maintain that a long-term treatment program for aphasia that is truly errorless could only exist as a theoretical construct. Perhaps in a patient without any phonological processing problems or attentional deficits, a single treatment session could be truly errorless, but it is unrealistic to propose a treatment method that requires an aphasia patient to complete a multi-week program without ever producing an error. Moreover, it has been suggested that, even if a treatment could be administered in such a way that no error is ever produced orally, it is possible that a patient considers multiple erroneous responses before finally choosing the correct oral response (Clare & Jones, 2008). It is impossible to know if this happens or whether internally-generated errors could also be paired with the picture stimulus. Therefore, we use the term errorless (EL) here to mean that we designed the treatment to be as errorless as possible and that in practice it greatly reduces the number of errors produced in comparison to an errorful (EF) treatment (See Table 3).

The Current Study

To maximize the likelihood of uncovering an effect of condition, we treated a group of patients in an SR EL and an EF treatment until they reached either a pre-determined criterion or a plateau (Experiment 2). Both treatments continued until they were either shown to be effective or shown not to be, allowing us to ask which one would emerge as superior and at what point. All patients were treated in both conditions in every session and the order in which treatments were administered in the session was counterbalanced. It was also our goal to determine if
statistical differences exist between EF and EL in terms of how treatment effects are maintained in the long-term. Therefore, we tested maintenance of both treatment effects at longer post-treatment intervals than has been done previously.

Another major question we examine here is the degree to which explicit memory impairment might interact with performance on the two treatment paradigms initially and at maintenance testing (Experiment 1). Previous studies have administered standardized explicit memory tests and correlated these with treatment performance. In Experiment 1, we report the results for control subjects and people with aphasia on our own test of explicit verbal recognition memory using auditory word stimuli. We developed the test in order to assess explicit memory in the same domain in which it would be used during treatment. Because the theoretical motivation of EL treatments is rooted in studies of memory-impairment, it is important to determine the role that explicit memory plays in EF and EL treatment outcomes.

**Experimental Design and Methods – Experiment 1**

**Experiment 1 – Verbal Recognition Memory Testing**

**Participants.**

*Patients.* Seven persons with aphasia participated. They were recruited through local hospitals and stroke support groups and gave informed consent prior to beginning testing. The protocols were approved by the Georgetown University Institutional Review Board and the research was conducted in compliance with the Helsinki Declaration. They were at least one year post-stroke, were able to follow basic commands and instructions and did not show any signs of dementia. (See Table 1 for demographic information)
Control Participants. Fifty-five normal subjects over the age of 40 participated in the norming of our newly-developed explicit verbal recognition memory task and the assessment of name agreement for our treatment pictures. They ranged in age from 40 to 83 and had education levels from 12 to 20 years. In order to be included in the study, controls had to receive a score of 27 or above (out of 30) on the Mini Mental State Exam (Folstein et al., 1975) and they had to pass a hearing test.

Stimuli. Eight words were chosen from each of ten semantic categories, five natural and five man-made, for a total of 80 words. All words were then recorded by the same speaker. Participants listened to the recorded words through circumaural headphones and made responses on a keyboard containing only the keys they would need for the experiment. The words used were mid to lower dominance nouns, as defined by rankings between 5 and 27 on Battig and Montague’s (1969) ratings of category membership. That is, the words were not among the nouns most frequently named as being typical members of a given category.

Procedure. In order to assess explicit memory, we designed a task that could be administered in a single session and that would assess explicit recognition memory. The test used recorded, auditory word stimuli presented through headphones, as this was what would be used in treatment. The task consisted of a study phase for 40 words and, after a 20-minute delay, a test phase in which the studied words were presented again with 40 unstudied distracter words from the same semantic categories as the studied words. The studied or unstudied words were counterbalanced such that half of the participants studied one set of 40 words while the other half of the participants studied the other half.
**Study Phase.** Subjects rated a list of 40 auditorily presented words on either concreteness or frequency using a scale with five gradations, indicated on the keyboard by increasingly larger circles. They were told before the task began that they would be asked to recall these words later. For deeper encoding of the studied stimuli, subjects then rated the same 40 words again on whichever scale, concreteness or frequency, they had not used in the first rating task. The procedure for the concreteness rating study task was modeled after Paivio and Yuille (1968) but with a five rather than a seven point scale, and with simplified language for use with patients. Examples were given in both tasks and responses were noted by the examiner to be sure the participants understood the task.

Following this task, the participants were asked to relax and watch a video for 20 minutes. The video was a slapstick comedy, chosen because of its low verbal content to minimize interference with memory for the studied words.

**Test Phase.** Participants were then tested on their ability to remember which words they had rated by answering the question “did you hear this before?” for all 80 words (40 studied, 40 unstudied). The task was simply to press the yes or no key to indicate whether the word had been presented previously.

The following variables were counterbalanced across subjects: order of rating tasks, order of presentation of items in the study phase and test phase, items for which “yes” is the correct response and items for which “no” is the correct response, the positioning of yes and no keys on the keypad, and the lists of what was studied and what was unstudied.
Results – Experiment 1

Control Participants

Controls performed well but not at ceiling on our test of explicit memory. The average score on the test of 80 questions was 91.9% (.04%) correct. Accuracy ranged from 81.0% to 97.5%. Subjects were divided into three age groups: 40-55, 56-68, and 69-83. Mean accuracy for the 40-55 group was 93.5% (.03%) (Range: 86.3%-97.5%), for the 56-68 group 91.0% (.05%) (Range: 81.3%-97.5%) and for the 69-83 group 89.8% (.05%) (Range: 81.0%-96.3%). Control participants were also divided into five different education groups: high school degree (12 years), some college or education past high school (14 years), college degree (16 years), master’s degree (18 years) and PhD or professional degree (20 years). There was a main effect of age on performance (F(2,53)=3.78; p= .029), but education level did not affect performance (F(4,51) = 0.53; p>.1).

Patients

All patients were impaired on the explicit memory task compared to controls (i.e., more than 2 standard deviations below the mean for their age group). They ranged from more than 2 standard deviations (SD) to more than 11 SD below their age-group mean (See Table 3).

Experimental Design and Methods – Experiment 2

Experiment 2 – Anomia Treatment

Participants

The same control participants and 7 persons with aphasia participated in Experiments 1 and 2. The only task of the control participants was to name our set of pictures to determine
name agreement. 132 pictures were above 83% name agreement and were therefore determined to be suitable for use as treatment items for the patients.

The participants with aphasia were tested on multiple language and cognitive tests before beginning treatment (see Table 1 and test descriptions below). We developed criteria for placing patients into categories of phonological, semantic and mixed anomia. A diagnosis of phonological anomia was given to those patients who were impaired on our pseudoword repetition task but within normal range on our word-to-picture matching tasks and on the Pyramids and Palm Trees (Howard & Patterson, 1992). Semantic anomia diagnosis was based on intact pseudoword repetition scores, but below-normal scores for one of our two semantic processing tests: word-to-picture matching and the Pyramids and Palm Trees test. Patients who were impaired on both types of tests were given a diagnosis of mixed anomia.

**Stimuli**

**Picture Stimuli.** Picturable nouns with two or three syllables and five to seven phonemes were chosen as treatment stimuli. Black and white line drawings were then gathered from various sources on the internet as well as Snodgrass and Vanderwart (1980). Pictures used in treatment as well as the 2\textsuperscript{nd} exemplars used to assess generalization were normed for name agreement by the same control subjects from Experiment 1. Pictures were included in the study only if at least 83% of controls produced the exact same naming response to the picture.

**Phonological Cues.** The name of each item was divided into three phonological cues. In order to match the three cues across items, it was necessary to control the number of phonemes and choose words that could be split into three cues that did not require the use of a schwa.
When splitting the words into cues, we retained as many phonemes as possible in the third (longest) cue and, where possible, retained two syllables in the third cue for three syllable words. We did this so that the first fade after repetition in the EL condition would contain as much information as possible, providing support and reducing the chance for error (See Figure 1). The cues to be used in treatment were then recorded so that there would be no difference in how the cues were presented in each condition and on each treatment day. The recorded voice was a female speaker and each cue was recorded separately, rather than recording the whole word and then splicing it into three recordings. This allowed us to keep the cues sounding natural while at the same time ensuring that they did not change from day to day or between conditions or patients.

We chose to use phonological, ascending (EF) and descending (EL) cues in our study as a way to repair the link between phonology and semantic representations. There is some debate as to whether treatment cues should be at the level of each patient’s deficit (i.e., semantic cues for patients with damage to the semantic system; Hillis, 1989; Wambaugh, 2003). However, most models of word retrieval include interactive components between levels of representation (phonologic, semantic), indicating that phonological cues could benefit patients with deficits at any level (Hillis 1993; 1998). In recruiting subjects, it has been our experience that most fall under the category of mixed anomia and so could be expected to benefit from either or both types of cues. Therefore, we report testing for semantic and phonologic deficits in our participants, but because our main question of interest focuses on the EF/EL dichotomy, all patients were treated the same way, with ascending and descending phonologic cues given in a spaced retrieval design (explained below).
Procedure – Pre-treatment Testing

**Standardized Pre-testing.** Patients completed the Boston Diagnostic Aphasia Examination (BDAE; Goodglass, Kaplan & Barresi, 2001) as well as other standardized language and cognitive skill tests before beginning treatment.

**Modified Wisconsin Card Sorting Test (MWCST).** Executive function was assessed using the MWCST (Lineweaver et al., 1976), a version of the Wisconsin Card Sorting Test (WCST) designed to be used with aphasic patients. As in the WCST, patients are asked to sort cards by attributes such as color, shape and number. After six consecutive correct responses in sorting the cards, the patient is told to "find another rule" (i.e., sort by a different attribute). This procedure continues until the patient gets through six categories or until all 48 cards are administered. Perseverative errors were the dependent measure we were interested in and we compared our patients to normative data based on age and education (Lineweaver et al., 1976).

**The Biber Figure Learning Test (BFLT).** The BFLT (Glosser et al., 1989) was also administered to each patient in order to assess learning, recognition memory, and recall in a non-verbal domain. In this test, 10 novel, geometric shapes are presented serially for 5 learning trials. Drawings are scored using a scale of 1-3 points for each drawing so that the maximum score for each learning trial is 30 points. After a 20 minute delay, free recall of the shapes is assessed using the same scoring system. Delayed recognition memory is also tested at the same time-point, by asking the patient to pick out each studied drawing from among 3 distracters.

**The Pyramids and Palm Trees Test (Howard & Patterson, 1992).** In order to assess semantic processing, the Pyramids and Palm Trees Test was administered to each patient. In this test, the participant is asked to look at a picture and then choose which of two other pictures best
“goes with” the first picture. No spoken language is used during the test. The test contains 52 items and normal range is considered 49 and above.

**Study-Specific Pre-testing.** In addition to the standardized testing, participants completed several other tests developed in our laboratory.

**Pseudoword repetition test.** The pseudo-word repetition task was developed in our laboratory and consists of 30 pseudowords ranging from one to five syllables. The patient is asked to repeat each word after the experimenter. Normal performance on this task is above 26/30 correct (unpublished normative data from our laboratory).

**Word-to-Picture Matching.** In the word-to-picture matching test, participants are presented with 6 semantically related pictures on a screen and asked to point to the item spoken by the experimenter. The test contains 48 items and normal range is considered 46 and above (Rogers & Friedman, 2008).

**Repetition Test.** We created a single-word repetition task comprised of words matched for number of syllables to all of the words from which the treatment stimuli would be selected. It was necessary for repetition to be relatively intact to ensure that the EL condition, which begins with whole word repetition, was as errorless as possible. In order to keep errors to a minimum, but not exclude all patients who had difficulty with repetition (a common deficit in the aphasic population), cut-off was set at 80% correct. The test was administered through headphones. If patients could not meet the cut-off when using the headphones, they were required to repeat accurately under conditions that would be used in treatment, i.e., repeating
with the picture present and, in two cases, with the cues presented live. Patients VZN, OJN, XTC and DBR achieved scores higher than 80% on the repetition test using headphones. Patient QPR repeated only 40% of the items correctly when using the headphones. However, with the picture present, her repetition was 100% correct. Patients JLC and YLC were unable to complete the repetition test with the headphones. However, with the picture present using live repetition, they scored 80% and 100%, respectively.

Procedure - Treatment

**Overview.** Treatment consisted of baseline testing to determine suitable treatment items followed by treatment 2 times per week until criterion or plateau was reached. We minimized administration of probe tests because they create an errorful situation. They were given only once every 4 treatment sessions. This was followed by maintenance testing at least 1 month and at least 6 months after the end of treatment. In order to assess generalization, patients were also tested before and after treatment on naming of a set of different exemplars of each treatment item (See Table 4).

1. *Baseline Picture naming.* Patients were asked to name the set of 132 pictures determined to have high name agreement among control subjects (see Stimuli). Treatment items were chosen individually for each participant. Sixty items that participants missed at least 3/4 times on baseline testing were randomly assigned to the EF, EL and untrained conditions. Each set was matched for initial phoneme, semantic category, number of syllables and frequency. For QPR and OJN, the BNT was used as the untrained word list, because these patients missed fewer than 60 words consistently across baselines.
2. Second exemplar testing. In the fourth baseline session, patients were given a set of second exemplars of the same 132 items. Only items whose second exemplars were named incorrectly were used in treatment, because items were tested again after treatment to assess generalization to a novel picture.

3. Pre-treatment probe. The 40 treatment words were tested once more before the first treatment session to ensure a stable baseline. Subsequent probes to assess treatment progress occurred only once every 4 treatment sessions to minimize the errorful situation of the tests.

4. Treatment. The method of ascending and descending cues (Linebaugh & Lehner, 1977) was used for the EL (descending cues) and EF (ascending cues) treatments. Treatment sessions were one hour, twice a week. The EF and EL conditions alternated within a session and were counterbalanced across sessions so that each patient practiced the treatment words twice per session and never began with the same condition in two consecutive sessions.

A descending cueing hierarchy should be more engaging for the patient than simple repetition, though it does allow for the possibility of errors after the initial cue is faded. Based on previous research showing that making a descending cue paradigm more effortful improves initial learning (Lacey et al., 2004), we incorporated a spaced retrieval (SR) design into our treatment paradigm as well. We alternated between two treatment items so that each cue was not given immediately after the preceding cue. This required patients to make some effort to retrieve the item after being forced to attend to the alternate item in between. Though the inclusion of sufficient effort is more an issue for the EL condition, we made this manipulation in both
conditions in order to keep them parallel (see examples of each condition below for clarification).

Feedback was always provided in both conditions. On the rare occasion an error was made in EL, the experimenter asked the patient to repeat the complete word, then ended the trial so as to minimize any potential association of the error with the picture stimulus.

*Treatment Example – Errorful (insert figure 1 about here).* Figure 1a shows an example of an EF trial with the words “cactus” and “volcano” in which the participant required the full set of cues for cactus but was able to repeat it successfully and required 3 cues before producing volcano successfully.

*Treatment Example – Errorless.* Figure 1b shows an example of an EL trial with the words “banana” and “envelope” in which the participant did not make any errors during the trial. If the participant had produced an error on banana, for example, the SLP would have spoken the word “banana”, asked the subject to repeat it and then continued treatment only for “envelope”.

5. **Hitting Criterion or Plateau.** Criterion was 90% correct for two consecutive probe tests. A plateau was defined as less than 5% improvement over the previous 6 probe tests. Because of the infrequent probes, the plateau represents 12 weeks (24 sessions) of treatment with less than 5% improvement. We set this conservative plateau because our long-term treatment was designed to determine any possible differences between EF and EL and we wanted to ensure that both treatments were given sufficient time to elicit any possible change in performance.

Because probe tests occurred once every four sessions, it was sometimes necessary to continue treatment in one condition when the other condition reached criterion or plateau first. In this case, we continued to alternate between EF and EL treatments to keep the structure of the
sessions the same, but used “filler” words (words the patient was able to name at baseline) instead of treatment words, in essence allowing a sham treatment to fill in for the condition on which criterion or plateau had already been reached.

6. Maintenance Testing. Patients were tested for maintenance at 2 time-points after criterion or plateau was reached in each condition. Maintenance was tested in all patients, regardless of treatment outcome. Maintenance time-point one was at least 1 month post-treatment for each patient and time-point 2 was at least 6 months post-treatment. Due to patient schedules, it was often difficult to test them at exactly 1 and 6 months post-treatment, but we always erred on the side of testing further from treatment. Because the treatments differed in terms of when patients hit criterion or plateau, this measurement often had to be taken for each condition on a different day. 2nd exemplar generalization pictures were also re-administered at both maintenance time-points.

Results – Experiment 2

Results - Standardized testing

Table 1 shows results for the language and cognitive tests. Scores for participant QPR indicate that the anomic deficit is at the level of phonology according to our diagnostic criteria. VZN’s score on the Pyramids and Palm Trees test indicates semantic anomia, though her Word-to-Picture matching score was within normal range. All other patients were semantically and phonologically impaired (mixed anomia). Most patients also showed signs of executive functioning deficits on the MWCST.
Results - Treatment

Table 2 summarizes the results for pre-treatment, post-treatment and maintenance probe tests, number of sessions to criterion or plateau as well as the memory test results for each patient. Figure 1 shows pre-treatment, post-treatment, and maintenance probes for all patients broken down by EF and EL. Response to treatment followed two different patterns, though they were not related to the EF and EL conditions. Patients either reached criterion (two consecutive probes at 90% accuracy or above), or they came to a plateau and treatment was discontinued. Patients QPR and VZN reached criterion in both the EF and EL conditions. Patient OJN reached criterion only in EF, but plateaued high in EL at 16/20 (80%). All other patients came to a plateau in both conditions, albeit at different time-points for EF vs. EL (see Figure 2 and Table 2 sessions to plateau).

Patients DBR, XTC, and YLC came to a plateau in treatment, but achieved learning significantly above the pre-treatment probe in at least one of the two conditions. Initial learning significantly above baseline was demonstrated in twice as many patients in EF than in EL and it also maintained, even at the longest time interval. DBR maintained for time-point 1 (binomial test: \( p < .001 \), two-tailed) and 2 (binomial test: \( p < .05 \), two-tailed) in EF but not in EL (Binomial test: \( p > .1 \), two-tailed). XTC also maintained for time-point 1 and 2 (Binomial test: \( p < .01 \), two-tailed, for both time-points) in EF, but not in EL (Binomial test: \( p > .1 \), two-tailed). YLC maintained at time-point 1 in EF (Binomial test: \( p < .05 \), two-tailed) and EL (Binomial test: \( p < .05 \), two-tailed), but after 6 months EF was still significantly above pre-treatment levels (Binomial test: \( p < .05 \), two-tailed) while EL was not (Binomial test: \( p > .1 \), two-tailed). The only significant difference between EF and EL was for DBR on her final treatment probe (EF=14, EL=5; \( \chi^2 = 4.9, p < .05 \)).
Results - Maintenance

Table 2 shows maintenance scores for each patient. QPR maintained significantly above her pre-treatment probe score at maintenance time-points 1 and 2 (Binomial test: $p<.001$, two-tailed for both time-points) in EF as well as EL (Binomial test: $p<.001$, two-tailed, for both time-points). There were no differences between EF and EL. VZN also maintained above pre-treatment levels at maintenance time-point 1 (Binomial test: $p<.001$, two-tailed) and 2 (Binomial test: $p<.01$, two-tailed) in EF as well as time-point 1 (Binomial test: $p<.05$, two-tailed) and 2 (Binomial test: $p<.001$, two-tailed) in EL. There were no significant differences between EF and EL. OJN maintained his EF treatment words significantly above pre-treatment levels at time-points 1 and 2 (Binomial test: $p<.05$, two-tailed, for both time-points), but not his EL words (Binomial test: $p > .1$, two-tailed).

Results - Untrained items

For six of the seven patients, at the final treatment probe, accuracy on the EF items was significantly different from accuracy on the untrained items. For four of the patients, at the final treatment probe, accuracy on the EL items was significantly different from accuracy on the untrained items (QPR ($\chi^2 = p<.001$ for both conditions), VZN (EF $p<.01$; EL $p<.001$), OJN (EF $p<.01$; EL $p<.05$), DBR (EF only $p<.01$; EL $p>.10$ NS), XTC (EF only $p<.05$; EL $p>.10$ NS), YLC (EF $p<.05$; EL $p<.02$), JLC (no treatment effects)).

Results - Sessions to criterion/plateau
Table 2 lists the number of treatment sessions required for each patient in each condition to reach criterion or plateau. Of the 3 patients who reached criterion, 1 reached it in both conditions simultaneously, one reached it in EF first and the third reached it only in EF. 3 of the 4 patients who did not reach criterion, reached plateau (i.e. stopped improving) sooner in the EL than the EF condition.

**Results – Errors**

In Table 3 a tally of the errors made during all treatment sessions shows that errors were far less frequent in the EL condition, especially after the initial treatment sessions. Each cell in the table shows the total number of errors produced over four treatment sessions, because this was the number of sessions after which each probe test occurred.

**Results – Generalization**

Table 4 shows the results for the 2nd exemplar generalization items compared to results for the treated items. Accuracy on the 2nd exemplars followed the same pattern as the treated items, indicating generalization to untrained pictures of the same concept.

**Discussion**

**Summary and response to past literature**

We found no advantage to EL learning in patients with varying degrees of explicit memory impairment after a large number of treatment sessions, despite our manipulations to make the EL condition more effortful. Six of the seven patients displayed learning significantly greater than their baseline probe and their untrained word probe in both conditions or in the EF
condition only. At maintenance time-point 1, performance was significantly above baseline for six patients in EF and 3 patients in EL. At maintenance time-point 2, at least 6 months post-treatment, performance was significantly above baseline for six patients in EF and 2 patients in EL. Contrary to what theories of EL learning would predict, the patients who had lower verbal explicit memory scores did not show greater learning or maintenance when they were prevented from making errors. Two of the three patients who reached criterion did so in fewer sessions of EF treatment than EL. Three of the four who did not reach criterion took more sessions to reach a plateau in EF than in EL, indicating that learning was still occurring in EF after it ceased to be measurable in EL. Essentially, there were 2 patterns of treatment effect: those who reached criterion and those who did not. There was no clear effect of the EF/EL variable on either group, but there were some signs that EF treatment could be more effective.

We tested performance in several ways in an attempt to reveal differences between the two variables if they exist. The EF/EL variable did not differ in terms of generalization to a second exemplar; performance on the second exemplars generally followed the same pattern as the treatment items (see Table 4). Previous studies have suggested that executive function and self-monitoring are the most important variables in anomia treatment outcome (Fillingham et al. 2005, 2006). Our measure of executive function, number of perseverations on the MWCST, followed a similar pattern to our memory measures in that patients who were more impaired did not benefit more from EL learning than those who were less impaired (see Table 1).

The current study confirms and extends the findings of Fillingham et al. (2006). Indeed, data for their short-term study and our long-term study are similar in several respects. Out of their 11 patients, 5 maintained in both conditions and 4 maintained only the EF items (2 of these 4 maintained EF items significantly better than EL items). The two patients in our study who
reached criterion in both conditions also maintained the treatment effect significantly above baseline after at least 6 months. OJN, DBR, XTC and YLC maintained only the EF treatment effect after at least 6 months (see Table 2). Similar to our findings, others have found no support for a clear advantage to EL (Fillingham et al., 2006; Conroy et al., 2009). However, they suggest that, since the two treatments do not seem to differ, EL may be the better option because their patients preferred it to EF. However, we have found, as have others (Abel et al., 2005) that some evidence points to EF being more effective.

The Fillingham patients’ self-reported preference for EL may have been influenced by the design. The study did not counterbalance the order in which patients were given the two treatments. Patients might have been more interested and engaged in whatever treatment came first (which was EL, in all 11 cases) because of greater arousal due to novelty. Contrary to what Fillingham et al. (2006) reported, our patients did not find EF frustrating. YLC, perhaps our most severely impaired patient, seemed to enjoy the challenge of trying again when he made an error during EF treatment. Discrepancies in patient preferences could be related to anything from personalities to protocols. Regardless of the preferences of any individual patient, it is clear that EL is not unequivocally superior to EF treatment for anomia. The EF/EL question remains unresolved.

The Role of Effort

Our paradigm did not use a literal definition of EL learning in that we did not expect to prevent 100% of errors, but we argue that our paradigm tested a more ecologically valid definition of EL learning that takes into account the need for the patient to be engaged in the task. We previously found that a traditional descending cue paradigm could be more effective if
we made it more challenging (Lacey et al., 2004). Based on these findings, we added the SR manipulation to the current study as well. Our SR manipulation required patients to switch between two items. This might be expected to cause problems for perseverative patients, but the patients who made more perseverative errors on the MWCST did not have difficulty with it. The SR manipulation did not cause the EL condition to become errorful. However, making the EL condition more effortful did not make it more effective than EF.

The potential confound of effort is intertwined with the problem of controlling errors, in that 100% error elimination can also mean elimination of effort, lack of attention and subsequent lack of encoding (Robertson & Murre, 1999). The problem of controlling for effort in the comparison of these two paradigms seems almost impossible to overcome. A change that makes a paradigm more effortful usually carries with it a greater possibility for errors, while a change that lessens the likelihood of errors almost by definition decreases the effort required to complete the task. The EF/EL question in anomia treatment may remain unresolved for this reason. It may be more beneficial to investigate the effects of level of engagement in therapy and to determine individualized treatments that are sufficiently challenging and minimally frustrating for aphasia patients.

**The Role of Explicit Memory**

The potential benefits of EL learning may only be realized in populations with more severe memory impairments than the patients presented here. This raises the question: what level of impairment would be severe enough to show a difference? The current study includes stroke patients at a variety of severity levels in terms of language as well as memory function, yet even the most severe patients did not show an advantage to EL learning.
The patients who did not reach criterion in treatment also scored lower on our verbal recognition memory test as well as on the final learning trial and delayed recall of the BFLT (which requires recall of the shapes under unsupported conditions, similar to a naming probe for treatment pictures). However, they achieved normal performance on the spatial recognition memory portion of the BFLT (see Table 2), indicating that certain memory systems are still intact and potentially available to help them to handle error production in a way that amnesics would not be able to. Glosser et al. (2002) tested a group of Alzheimer’s Disease (AD) patients on the BFLT and found that they performed at about 20% accuracy on the final learning trial, similar to the performance of our group of patients who did not reach criterion. However, the average delayed recognition memory score for the AD patients was 62%, whereas our group of patients who did not reach criterion had an average recognition memory score of 90%. Spatial recognition memory, which is thought to be dependent on hippocampal integrity (Broadbent et al., 2007), seems to be intact in our patients. AD patients and others who can be shown to have explicit memory deficits, even when explicit memory is tested under supported conditions (i.e. recognition memory), may be a more appropriate population for EL learning techniques. Aphasia patients who show signs of intact explicit memory function, even if it is only on spatial memory tests, may benefit from treatments that allow for some errors to be produced.

Because of the way we designed our treatments, the EL condition actually provided more repetitions of the treatment word. However, despite the extra repetitions in EL that our design allowed, we found no advantage to the EL condition. We chose to let production of an error dictate the end of a trial in the EL condition and production of the correct name dictate the end of a trial in the EF condition. Previous studies have used a similar design to ours (Fillingham et al., 2006) while others have chosen to equate their EF and EL conditions so that patients were given
the same number of repetitions of the item in both conditions (Fillingham et al., 2005). Neither design showed a clear advantage to EL. Moreover, the patient who had the most errors in the EL condition in our study, QPR, also had the most success in treatment and treatment maintenance. We originally hypothesized that errors paired with the picture stimulus during initial learning would have a detrimental effect on the maintenance data, but this theory was not borne out. The motivation behind EL treatments is supported by Hebbian learning theory and by multiple studies with amnesics, but it does not seem to translate to anomia treatment for stroke patients.

**Is EF Learning More Appropriate Than EL Learning for Persons with Aphasia?**

It is well-known that attentional engagement is extremely important for memory encoding and subsequent retrieval. Testing in control subjects has been shown to improve recall (Tulving, 1967), which may be partly due to a testing situation being more attentionally engaging. Forced guessing is associated with production of more errors but not associated with decreased accuracy on subsequent tests (Pashler et al., 2007). A recent study in which control subjects were asked to answer questions about fictional facts shows that testing on which subjects fail completely (i.e. 100% errorful) is associated with improved retention compared to studying of the same fictional material without being tested (Kornell et al., 2009). The effect of testing on patient populations is more complicated. The current study indicates that only in populations with severe, multi-domain explicit memory impairments might the benefits of error reduction in EL learning override the benefits of superior attentional engagement and deeper encoding inherent in EF learning. Moreover, the current study indicates that these benefits in EF are associated with better maintenance of treatment for patients at all levels of severity.
The literature that initially established an advantage to EL learning in humans (e.g., Wilson et al., 1994) does not directly relate to the type of learning that is important in aphasia rehabilitation. Rather than testing retention of re-learned, previously intact associates over several weeks, as occurs in anomia rehabilitation, many EL learning studies from the amnesia literature assessed new learning of arbitrary paired-associates in a single session. However, one study that tested learning of loosely-related paired associates (e.g., card-spade) (Squires et al., 1997) in people with memory impairments can provide some insight into the EF/EL question in people with aphasia. These authors report an advantage to EL for recalling the related pairs at immediate test, but no difference between EF and EL after a delay. Analysis of the data showed that there were several items that participants could not recall at immediate test but were able to recall after a delay. This phenomenon occurred much more frequently in EF than in EL. The authors interpret this result to be due to interference from errors produced during learning that interrupt production of the correct response at immediate test. At delayed test, the interference has resolved and patients are able to show recall of these items. If EL learning is associated with successful immediate recall in associated pairs while EF learning is associated with successful delayed recall after a resolution of interference, it could help to explain the possible advantage to EF that we observed over our long-term training.

A more recent study in a control population (Kornell et al., 2009) revisited the issue of errors produced when learning weakly associated pairs (e.g., pond-frog) in an EF (guessing the associate) compared to an EL (reading of the two associates) paradigm. They tested their subjects after a delay of approximately 38 hours and found that retention of the items learned in the EF paradigm was superior to retention for those learned in EL. When subjects were tested on arbitrary material, EF learning was equal to EL, but when the two associates were related, EF
learning was superior. The authors hypothesized that the associated pairs were more similar to real-life learning situations in which several potential answers are considered before choosing a final answer, a process which most likely activates the semantic network of the correct answer and may be the mechanism for improved retention. A similar phenomenon may be at work in EF treatments for anomia. By providing the picture first without its associated phonology, EF treatments may activate the semantic network of that item, thereby creating deeper encoding and improving retrieval at maintenance testing. The patients in the current study may have benefited from the greater effort required and the potentially deeper encoding in the EF condition and, over the long-term, any interference effects their errors may have caused had time to resolve. More generally, as Clare & Jones (2008) suggested, EF learning may train patients to deal with errors so that when an errorful response later comes to mind during testing they are better equipped to suppress it.

**Conclusion**

Anomia rehabilitation concerns the re-learning of information that the patient once could produce without difficulty. Although many persons with aphasia also suffer from memory impairment, they may not be impaired to a degree that makes EL learning approaches necessary. Mounting evidence shows that attentional engagement is as important as error-management in therapy, if not more so. Attention may be particularly difficult to manage with the kind of familiar material that aphasia patients must focus on in treatment. Approaches that allow errors to occur, if they are more engaging for the patient, may actually be more effective in the rehabilitation of naming.
Figure 3.1a. Errorful Treatment Example
Figure 3.1. The figure shows an example of a patient’s responses during an errorful and an errorless trial.
Figure 3.2 Naming accuracy on test probes for all patients at multiple time-points

EL learning and maintenance: All patients

![Bar chart showing naming accuracy for EL learning and maintenance.]

EF learning and maintenance: All patients

![Bar chart showing naming accuracy for EF learning and maintenance.]

Figure 3.2. The figure shows the accuracy of all patients pre-treatment, post-treatment and at two maintenance time-points at least 1 and 6 months post-treatment. Starred bars indicate significant difference from pre-treatment probe.
Table 3.1 Descriptions and test scores of study participants with aphasia

<table>
<thead>
<tr>
<th>Background information</th>
<th>QPR</th>
<th>VZN</th>
<th>OJN</th>
<th>DBR</th>
<th>XTC</th>
<th>YLC</th>
<th>JLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Age</td>
<td>39</td>
<td>60</td>
<td>59</td>
<td>84</td>
<td>48</td>
<td>64</td>
<td>38</td>
</tr>
<tr>
<td>Time Post-onset</td>
<td>3 years, 9 months</td>
<td>12 months</td>
<td>1 year, 1 month</td>
<td>2 years, 9 months</td>
<td>1 year, 4 months</td>
<td>3 years, 8 months</td>
<td>1 year, 7 months</td>
</tr>
<tr>
<td>Etiology</td>
<td>Stroke</td>
<td>Multiple Hemorrhages</td>
<td>Hemorrhage</td>
<td>Stroke</td>
<td>Stroke</td>
<td>Hemorrhage</td>
<td>Stroke</td>
</tr>
<tr>
<td>Lesion location</td>
<td>L insula, parietal-frontal, middle temporal</td>
<td>L posterior temporal, L parietal-occipital, R posterior inferior frontal</td>
<td>L inferior frontal</td>
<td>L posterior temporal-occipital, parietal-occipital</td>
<td>L parietal-frontal, basal ganglia</td>
<td>L inferior superior parietal, superior occipital</td>
<td>L frontal-parietal, insula, basal ganglia</td>
</tr>
<tr>
<td>Education</td>
<td>12</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>12</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Aphasia type</td>
<td>Anomic</td>
<td>Anomic</td>
<td>Fluent</td>
<td>Anomic</td>
<td>Fluent</td>
<td>Nonfluent</td>
<td>Nonfluent</td>
</tr>
<tr>
<td>Language Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Naming Test Score (max 60)</td>
<td>19</td>
<td>32</td>
<td>29</td>
<td>24</td>
<td>9</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>PW Repetition (max 30)</td>
<td>2</td>
<td>26</td>
<td>15</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>CNC</td>
</tr>
<tr>
<td>Pyramids and Palm Trees (max. 52)</td>
<td>49</td>
<td>40</td>
<td>39</td>
<td>36</td>
<td>42</td>
<td>35</td>
<td>48</td>
</tr>
<tr>
<td>Word-to-picture Matching (max 48)</td>
<td>46</td>
<td>47</td>
<td>33</td>
<td>41</td>
<td>43</td>
<td>31</td>
<td>42</td>
</tr>
<tr>
<td>Repetition of 132 words matched to potential treatment words (through headphones)</td>
<td>&gt;80% w/ picture</td>
<td>&gt;80%</td>
<td>&gt;80%</td>
<td>&gt;80%</td>
<td>&gt;80% live repetition w/ picture</td>
<td>&gt;80% live repetition w/ picture</td>
<td></td>
</tr>
<tr>
<td>Cognitive Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impairment category on mWCST perseveration score (after conversion to standard score for age and education)</td>
<td>Average</td>
<td>Low Average</td>
<td>Mild Impair.</td>
<td>Mild Impair.</td>
<td>Mild-Mod. Impair.</td>
<td>Mild-Mod. Impair.</td>
<td>*</td>
</tr>
</tbody>
</table>

Grey boxes indicate patient is within normal range on the task. CNC=could not complete
* JLC was not tested on this task
Table 3.2 – Treatment Outcomes and Memory Test Scores

<table>
<thead>
<tr>
<th></th>
<th>QPR</th>
<th>VZN</th>
<th>OJN</th>
<th>DBR</th>
<th>XTC</th>
<th>YLC</th>
<th>JLC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Probe Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment Probe</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Final Probe</td>
<td>19</td>
<td>20</td>
<td>18</td>
<td>19</td>
<td>19</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Maintenance 1&lt;sup&gt;1&lt;/sup&gt;</td>
<td>18</td>
<td>19</td>
<td>18</td>
<td>14</td>
<td>10</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Maintenance 2&lt;sup&gt;1&lt;/sup&gt;</td>
<td>18</td>
<td>20</td>
<td>14</td>
<td>17</td>
<td>11</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Sessions to criterion</td>
<td>16</td>
<td>16</td>
<td>24</td>
<td>32</td>
<td>16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sessions to plateau</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>36</td>
</tr>
<tr>
<td><strong>Explicit Memory Testing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Verbal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition Memory</td>
<td>74%</td>
<td>78%</td>
<td>76%</td>
<td>69%</td>
<td>55%</td>
<td>55%</td>
<td>68%</td>
</tr>
<tr>
<td><strong>Spatial (BFLT)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Learning (accuracy on final learning trial)</td>
<td>100%</td>
<td>70%</td>
<td>47%</td>
<td>33%</td>
<td>17%</td>
<td>0%</td>
<td>30%</td>
</tr>
<tr>
<td>Delayed Spatial Recall</td>
<td>100%</td>
<td>70%</td>
<td>43%</td>
<td>47%</td>
<td>7%</td>
<td>0%</td>
<td>57%</td>
</tr>
<tr>
<td>Delayed Spatial Recognition</td>
<td>100%</td>
<td>100%</td>
<td>80%</td>
<td>60%&lt;sup&gt;2&lt;/sup&gt;</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Shading indicates significant difference from pre-treatment probe. The only significant difference between EF and EL was for DBR’s final probe. All patients have impaired verbal explicit memory relative to normal controls (more than 2SD below mean for their age group).

<sup>1</sup> Due to patients’ schedules, it was not always possible to collect maintenance data at exactly 1 and 6 months post-treatment.

<sup>2</sup> Because DBR is 84 years old, this score is actually within normal range (mean for 80+ = 84% +/- 28%)
Table 3.3 – Errors produced during treatment

<table>
<thead>
<tr>
<th>Errors produced in 4 treatment sessions before probe</th>
<th>QPR</th>
<th>VZN</th>
<th>OJN</th>
<th>DBR</th>
<th>XTC</th>
<th>YLC</th>
<th>JLC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Probe 1</strong></td>
<td>EF</td>
<td>EL</td>
<td>EF</td>
<td>EL</td>
<td>EF</td>
<td>EL</td>
<td>EF</td>
</tr>
<tr>
<td>229</td>
<td>64</td>
<td>44</td>
<td>178</td>
<td>39</td>
<td>251</td>
<td>20</td>
<td>289</td>
</tr>
<tr>
<td><strong>Probe 2</strong></td>
<td>109</td>
<td>21</td>
<td>13</td>
<td>0</td>
<td>34</td>
<td>4</td>
<td>78</td>
</tr>
<tr>
<td><strong>Probe 3</strong></td>
<td>46</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>61</td>
</tr>
<tr>
<td><strong>Probe 4</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td><strong>Probe 5</strong></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>4</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td><strong>Probe 6</strong></td>
<td>0</td>
<td></td>
<td>33</td>
<td>1</td>
<td>104</td>
<td>0</td>
<td>102</td>
</tr>
<tr>
<td><strong>Probe 7</strong></td>
<td>0</td>
<td></td>
<td>25</td>
<td>1</td>
<td>90</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td><strong>Probe 8</strong></td>
<td>0</td>
<td></td>
<td>20</td>
<td>1</td>
<td>109</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Probe 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>93</td>
</tr>
<tr>
<td><strong>Probe 10</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table shows the total number of errors produced over all trials during the 4 treatment sessions that occurred before each test probe.
Table 3.4 - Generalization

<table>
<thead>
<tr>
<th></th>
<th>QPR</th>
<th>VZN</th>
<th>OJN</th>
<th>DBR</th>
<th>XTC</th>
<th>YLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe Scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Probe</td>
<td>19</td>
<td>20</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>20</td>
<td>16</td>
<td>15</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2nd Exemplar Final Probe</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>15</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Maintenance 1</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>15</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Maintenance 2</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>15</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2nd Exemplar Maint 1</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>15</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2nd Exemplar Maint 2</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>15</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Scores tended to be similar in the treatment items and 2nd exemplars, indicating that the treatment does generalize to different exemplars of the same items. 

*OJN suffered a seizure just before we were to test the 2nd exemplars. However, he recovered and returned for his first maintenance test 2 months after his final probe. YLC’s data for the EL final probe 2nd exemplars are not available.
Conclusions and Future Directions

The preceding chapters have presented studies focusing on generalization and maintenance of behavioral interventions for acquired reading disorders and naming deficits. Studying these variables is of great importance to public health. Stroke survivors have a limited number of sessions for cognitive treatment and must make the most efficient use of that time.

These variables are also important to cognitive rehabilitation research because they can be studied across a wide range of patients and can provide a common thread between case studies of different treatments. Converging evidence, even from case studies, can be used to improve existing treatments. In Chapter 3 we demonstrate a naming treatment that maintains well. As we see from the data for the 2nd exemplar pictures, the treatment effect also generalizes to other depictions of the same object (see Table 3.4). Both the EF and EL treatment effects generalize to other exemplars, but not to novel concepts. One could imagine drawing from these data and others to create a more efficient naming treatment that generalizes and maintains using Kiran & Thompson’s (2003) low typicality items to increase generalization and an errorful treatment paradigm with sessions until criterion is reached to achieve better maintenance of treatment.

Chapters 2 and 3 also generate ideas for new variables that might be studied. Though it is difficult to operationalize, the variable of effort emerges as an important concept that could be pursued further. In Chapter 3, we discuss the possibility that increased effort and attentional engagement could explain the maintenance of treatment effect in the EF condition, but increased effort could also explain some of the studies in the literature showing that treatment with more difficult items, which are presumably more attentionally demanding, is associated with generalization. This has been shown in naming treatments (Kiran & Thompson, 2003) and
syntax treatments (Thompson et al., 2003). It could also be one of the underlying reasons for the context effect we discuss in Chapter 2. We show that text reading is improving from the bottom up, but the context of phrases also seems to be facilitative for the people with phonological text alexia. The context provided by the phrase could offer supportive semantic and grammatical cues, essentially making reading of the difficult words easier. However, there is evidence that patients with phonological text alexia have difficulty reading words that are low in semantic value when they are surrounded by concrete words, as they might be in a phrase (Friedman & Lott, 1995). Therefore, reading the words within phrases could also be more difficult for these patients which, the data from Chapter 3 tells us, could be associated with greater engagement and better initial learning.

Complicating matters further, what is difficult enough to be engaging for one patient can be fatiguing or distracting for another. One example from our own lab is patient TJN, a person with phonological alexia who was originally enrolled in the MOR study but was unable to complete it because it was simply too tiring for her. We designed another treatment for her in which she practiced difficult words within short phrases (Snider et al., 2008). Not only did she show measurable generalization in this treatment, but she reported greater ease with reading material in her everyday life as well.

Although the design of our MOR study required that we work with patients less impaired than TJN, our anomia study included a wide range of severities. If we had chosen the number of treatment sessions based on what would occur in the clinic (about 10 sessions; Fillingham et al., 2006), we would not have known that the more severe patients could eventually learn significantly above baseline levels. More importantly, we would not have learned that they do not seem to benefit from a treatment that prevents them from making errors. In this case, what
was engaging for the milder patients seemed to work for the severe patients as well. Even the patients who demonstrated perseverative tendencies on the MWCST were able to alternate between two treatment pictures. Indeed, the alternating pictures did not seem to make the EL condition difficult enough. There is always a trade-off between effort and error, however, and making both treatments more difficult might also have been problematic. In piloting the study in Chapter 3, we discovered that, even in a moderate patient (similar to patient OJN; see Tables 3.1 and 3.2), trying to alternate between five different treatment pictures instead of two was too difficult for him. However, future studies might investigate effort by manipulating the number of intervening items as a way of operationalizing level of effort. Although the more severe patients in Chapter 3 were able to learn above baseline levels, we continue to look for ways in which naming treatment could be more effective for them. We are currently investigating the use of idiomatic phrases and orthography in addition to phonologic cues to improve their production of picture names.

In the preceding studies, we report four out of seven people with anomia who were unable to learn 40 picture names over many weeks of therapy and six people with alexia who read dutifully, daily for eight weeks, only to achieve effects sizes smaller than any yet reported in the literature. These results might seem to represent a large deposit of time and energy for a comparatively small return. It is true that behavioral intervention is slow and laborious, but it is currently the best option available. Recent findings showing improved naming after repetitive trans-cranial magnetic stimulation (rTMS) to right pars triangularis (anterior Broca’s area) may point the way to more advanced techniques for aphasia rehabilitation (Naeser et al., 2005). The hypothesized mechanism for improvement after rTMS is suppression of maladaptive disinhibition due to missing input from damaged left hemisphere language areas, perhaps
allowing intact portions of the left hemisphere to resume functioning. Data from our own laboratory are consistent with the idea that improved language at the chronic stage is related to a shifting from right hemisphere language homologs to recruitment of preserved tissue in left hemisphere perilesional areas (Kurland et al., 2008), but we achieved this shift with behavioral intervention alone. However, unpublished 6 month maintenance data for the same patient shows a drop in performance and a reduction in left hemisphere activation. Conversely, in Naeser’s study (2005), the patients maintained naming improvements 8 months after rTMS, though it should be mentioned that the improvements were small (an average of 4 items gained). An unpublished follow-up to Naeser’s study shows that the improvements in naming also generalize to spontaneous speech as measured by mean length of utterance and other measures in a standardized picture description (Hamilton et al., 2010). These exciting data show generalization and maintenance as a result of only ten days of twenty-minute treatments. However, these preliminary data do not diminish the need to continue to understand how, why and for whom behavioral treatments are effective. It is likely that treatments like rTMS in combination with behavioral intervention will be the future of aphasia rehabilitation.
Appendix A

List of Publications and Copyright Permissions


*Neuropsychological Rehabilitation*, In Press.

Our Ref: KB/PNRH/P3189

23rd August 2010

Dear Elizabeth Lacey

Thank you for your correspondence requesting permission to reproduce the following material from our Journal in your doctoral dissertation:

Multiple Oral Re-reading treatment for alexia: The parts may be greater than the whole – vol 20 no 4 2010 pp 601-623

We will be pleased to grant entirely free permission on the condition that you acknowledge the original source of publication and insert a reference to the Journal’s web site:

http://www.informaworld.com

Thank you for your interest in our Journal.

Yours sincerely

Karin Beesley
Permissions Administrator
Taylor & Francis Group

Taylor & Francis Group is a trading name of Informa UK Limited, registered in England under no. 1072954
Appendix B: Examples of untrained reading passages from Chapter 2

MOR Training passage Example

The Training passage taken from a 6-8th grade reading workbook (Instructional Fair, Inc., 1990) contained the following words, 60-80% of which were re-used in the untrained passages. The words in the Training passage were coded as either Content, Functor or Other (except the articles a. an and the which were excluded from all analyses). There are 112 Content words (60-80%=67-90), 115 Functor words (60-80%=69-92), 33 articles and 40 words coded as “other”.

The passage was about King Arthur.

<table>
<thead>
<tr>
<th>Content</th>
<th>Functor</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>told</td>
<td>have</td>
<td>stories</td>
</tr>
<tr>
<td>king</td>
<td>been</td>
<td>knights</td>
</tr>
<tr>
<td>arthur</td>
<td>about</td>
<td>years</td>
</tr>
<tr>
<td>thousand</td>
<td>and</td>
<td>legends</td>
</tr>
<tr>
<td>arthur</td>
<td>his</td>
<td>lived</td>
</tr>
<tr>
<td>court</td>
<td>for</td>
<td>knights</td>
</tr>
<tr>
<td>magic</td>
<td>more</td>
<td>arthur's</td>
</tr>
<tr>
<td>city</td>
<td>than</td>
<td>protected</td>
</tr>
<tr>
<td>camelot</td>
<td>according</td>
<td>arthur's</td>
</tr>
<tr>
<td>city</td>
<td>to</td>
<td>carved</td>
</tr>
<tr>
<td>was</td>
<td>and</td>
<td>trimmed</td>
</tr>
<tr>
<td>beautiful</td>
<td>his</td>
<td>knights</td>
</tr>
<tr>
<td>arthur</td>
<td>in</td>
<td>perfectly</td>
</tr>
<tr>
<td>was</td>
<td>of</td>
<td>knights</td>
</tr>
<tr>
<td>wise</td>
<td>as</td>
<td>insisted</td>
</tr>
<tr>
<td>good</td>
<td>as</td>
<td>seats</td>
</tr>
<tr>
<td>brave</td>
<td>and</td>
<td>quarrels</td>
</tr>
<tr>
<td>palace</td>
<td>and</td>
<td>knights</td>
</tr>
<tr>
<td>weak</td>
<td>who</td>
<td>quarrels</td>
</tr>
<tr>
<td>fought</td>
<td>in</td>
<td>fights</td>
</tr>
<tr>
<td>evil</td>
<td>and</td>
<td>plates</td>
</tr>
<tr>
<td>wonderful</td>
<td>and</td>
<td>knights</td>
</tr>
<tr>
<td>great</td>
<td>one</td>
<td>knights</td>
</tr>
<tr>
<td>court</td>
<td>of</td>
<td>swords</td>
</tr>
<tr>
<td>was</td>
<td>most</td>
<td>jumped</td>
</tr>
<tr>
<td>famous</td>
<td>in</td>
<td>declared</td>
</tr>
<tr>
<td>round</td>
<td>before</td>
<td>needed</td>
</tr>
<tr>
<td>table</td>
<td>was</td>
<td>fighting</td>
</tr>
<tr>
<td>round</td>
<td>that</td>
<td>suddenly</td>
</tr>
<tr>
<td>table</td>
<td>in</td>
<td>knights</td>
</tr>
<tr>
<td>built</td>
<td>of</td>
<td>watched</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>arthur</td>
<td>was</td>
<td>building</td>
</tr>
<tr>
<td>had</td>
<td>at</td>
<td>arthur's</td>
</tr>
<tr>
<td>great</td>
<td>for</td>
<td>solved</td>
</tr>
<tr>
<td>long</td>
<td>every</td>
<td>knights</td>
</tr>
<tr>
<td>table</td>
<td>and</td>
<td>carried</td>
</tr>
<tr>
<td>stood</td>
<td>his</td>
<td>knights</td>
</tr>
<tr>
<td>middle</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>great</td>
<td>each</td>
<td></td>
</tr>
<tr>
<td>hall</td>
<td>was</td>
<td></td>
</tr>
<tr>
<td>fine</td>
<td>at</td>
<td></td>
</tr>
<tr>
<td>chair</td>
<td>for</td>
<td></td>
</tr>
<tr>
<td>set</td>
<td>every</td>
<td></td>
</tr>
<tr>
<td>table</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>table</td>
<td>his</td>
<td></td>
</tr>
<tr>
<td>knight</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>lad</td>
<td>each</td>
<td></td>
</tr>
<tr>
<td>seat</td>
<td>was</td>
<td></td>
</tr>
<tr>
<td>gold</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>were</td>
<td>with</td>
<td></td>
</tr>
<tr>
<td>happy</td>
<td>most</td>
<td></td>
</tr>
<tr>
<td>take</td>
<td>to</td>
<td></td>
</tr>
<tr>
<td>seat</td>
<td>any</td>
<td></td>
</tr>
<tr>
<td>have</td>
<td>at</td>
<td></td>
</tr>
<tr>
<td>head</td>
<td>all</td>
<td></td>
</tr>
<tr>
<td>table</td>
<td>but</td>
<td></td>
</tr>
<tr>
<td>trouble</td>
<td>few</td>
<td></td>
</tr>
<tr>
<td>spread</td>
<td>that</td>
<td></td>
</tr>
<tr>
<td>one</td>
<td>they</td>
<td></td>
</tr>
<tr>
<td>night</td>
<td>should</td>
<td></td>
</tr>
<tr>
<td>dinner</td>
<td>at</td>
<td></td>
</tr>
<tr>
<td>broke</td>
<td>of</td>
<td></td>
</tr>
<tr>
<td>became</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>began</td>
<td>at</td>
<td></td>
</tr>
<tr>
<td>fly</td>
<td>out</td>
<td></td>
</tr>
<tr>
<td>threw</td>
<td>amon</td>
<td></td>
</tr>
<tr>
<td>drew</td>
<td>soon</td>
<td></td>
</tr>
<tr>
<td>began</td>
<td>to</td>
<td></td>
</tr>
<tr>
<td>duel</td>
<td>as</td>
<td></td>
</tr>
<tr>
<td>arthur</td>
<td>them</td>
<td></td>
</tr>
<tr>
<td>had</td>
<td>at</td>
<td></td>
</tr>
<tr>
<td>feet</td>
<td>each</td>
<td></td>
</tr>
<tr>
<td>fought</td>
<td>other</td>
<td></td>
</tr>
<tr>
<td>put</td>
<td>their</td>
<td></td>
</tr>
<tr>
<td>death</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>arthur</td>
<td>that</td>
<td></td>
</tr>
<tr>
<td>knew</td>
<td>anyone</td>
<td></td>
</tr>
<tr>
<td>better</td>
<td>who</td>
<td></td>
</tr>
<tr>
<td>way</td>
<td>would</td>
<td></td>
</tr>
</tbody>
</table>
stop
was
knock
castle
door
old
man
know
solve
problem
said
arthur
strange
man
set
round
table
problem
table
was
big
seat
light
king
went
war
round
table
has
head
matter
sat
were
equal

be
to
but
that
he
to
there
on
and
i
how
to
he
and
as
and
his
about
was
enough
to
all
yet
enough
to
be
along
when
to
and
since
no
no
where
they
always
The following passage was created using 60-80% of the phrases from the Training passage described above. There are 298 words and 27 re-used phrases. Bold type indicates content words from the Training passage. Underlining indicates functor words from the Training passage. Italics indicate “other” words from the Training passage.

Video games are quite odd these days. Many games are made [according to the legends] of [every knight and his lady] from [more than a thousand years] ago. My daughter usually enjoys these types of games. She thinks they are the [most wonderful things]. However, she received a game, which came in a box that looked [carved and trimmed with gold], which she did not like. The game’s theme is some [wise and good] warriors getting into a battle [one night at dinner].

The main goal is to [seat all the knights] in the proper order [at the table]. You start by [building a round table]. You [solve your problem] by listening to which chairs the men want. Then you place the men in the correct order. This task isn’t very easy. My daughter placed two men together at the [great long table]. These men [began to duel]. She discovered that the fighters could not just [take any seat at all], so she [needed a better way]. She tried a different order. This time the men took their chairs and [threw them at each other]. She tried to [stop the fighting] but then a huge riot [broke out among the knights]. The whole room [drew their swords]. The men basically [went to war]. The soldier she placed at the [head of the table] was [put to death]. The soldier’s headless ghost [jumped to his feet] and [began to fly] around the room laughing. A ghost, which [has no head], is quite scary. Next, there was a [knock on the
castle door]. A wizard informed my daughter that she hadn’t [protected the weak] very well.

The whole game then started over from the beginning.

My daughter decided that the game wasn’t much fun, so she read a book instead.
Marcy decided to go swimming and cool off. It was very hot, yet she was staying in a hotel that had no air conditioner. At the beach, Marcy began searching for an area that attracted no more than a few people. She walked and walked until she found an area in which the ocean met jagged cliffs. At the bottom of the cliffs she found a fine beach without anyone there. She went down the sides of the cliff. When she got close to the bottom, seagulls flew out of each and every crevice and surprised her. Marcy lost her grip and fell. She tried to grab something, but she knew she had little hope.

Marcy opened her eyes and blinked them twice. She saw before her a man without legs. Instead, a fin-like tail took their place. He moved his hands towards her and made noises. He must have been trying to communicate. He led her to his beautiful underwater world among the fish. How should she be able to breathe in a place that was filled with water? They swam into his great palace. The array of food in his dining room included the most seafood Marcy had ever seen and enough desserts to feed a small country. After they finished the big dinner, he invited Marcy to a wonderful playroom unlike any other. She spent hours playing games similar to basketball and baseball, and some unfamiliar games. She felt tired and decided to rest for a while.
The warm sand and cool breeze felt good. As she woke up, she realized she was sitting back at the cliffs again, where she'd fallen, and she was feeling dizzy. She wondered if she had been dreaming it all. Just then, however, she noticed her clothes were soaking wet!
MOR Content Passage Example

The following passage was created using 60-80% of the content words from the Training passage about King Arthur described above. There are 298 words and 70 re-used content words. Bold type indicates content words from the Training passage. Underlining indicates functor words from the Training passage. No words coded as “other” from the Training passage were re-used.

My **good** friend **has** terrific parties. Last **night**, a whole bunch **of** us **went** together. **When** we walked through the **door**, we noticed everyone **was** wearing costumes. There **was** a **brave** soldier, a **beautiful** butterfly, a **king**, a scary monster **with** six **feet**, **and** a **famous** actor.

The party **began** with a **man** doing **magic** tricks. Then the **man** **told** spooky tales. Next, we played many **great** games. We divided into two **equal** teams. We **fought** a **war** **and** pretended to **duel** over the **palace** **and** the **city**. We **became** quite wild. Someone **threw** a pillow into the **middle** **of** the room. Everyone **had** fun until a **chair** tipped over. We saw it **knock** into a **fine**, **old** vase. We saw the vase **fly** across the room. We could not prevent it from breaking. We **knew** this **was** a **great problem**. We **were** in **trouble**.

We **fought** over what we should do. Finally, someone **said** something **wise**. She **suggested** we **solve** our **problem** by making up **for** it. We prepared a **big dinner**. Afterwards, we **set** dishes made **of** **gold** **on** the **table**. Then we **spread** the food upon the **table**. Just **as** my friend's parents came into the dining room, we **sat** them **down** at the **table**. The mother chose the **seat** at the **head** **of** the **long**, wooden dining **table**, while the father chose the **seat** at the **other** **head** **of** the dining **table**. The parents looked worried, **but** **began** to eat the meal. **They** looked up from the **table** **in** a **happy way** and exclaimed they **had** never eaten a **better** meal! The
mother stood up. She forgave us for whatever evil we did. She then led us toward a round bucket filled with a thousand prizes. We know this was the best party ever!
References


Effects of training multiple form classes on acquisition, generalization and maintenance of word retrieval in a single subject. *Aphasiology, 12*(7/8), 575-585.


