CHOOSING REMEDIATION TARGETS FOR NAMING DEFICITS IN PROBABLE ALZHEIMER DISEASE: DOES TYPICALITY MATTER?

By

CLAUDIA A. MORELLI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2006

This dissertation is dedicated to my incredible husband, David Efros, my amazing mother, Claire Morelli and my sisters; and in memory of my wonderful father, John. They have been an integral part of my life and their belief in me has been a fortress of encouragement, hope, and steadfast love.

ACKNOWLEDGMENTS

First, I thank and acknowledge each participant and their family and/or caregivers for their time, effort, and for extending themselves and their home to allow me to work with them to complete this research project. I will always remember them and the devastating effects of Alzheimer disease, and this will influence the focus of my future clinical research. I also extend my gratitude and appreciation for the UF Memory Disorder Program, especially David B. Efros, Dr. Kenneth M. Heilman, Dr. Glen Finney, Dr. Kimford Meador, Dr. Catherine Price, and Jennifer Rembisz, whose kind and diligent efforts resulted in referrals for me to recruit these participants.

I also express my deep appreciation for my committee. First, I want to thank my mentor, Dr. Lori J.P. Altmann, for her vast knowledge of neurolinguistics that provided a refreshing perspective, especially with regards to research in adults with probable Alzheimer disease. Working with Dr. Altmann allowed me to broaden my thought process and expand my approach to this dissertation. These qualities, combined with her guidance, leadership, and patience have been of great value to me during the exciting and challenging process of becoming a Ph.D. I also acknowledge Dr. Diane Kendall for her expert opinion on clinical research in adult neurogenics, and for the many opportunities to learn more about single subject design, especially during my pre-doctoral fellowship with her at the VA Brain Rehabilitation Research Center. Her outstanding clinical skills and perspective as a compassionate speech-language pathologist combined with her insights have had a great influence on me and have motivated my interest in theoretically driven treatments. I also gratefully acknowledge Dr. Ira Fischler who shared his extensive knowledge of psychology, which played an important role in this dissertation. Dr. Kenneth Heilman, with his behavioral neurology expertise combined with his incredible commitment and compassion for patients, has taught me so much in such inspiring and meaningful ways. Each of these remarkable individuals has contributed to my development by their expertise and great role modeling.

I thank the Department of Communication Sciences and Disorders at the University of Florida, especially Dr. Christine Sapienza, Dr. Samuel Brown, and Dr. Scott Griffiths, for their support and encouragement, as well as the wonderful opportunities to teach the undergraduates. Also, I will always remember the incredible opportunities to guest lecture for Dr. Howard Rothman and Dr. Bonnie Johnson (including the lecture swaps with Dr. Johnson). I extend my appreciation to the entire faculty and staff for their professionalism, collegial, and positive approach. In particular, Idella King, Cassie Mobley, Debbie Butler, and Addie Pons deserve a special thanks. The computer support from Neal Musson will always be appreciated.

I also want to acknowledge the members of Dr. Altmann's Language over the Lifespan Lab. First I thank Alexia Frederiksen, who spent hours developing the stimuli, coordinating and working with us, and whose attention to details was truly awesome. Her dedication to the project was sincerely appreciated and always will be. Additional thanks go to Katie Chiarella, Andrea Holt, and Angel Ziesk for their wonderful help in preparing the stimuli; and to Maisa Haj Tas, Ashley Mullen, and Deb Gober for lending their incredible artistic talent. Alexia Frederiksen, Erin Hunt, and Rachel Hogue deserve acknowledgement for their excellent questions and feedback during the protocol training, which prepared them for testing the young adults for a related project. Becca Huy also played an important role by transcribing a large amount of the data and working closely with me with on the reliability checks. Her devoted attention to the details of accurate transcription, combined with her enthusiasm, and willingness to help was wonderful. Also, I thank Charlene Cohen-Deroy for her diligent efforts and commitment to excellence in coding and scoring the Word Reading subtests, as well as calculating the phoneme length for the stimulus words. All of these students demonstrated a keen awareness of the importance of maintaining the integrity of the data. Without the combined efforts of these wonderful students, this dissertation would have taken longer. Finally, I thank the collective insights of the members of the Language over the Lifespan lab, especially those who provided suggestions for the semantic cues. It was their integrative efforts that helped to improve various aspects of this study.

I also would like to thank the Director and members of the Brain Rehabilitation Research Center, at the Malcom Randall Veterans Affairs Medical Center, Dr. Leslie Gonzalez Rothi, Dr. Stephen Nadeau, and Dr. Jay Rosenbek, for the pre-doctoral fellowship opportunity that Dr. Diane Kendall provided. I also thank Haijing Qin for her statistical support. I acknowledge other investigators and clinicians, especially Nan Musson, Susan Leon, and Amy Rodriguez. I appreciated the ongoing support of the staff, especially Brenda Stidham, Sue Nadeau, Joy McCallum, Sandy Davis, Lisa DeEmmanuel, and Lynn Dirk. Further appreciation is extended to the incredible VA library staff, for their expertise, time, and consideration: Marsha H. White, Ellen L. Umans, Marylyn Gresser, and Harold A. Boyce.

Two fine clinicians and professors have played an important, yet earlier role in my career. I extend my thanks and appreciation to Dr. Robert C. Marshall, for inspiring me years ago in my master's program, continuing his belief in me, and his incredible mentoring in a supportive, enthusiastic and gentle manner. If it were not for him, I would never have started this Ph.D. endeavor. Also, I thank Donna Chadwick for teaching me the fundamental aspects of being an effective clinician during my undergraduate program.

My friends have also been a great source of inspiration and hope. Dr. Sally Ann Giess, whom I met in my first year at the University of Florida, continues to be a great friend, and role model. I acknowledge my other friends and colleagues from the UF CSD Department: Dr. Ann Marie Knight, Dr. Amber Hollingsworth, Dr. Judith Wingate, Dr. Lynnette Bardolf, Susan Leon, and Maisa Haj Tas. I thank The Reverend Jeremy Hole for his support and encouragement. I will always appreciate my dear friends Ann Case and Nancy Merrifield for their wonderful presence in my life.

Most importantly, I am forever grateful for my family. Their patience, ongoing love, support, and belief in me have been truly wonderful. Thus, I thank them: my husband David, mother, Claire, and my sisters and their families, Jeanmarie; Donna and Ryan, Buzz and Charlie; Claire and Ross, Martina and Fabiana; as well as Everett, my father-in-law and Marie, my mother-in-law. Also, I thank my cousins and their families: Marybeth, Mary Catherine, Luigi, Lou and Joey, and Lee and Mary Anne. Finally, I would like to acknowledge all those who came before, paving my way through their loving support and kindness, including my father John; and grandmothers, Florence McNulty and Anna Efros Kramer.

This dissertation was partially supported by my teaching and research assistantships from the UF Department of Communication Sciences and Disorders. It was also partially funded by Dr. Altmann's Advancing Academic-Research Careers award, which she had received from the American Speech-Language-Hearing Association. I am grateful for these funding sources.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CHOOSING REMEDIATION TARGETS FOR NAMING DEFICITS IN
PROBABLE ALZHEIMER DISEASE:
DOES TYPICALITY MATTER?

By

Claudia A. Morelli

August 2006

Chair: Lori J.P. Altmann
Major Department: Communication Sciences and Disorders

The purpose of this single session exploratory study was to determine if remediation targets could be selected to improve picture naming and promote generalization to untrained items in 12 adults with early probable Alzheimer disease (PAD). Because theories support competing predictions about the relative effects of training typical versus atypical semantic category members with respect to subsequent generalization, this study contrasted the effects of training typical and atypical semantic category exemplars. Specifically, it examined changes in picture naming and category generation following initial repetition of the items at pretest followed by a semantic training. Stimuli included 24 items from each of 3 semantic categories, half of the items were typical and half atypical based on Rosch's norms. Two categories received training, one using a subset of typical items, one using a subset of atypical items; the third category remained untrained to track effects of repetition. The untrained category showed nonsignificant improvements at post-test. Accuracy scores improved for all trained categories; however, only trained, typical items were named significantly faster at post-test. Generalization was found in the untrained typical items in the categories that were trained with typical items, which were named more accurately at posttest. Neither atypical items from the categories that were trained with typical items nor any of the items in categories that were trained with atypical items showed improved accuracy after training. These findings are consistent with those of other researchers investigating the semantic deterioration in adults with PAD. Several studies have found an advantage for typical items in this population. These findings are attributed to the redundant connections among features in typical items, which allow them to be more resilient in the face of progressive damage, allowing them to be more responsive to intensive semantic training. Atypical items lack the redundancy of connections and, thus, are more vulnerable to damage. These findings are extremely encouraging for the development of principled strategies for choosing items to encourage generalization in the remediation of anomia, as well as for the development of lexical-semantic treatment paradigms for individuals with early PAD.

CHAPTER 1
INTRODUCTION

Probable Alzheimer disease (PAD), a progressive neurodegenerative disease, is the most prevalent type of acquired cognitive dysfunction (Thal, 1999) and affects over four million Americans (Kawas & Katzman, 1999). The prevalence of PAD accounts for 50-80% of patients with dementia (Chui, 1989) and comprises the greatest percentage of the dementias in the geriatric population (Chui, 1989; Henderson & Finch, 1989; Sclan & Kanowski, 2001). Furthermore, PAD is already considered to be a major public health concern, and by 2025, the number of Americans affected by the disease is expected to be 12 million (Rosenberg, 2005) and 14 million by 2050 (Katzman & Fox, 1999). Although there is no known cure for PAD, and only symptomatic treatment is available at this time, current research endeavors are aiming to halt the progression of the disease as well as prevent its occurrence (Petersen et al., 2001; Rosenberg, 2005). Reducing the disability until this occurs is important (Rothi et al., 2005). Furthermore, when drug therapies are available that halt the progression of the disease, there will be a demand for linguistic and cognitive rehabilitation for this population. Therefore, it is incumbent upon speech-language pathologists to develop and test appropriate treatments for this population now, so that tested methods are available that address the language difficulties in PAD (ASHA, in press), which largely revolve around word finding

Word finding difficulties (i.e., anomia), as measured by poor performance on picture naming tasks have been widely reported in adults with PAD (Barbarotto, Capitani, Jori, Laiacona, & Molinari, 1998; Bayles, 1982; Bayles & Tomoeda, 1983;

Benson & Geschwind, 1985; Heilman, 2005; Hodges & Patterson, 1995; Hodges, Patterson, Graham, & Dawson, 1996; Huff, Corkin, & Growdon, 1986; Huff, Mack, Mahlmann, & Greenberg, 1988; Kirshner, Webb, & Kelly, 1984; Martin & Fedio, 1983; Smith, Faust, Beeman, Kennedy, & Perry, 1995; Smith, Murdoch, & Chenery, 1989; Williams, Mack, & Henderson, 1989; Williamson, Adair, Raymer, & Heilman, 1998). However, to date, there are only three PAD anomia treatment studies (Abrahams & Camp, 1993; Ousset et al., 2002; Rothi et al., 2005), and only one has reported generalization to untrained items (Abrahams & Camp, 1993). Consequently, developing treatments for this population that reduce the word finding difficulty is important (Rothi et al., 2005), especially those that promote generalization.

Generalization of learning, an essential objective for rehabilitation (Kearns, 1989; Thompson, 1989), has been described as an observation of an occurrence of a particular trained behavior in a context that has not been trained (McReynolds, 1989). Generalization is important for several reasons, including the clinical accountability of the speech-language pathologist to use methodologies to measure the effectiveness of treatment (Kearns, 1989). The mechanism underlying generalization has been a long-standing question (Martin, Laine, & Harley, 2002), and although cognitive models have been advanced over time, the construct of generalization remains a mystery (Francis, Clark, & Humphreys, 2002). Perhaps due to this, there is still a lack of generalization reported in the literature (Francis et al., 2002; Kearns, 1989; Kiran & Thompson, 2003b; McNeil et al., 1998; McReynolds, 1989; Nickels, 2002; Thompson, 1989; Thompson, Ballard, & Shapiro, 1998; Thompson et al., 1997; Thompson, Shapiro, Kiran, & Sobecks, 2003). Moreover, attempts to replicate methods that have promoted generalization in initial studies have not always been successful (McReynolds, 1989; Thompson, 1989).

There are two types of generalization that, according to Thompson (1989), play an important role in aphasia rehabilitation: response generalization (i.e., when an untrained response occurs after other responses have been trained (e.g., production of different, untrained items) and stimulus generalization (i.e., when there is carryover of a trained behavior to a different, untrained stimulus condition, e.g., outside of the clinic). Without the former, there would be an inordinate number of responses to train, and without the latter, trained behaviors would occur only in the clinic setting (Thompson, 1989). Although both of these are important, the current study focuses only on response generalization.

Several investigators (Hillis, 1998; Kiran & Thompson, 2003b; Plaut, 1996; Thompson, 1989) have suggested that treatment studies might be able to influence the recovery of language if they target factors that are modifiable. A principled selection of materials for training as well as for testing generalization is just one example of this (Hillis, 1998; Kiran & Thompson, 2003b; Plaut, 1996; Thompson, 1989; Thompson et al., 2003). One technique that has successfully induced generalization is to have a structural relationship among and across the stimuli (Stokes & Baer, 1977; Thompson, 1989; Thompson et al., 1998; Thompson et al., 1997; Thompson et al., 2003). One approach recommends choosing stimuli that are more versus less complex, a factor that has been shown to be important for increasing generalization to untrained items (Kiran & Thompson, 2003b; Plaut, 1996; Thompson et al., 1998; Thompson et al., 1997; Thompson et al., 2003). This phenomenon has been referred to as the "complexity effect" and the Complexity Account of Treatment Efficacy (Thompson et al., 1998; Thompson et al., 1997; Thompson et al., 2003), which has been successfully used in syntactic treatment studies in the aphasia population.

Kiran and Thompson (2003b) have equated the notion of semantic complexity with category typicality (i.e., by the degree to which a semantic category exemplar is similar to the category prototype, as described Rosch, (1975)) and applied this to a semantic training in adults with aphasia. In this study, Kiran and Thompson (2003b) found generalization to untrained items after atypical items from a category were trained. Findings from this study replicated similar results from a connectionist model that were simulated in a computer experiment of acquired dyslexia (Plaut, 1996). In a similar study, Stanczak, Waters, and Caplan (2006) reported generalization for one of two participants with aphasia, after training atypical category exemplars. However, two other aphasia naming treatment studies (Mayer, Murray, & Karcher, 2004; Stanczak, Waters, & Caplan, 2005) did not find generalization following trainings with either atypical and typical items, although treatment effects were observed. Thus, it is important to further explore the effectiveness of training typical or atypical category exemplars as a strategy to encourage generalization (Murray & Clark, 2005).

All of these studies (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006) have tested the predictions of the Complexity Account of Treatment Efficacy (Thompson et al., 2003), and contrasted it with an older approach based on the family resemblance/prototype hypothesis by Rosch (1975). Arguments based on this view center around the benefits of training qualities or features that are shared among many members of a semantic category, so that many items in the category might benefit, and thus, it provides a rationale for training typical items, which contain a larger proportion of shared features. The structure of a semantic category is an essential element in the rationale for treatments based on typicality (Plaut, 1996); therefore, a population with preserved category structure such as individuals with PAD (Flicker, Ferris, Crook, & Bartus, 1987; Huff et al., 1986; Martin & Fedio, 1983; Nebes, Boller, & Holland, 1986; Salmon, Butters, & Chan, 1999; Schwartz, Marin, & Saffran, 1979; Schwartz, Kutas, Butters, Paulsen, & Salmon, 1996; Warrington, 1975) might be optimal for testing these hypotheses.

Some theories of how semantic representations are affected by PAD have specifically suggested that items that have many shared or intercorrelated features might be more preserved in PAD compared those with representations primarily consisting of distinguishing features, with relatively fewer shared features (Altmann, Kempler, & Andersen, 2001; Devlin, Gonnerman, Andersen, & Seidenberg, 1998; Gonnerman, Andersen, Devlin, Kempler, & Seidenberg, 1997). This literature combined with the evidence that the PAD population has worse performance on atypical items compared to typical items (Sailor, Antoine, Diaz, Kuslansky, & Kluger, 2004; Smith et al., 1995) provides further incentive to compare performance by adults with PAD on typical category exemplars, which have many shared features, and atypical category exemplars, which have relatively fewer shared features.

This exploratory study extends the small body of research comparing the effects of training typical and atypical category exemplars in adults with aphasia (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006) to those with mild-moderate PAD using a modified version of the semantic feature generation task (Boyle, 2004; Boyle & Coelho, 1995; Coelho, McHugh, & Boyle, 2000). Since there were no reports of this type of training with the PAD population, it seemed appropriate to first apply it in a single session to determine its feasibility. Thus, the first aim of the study was to determine if there is a facilitation effect (Hickin, Best, Herbert, Howard, & Osborne, 2002; Howard, 1985; Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985; Patterson, Purell, & Morton, 1983), for trained items from either item repetition alone (Kendall, personal communication, 2005, Fuller et al. 2001; Hickin et al., 2002; Mayer et al., 2004; Patterson et al., 1993; Rothi et al., 2005) or from the repetition plus a semantic training (Drew & Thompson, 1999; Le Dorze, Boulay, Gaudreau, & Brassard, 1994; Wiegel-Crump & Koenigsknecht, 1973). The second aim of the study was to determine whether there was generalization of training to untrained items in the semantically trained categories. The third aim of the study was to examine generalization to an untrained task, category generation, using the same semantic categories as the naming and feature analysis tasks.

The structure of this study is as follows. The next two chapters provide reviews of the literature. Chapter 2 discusses the naming deficits (i.e., anomia) in adults with PAD, and the anomia treatment studies for this population. The third chapter provides background information on semantic theory, the two competing views, the relevant findings from the connectionist model (Plaut, 1996) and the treatment studies in adults with aphasia using semantic category exemplar training with typical and atypical items (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006). This is followed by a description of the current study, including rationale for applying these techniques to adults with PAD. Additional findings from the aphasia treatment literature are identified as they relate to the methodology for our study. Finally, our research questions and predictions based on the above literature are described. In Chapter 4, the study methods are explained in detail, followed by the results and discussion in Chapters 5 and 6, respectively. To foreshadow our results, this exploratory study provided good preliminary evidence that semantic training can have significant effects on the anomia found in PAD, and, with judicious choice of stimuli, this training may potentially generalize to other items in the same category.

CHAPTER 2
REVIEW OF THE LITERATURE 1: NAMING DEFICITS IN PAD

It is common for patients in the early stages of PAD to have picture naming deficits (Appell, Kertesz, & Fisman, 1982; Barker & Lawson, 1968; Bayles & Tomoeda, 1983; Chertkow, Bub, & Seidenberg, 1989; Flicker et al., 1987; Grossman et al., 2004; Kirshner et al., 1984; Lipinska & Backman, 1997; Schwartz et al., 1979; Warrington, 1975; Williams et al., 1989; Williamson et al., 1998). Furthermore, the exact nature of the naming deficit in PAD is controversial (Bell, Chenery, & Ingram, 2001; Nebes, 1992; Nebes, Brady, & Huff, 1989) regarding whether the deficit stemmed from impaired visual-perception, semantic memory, or lexical access (Nebes, 1989). An early theory suggested that visual misperception is the reason for the naming deficit, and this is based on the presence of visuo-perceptual errors, presumably due to difficulties perceiving the object (Barker & Lawson, 1968; Kirshner et al., 1984). These theories and studies providing evidence are discussed below, beginning with the semantic deficit, then the lexical access, and followed by the semantic and lexical access deficit. Next, a discussion of the few naming treatment studies that have targeted this population is provided (Abrahams & Camp, 1993; Fuller et al., 2001; Ousset et al., 2002; Rothi et al., 2005). These PAD naming therapy studies achieved a naming treatment effect (i.e., acquisition) but only one also reported generalization (Abrahams & Camp, 1993). This limited number of studies suggests a need for further research. Insights from all of these studies and additional research from the aphasia literature provided techniques that were applied in our exploratory study comparing the effects of training typical and atypical category exemplars.

Semantic Memory Deficit in Probable Alzheimer Disease (PAD)

Several studies have suggested that there is semantic memory loss or degradation and, consequently, a loss of information about semantic representations (Alathari, Trinh Ngo, & Dopkins, 2004; Hodges & Patterson, 1995; Hodges et al., 1996; Hodges, Salmon, & Butters, 1991, 1992; Huff et al., 1986; Huff et al., 1988; Margolin, Pate, Friedrich, & Elia, 1990; Martin, 1992; Salmon, Butters et al., 1999; Salmon, Heindel, & Lange, 1999; Salmon, Shimamura, Butters, & Smith, 1988; Schwartz et al., 1979). Evidence for this theory includes impaired naming with semantic errors related to the superordinate category or an associate item within the category (Bayles & Tomoeda, 1983; Martin & Fedio, 1983; Salmon, Butters et al., 1999); the presence of consistent naming responses at two different test periods (Henderson, Mack, Freed, Kempler, & Andersen, 1990) an association between the inability to name an item and the inability to recognize its name (Flicker et al., 1987; Huff et al., 1986; Huff et al., 1988); and a relationship between the naming failures and the lack of core information provided about the corresponding item (Hodges et al., 1996). In addition, deterioration of semantic memory in PAD has been described based on poor performance on explicit tasks such as picture naming tasks and category fluency (Bayles & Tomoeda, 1983; Chertkow & Bub, 1990; Flicker et al., 1987; Huff et al., 1986; Martin & Fedio, 1983); tasks requiring generation of semantic feature knowledge (Alathari et al., 2004), generation of verbal definitions (Garrard, Lambon Ralph, Patterson, Pratt, & Hodges, 2005; Hodges et al., 1996); as well as questions targeting feature knowledge (Chertkow et al., 1989; Giffard et al., 2002). Other tasks that have showed similar findings include making judgments about semantic relatedness

(Bayles, Tomoeda, & Cruz, 1999); and associating words, defining words, and ranking associations (Abeysinghe, Bayles, & Trosset, 1990).

There appears to be different interpretations or variants of this theory of semantic memory loss or degradation in PAD. For example, several investigators have found that, while attribute knowledge about a specific concept or exemplar in a category is impaired, superordinate category knowledge is preserved (Chertkow & Bub, 1990; Flicker et al., 1987; Giffard et al., 2002; Giffard et al., 2001; Hodges et al., 1991; Huff et al., 1986; Lukatela, Malloy, Jenkins, & Cohen, 1998; Martin & Fedio, 1983; Nebes et al., 1986; Salmon, Butters et al., 1999; Schwartz et al., 1979; Warrington, 1975). This led some researcher to assert that the attributes of concepts are degraded in PAD (Chertkow et al., 1989; Giffard et al., 2002; Giffard et al., 2001; Martin, 1992; Martin & Fedio, 1983).

Martin (1992) suggested that random damage from the pathological process of PAD resulted in changes to the semantic representations such that they would be degraded and thus more similar to one another. With progression of the disease, the ability to distinguish between items in the same category would be diminished. Thus, on a confrontation naming task, a semantic representation would be activated but it would lack the specificity needed to correctly name the item and, as a result, several lexical entries would be activated. Therefore, when the person with PAD sees a picture, the semantic representation that is activated may be underspecified in terms of the details. This is due to the absence of the essential attributes knowledge for that item such that the activated distinguishing features of that object are not strong enough to rule out other similar items (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997; Martin, 1992).

A property verification task has also been used in a group of adults with PAD (Smith et al., 1995). Participants were asked to verify brief statements containing information about an item's properties. The items were high or low typicality (and dominance, which relates to the relevance of the meaning of the item and is correlated with typicality) category exemplars. The statements were either distinguishing characteristics (i.e., distinctive) or shared features (i.e., common to the other items in that category). The results from the accuracy scores and reaction times indicated a degradation of property level information, particularly regarding both low dominance typicality items in the category. This was not interpreted as a loss of representations of the items' properties or a reorganization of relationships among properties of objects. Instead, Smith et al. (1995) offered this as evidence that the representations of category exemplars that are low-typical and low dominant have been degraded by the Alzheimer pathology. Furthermore, these investigators hypothesized that task demands might be a reason for differences in the literature. Smith et al. (1995) suggested that implicit knowledge allows adults with PAD to have faster verification response times. Thus, in a category verification task, the demands probably do not need the full semantic specification. Smith et al. (1995) also suggested that while explicit knowledge is needed to assess information could be impaired, for example, tasks requiring the participant to use relevant information in a ranking task. This example comes from Grober, Buschke, Kawas, and Fuld (1985) who reported that performance on an attribute ranking task showed that attributes about a concept are preserved, but that the *organization* of semantic information is altered by the disease process. Another example from Smith et al. (1995) is that the naming process involves computation such that an object's properties are activated to allow for distinction among other category members. If Alzheimer disease processes affect both distinguishing and shared features that are low dominant and low typicality, this could interfere with complete activation of the object representation (Smith et al., 1995). Consequently, it would reduce the accuracy on a naming task. Better performance is seen when contextual information is provided, yet when the task involves full semantic representation without the support, performance declines (Smith et al., 1995).

While the evidence above for a progressive deterioration of the semantic system comes largely from tasks requiring explicit access of semantic representations, evidence for an overall preservation of semantic category knowledge comes from tasks that require only implicit knowledge of word semantics.

For example, adults with PAD did not show a priming effect on a stem-completion task and had reduced number of productions for the second semantically related item on a free association task (Salmon et al., 1988). Hyper-priming has been reported on lexical decision tasks and found to be associated with degraded semantic representations (Chertkow et al., 1989; Giffard et al., 2002; Margolin, Pate, & Friedrich, 1996; Martin, 1992). To account for the hyper-priming in PAD, Martin (1992) suggested that changes in semantic activation are more robust for processing information that is degraded, such that a degraded semantic network benefits more from a semantic prime in than an intact semantic network. To account for the difference in performance on tasks requiring implicit versus explicit access to semantic representations, it has been suggested that participants with PAD have difficulty performing an intentional search through semantic memory, but perform relatively normally when relying on the automatic spread of activation in the semantic network (Chertkow & Bub, 1990; Nebes, 1992; Ober, Shenaut, & Reed, 1995).

All of these studies with explicit and implicit tasks suggest that at some level, the semantic system is impaired (due to loss or degradation), and different investigators have used different tasks and methodologies to address this. However, this theory does not go unchallenged; an alternative argument suggests that lexical access, specifically, access of the phonological word form from semantics, is the primary deficit in PAD.

Lexical Access Deficit in PAD

Several researchers have argued that the naming impairment in PAD results from impaired access to the phonological form of the word (i.e., a retrieval deficit) in the presence of an overall intact semantic knowledge structure (Albert & Milberg, 1989; Benson & Geschwind, 1985; Nebes, 1992; Nebes et al., 1986; Nebes et al., 1989; Nebes, Martin, & Horn, 1984; Neils, Brennan, Cole, Boller, & Gerdeman, 1988; Ober & Shenaut, 1999; Thompson-Schill, Gabrieli, & Fleischman, 1999). Along these lines, it has been suggested that the presence of semantic errors is actually an indication that knowledge about the item is intact, despite the lack of the ability to retrieve it, rather than in indication of a semantic impairment (Nebes, 1989).

Semantic priming paradigms (as discussed earlier) have been employed to evaluate the status of semantic memory via an implicit task, in order to reduce the participant's use of attentional mechanisms (Ober, 1999). There are reports of normal priming in adults with PAD (Nebes et al., 1984), as well as reports of hyper-priming which have been interpreted as being caused by abnormal attentional processes in the presence of preserved semantic memory (Hartman, 1991; Ober & Shenaut, 1995; Ober et al., 1995). In a meta-analysis of semantic priming studies, Ober et al. (1995) reported that hyper-priming was due to attentional mechanisms which co-occurred with large increases in reaction times found in controlled priming paradigms. These researchers argued that the hyper-priming represented evidence that automatic spreading activation was occurring among semantic representations in the semantic priming task (Ober et al., 1995). Further evidence for this comes from another study by Ober and colleagues (1995) in which adults with PAD, as well as healthy young and older adults participated in a series of lexical decision tasks. Results showed that priming effects were equal across the groups. The PAD group demonstrated longer reaction times on low frequency words before making lexical decisions, and this was interpreted as an indication that additional time was needed for reaching activation level threshold for these items. Ober et al. (1995) argued that these data reflected an intact semantic memory structure in PAD.

Two event-related potential (ERP) studies offered more support for an access deficit in PAD (Ford et al., 2001; Schwartz et al., 1996). Event related potentials measure electrical manifestations of particular psychological processes that occur in preparation for or in response to discrete events (Fabiani, Gratton, & Coles, 2000; Kutas & Hillyard, 1980). The N400 is a negative going deflection occurring at ~400 ms and occurs in response to anomalous information, specifically, semantic violations; for example it is larger when the prime is unrelated to the target compared to when it is related (Fabiani et al., 2000; Kutas & Hillyard, 1980). Schwartz et al. (1996) compared healthy older adults and adults with PAD to determine whether or not the specificity of a category prime had a differential effect on the degree of semantic priming in these two populations. Using a variation of a category verification task, the investigators asked the participants first to listen to a prime that was delivered auditorily as a name of a category (different levels, e.g. superordinate and subordinate) and then, to read the written presentation of the target name (i.e., a member of a category). While the most robust priming effects were found in the young adults, the smallest priming effect and slowest reaction times were found in the PAD group. Responsiveness to the category manipulation was revealed in priming effects that were similar across the groups, not just in the reaction times but also in the ERP N400 congruity effects being larger on some levels, for example, the subordinate category level. Thus, category level manipulation affected the PAD group, and this was interpreted as evidence for an intact semantic network. In contrast, longer response times and smaller priming effects were attributed to the demands of the task and the necessity of searching through memory while engaged in online processing required for this and other similar tasks (Schwartz et al., 1996).

In a more recent ERP study, Ford et al. (2001) examined priming based on age and dementia, and asked whether the N400 amplitude could be used to show specific semantic memory deficits for objects that could not be named. The participants (adults with PAD, healthy young controls and healthy older adult controls) completed a pretest confrontation picture naming task consisting of items from 12 semantic categories. The following week, the participants completed a picture naming verification task while ERPs were measured. The participants were instructed to press a button to indicate whether or not the prime (a picture) matched the target (a word). The consistent finding across the groups was that for a word that did not match the picture, there was a more negative N400 amplitude. The ability of the PAD group to correctly name pictures was not associated with the N400 priming effect or any corresponding scalp distributions. Thus, despite an inability to access the name of an item, there was evidence from the N400 results that the PAD group had sufficiently intact knowledge for priming responses at the cortical level (Ford et al., 2001).

It has been suggested that while the structure of semantic memory is not damaged in PAD, there is a generalized cognitive processing deficit, which is required for intentionally retrieving and evaluating information, and decision making (Nebes, 1992). This results in word retrieval or access deficits. Two ERP studies (Ford et al., 2001; Schwartz et al., 1996) also provide evidence for a fairly preserved semantic structure in PAD. Additional support for an access deficit in PAD comes from the finding by Ousset et al., (2002) that initial syllable cues and the sound representing the item were among the most effective cues during a oral naming to definition task, while providing the semantic category was the least effective.

Semantic Impairment and Lexical Access Deficit in PAD

The most likely explanation of these disparate findings is that the naming deficit in PAD is due to a combination of a breakdown in both semantic memory *and* retrieval abilities (Bowles, Obler, & Albert, 1987; Huff et al., 1988; Watson, Welsh-Bohmer, Hoffman, Lowe, & Rubin, 1999; Williamson et al., 1998). This has been suggested based on results from confrontation naming and fluency tasks.

Williamson et al. (1998) made predictions about performance of healthy older adults compared to adults with PAD on two confrontation naming tasks: the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983) and the Action Naming Test (Nicholas, Obler, Albert, & Goodglass, 1985), based on knowledge about the underlying anatomy and PAD pathology. More specifically, they predicted that, compared to action naming; object naming would be more impaired in PAD. This was because the disease process had resulted in more extensive damage to the left temporal area which is needed for naming objects, compared to the frontal areas required for naming actions. Compared to the control group, the participants with PAD were less accurate on both measures, but their overall performance was worse on the BNT, on object naming test, compared to the Action Naming Test. Williamson et al. (1998) suggested that object naming revealed significant impairments at the semantic specification and lexical levels. However, for action naming, the impairment was mostly at the semantic specification level, while the lexical level remained fairly intact. The authors also attributed the presence of more No Response errors on objects than on actions to the corresponding damage to the anatomy that underlies that function. Williamson et al (1998) concluded that compared to healthy older control participants, adults with PAD are not as accurate when they name objects and actions. Furthermore, the degree of deficit is more pronounced with objects, which was attributed to impairment not just at the semantic specification level, but also at the level of lexical access. Based on these findings, we designed our study to address both of these levels of deficits. The participants engaged in both lexical and semantic tasks to stimulate the system at both the lexical (e.g., repeating the correct name of the item to access the word form) and semantic levels (e.g., answering questions about semantic features of the item).

In summary, the naming impairment in PAD has been attributed to a loss of or degradation in semantic memory storage and/or organization, disruption of the retrieval of the information in the network, a combination of these, or other factors related to task demands. However, there is reason to believe that the semantic structure, at least in terms of semantic categories, is still fairly preserved in early PAD (Flicker et al., 1987; Huff et al., 1986; Martin & Fedio, 1983; Nebes et al., 1986; Salmon, Butters et al., 1999;

Schwartz et al., 1979; Warrington, 1975). For example, although Cronin-Golomb et al. (1992) reported longer reaction times on a category decision task as well as smaller number of items on a category generation task, the pattern was normal across categories and items were ranked by typicality in a normal manner. This was interpreted as being a reflection of intact semantic category organization (Cronin-Golomb et al., 1992).

Naming Treatment Studies in PAD

Although there is a plethora of research on the naming impairment in PAD, there are only a few naming treatment studies that involved participants with PAD (Abrahams & Camp, 1993; Fuller et al., 2001; Ousset et al., 2002; Rothi et al., 2005). Like the typicality training studies with the aphasia population (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006), these three PAD studies also had picture naming outcomes based on comparisons of performance before and after the intervention (i.e., a treatment effect), and examined generalization to untrained items. Although these three PAD treatment studies are very different from each other, they each have important implications for the future of word finding treatment for this population.

Using a single subject design study with two adults with dementia, Abrahams and Camp (1993) used spaced-retrieval training (SRT). Spaced retrieval training is a technique that intersperses increasing time intervals between presentation of a target by the clinician and recall of it by the participant. Participant 1 was diagnosed with progressive dementia (16/30 on the Mini Mental Status Exam; MMSE) (Folstein, Folstein, & McHugh, 1975) and Participant 2 had PAD (13/30 on the MMSE). Each participant was tested on the Boston Naming Test (Kaplan et al., 1983). From these results, nine training items from each participant were identified for use as a subset of training and control items. While two of these items were treated, the other seven were controls (Abrahams & Camp, 1993). Abrahams and Camp (1993) reported that the results of two different treatment targets for each participant showed improved performance (e.g., naming an item after a maximum of 300 seconds). Only Participant 1 was tested two weeks later. This showed maintenance effects for one of the two items, despite a six point drop in her MMSE score (10/30). Abrahams and Camp (1993) described generalization to untrained exemplars of the trained items: a colored drawing of the target; a second target; and an actual exemplar. The investigators reported that SRT was effective in adults with dementia because they not only have word-finding deficits, but also have cognitive deficits that affect the ability to remember new information. It was suggested that SRT uses procedural memory, which appears to be intact in this population, and thus enabled the participants to benefit from the anomia treatment (Abrahams & Camp, 1993).

Although both participants did not have PAD, the findings from this early anomia treatment study by Abrahams and Camp (1993) have value because they were the first to demonstrate a treatment effect after using SRT, in the PAD population, even with participants with low MMSE scores. Spaced retrieval training does not provide any semantic information about the item, but uses repetition. It was not clear if the items that showed generalization actually generalized. Instead, it could be that it was easier for the participant to name the color drawing.

Ousset et al. (2002) compared the effects of an experimental lexical therapy targeting both episodic and lexical stimulation in adults with mild PAD who were on an acetyl cholinesterase inhibitor. The average MMSE score was 21.1. The participants were divided into two groups of eight: one group received lexical therapy (with narratives and definitions) and the other, a control group, received occupational therapy (i.e., pottery, drawing, and conversations). The treatment was provided in 16 sessions, once a week, with a two week break at the midpoint. The pre and posttest picture naming stimuli consisted of 120 black and white line drawings from three categories. While 80 of these items from the three categories were a part of the lexical therapy, the remaining 40 were not. During each lexical therapy session, the participant first read aloud a narrative from a computer screen and then listened while the examiner read it. Then, the participant completed a naming to definition task on the computer. There were 20 definitions, half of the words were in the narrative to provide both semantic and episodic reinforcement; and half of the words were not in the narrative, to provide episodic reinforcement. When the participant did not respond or was incorrect, the computer randomly provided one of five different types of cues (i.e., semantic category, first syllable, first grapheme, presentation of the item as a color drawing; or the item's associated sound). If this did not result in a correct response, the computer provided the answer. Compared to the control group, Ousset et al. (2002) reported a significant improvement from pretest to posttest for the treated items in the lexical therapy group (i.e., a treatment effect) but generalization to untreated items was not significant. The investigators also commented about the possible benefit of episodic long-term memory reinforcing the association between the object's form and the name of it. Although the narratives (i.e., part of the lexical therapy) were designed to assist in memorization of the lexical labels, the participants were better at retrieving the words that were not in the narratives. Ousset et al. (2002) interpreted this reduced naming as a possible indication that working memory was over-extended by the semantic context in the narratives.

The analyses from the cues revealed that the presentation of the color drawings and the initial syllable were the most effective cues, and Ousset et al. (2002) hypothesized that both of these improved the process of searching the lexicon. The semantic category cue was the least effective. To account for this, the authors suggested that the naming deficit in their patients primarily affected their ability to access the phonological form from semantics. Ousset et al. (2002) questioned if the participants actually had a semantic deficit and used the training as a semantic intervention. Alternatively, it was hypothesized that the participants were a subgroup of the PAD population, such that their anomia might have been due to a retrieval deficit, and the lexical therapy provided rehearsal targeting episodic lexical information and linking it to objects and their corresponding names (Ousset et al., 2002).

This group study by Ousset et al. (2002) clearly showed treatment effects from the training and provided an innovative approach to remediate anomia. Trained words were unrelated to each other. It is unclear why Ousset et al. (2002) did not report generalization to untrained items, when the training items were from three categories. Also, further examination of the results indicated that there was generalization to *another task*. The participants showed better naming performance at the posttest compared to the pretest, but the treatment involved narratives and oral naming to definition, *not* naming pictures. Generalization to another task suggests that the treatment strengthened connections between semantic and phonological representations (Hillis, 1998). As mentioned earlier, the findings from Ousset et al. (2002), which indicated that the category cue was the least helpful on the naming task, could provide further support for the notion that semantic category structure in PAD is still intact. The rationale for the use of the semantic categories as the semantic cue was not addressed. Given the evidence of preserved superordinate semantic category knowledge (Chertkow & Bub, 1990; Flicker et al., 1987; Hodges et al., 1991; Huff et al., 1986; Lukatela et al., 1998; Martin, 1992; Martin & Fedio, 1983; Nebes et al., 1986; Salmon, Butters et al., 1999; Schwartz et al., 1979; Warrington, 1975) but degraded attribute knowledge (Martin, 1992) it is not surprising that the category cue had minimal effects. Perhaps a more potent semantic cue would have been to provide more distinguishing information about the item or possibly a picture of it with some foils. This might have led to better performance on the oral naming to definition task.

A recent single subject design study (Fuller et al., 2001; Rothi et al., 2005) also used an acetyl cholinesterases inhibitor, but with an errorless learning paradigm to improve confrontation naming in six adults with PAD. The scores for the patients on the MMSE were $\geq$ 10. In this study, the stimuli consisted of 100 words, evenly divided for high and low frequency, matched with black and white line drawings from eight semantic categories representing both natural kinds (e.g., animals) and artifact categories (e.g., clothing) with 10 exemplars per category. Three subsets of stimuli based on baseline performance were developed for each participant using three categories, allowing for training on two of these, while the third subset was used both for generalization and experimental control. To establish baseline stability, each participant named the 100 pictures over eight daily probes (i.e., over eight sessions). Sixty minute therapy sessions were provided four times per week, until criterion was met (90%) or after 20 sessions (20-35 sessions were needed to complete the entire protocol) (Fuller et al., 2001; Rothi et al., 2005).

The therapy included two similar treatment phases to determine if the improvements seen in Phase 1 would show response generalization to performance from Phase 2, which at that point would be an untrained set of stimuli. There were two treatment conditions, immediate and delayed repetition. For Phase One, Treatment Condition One involved simultaneous repetition in which the clinician presented the picture, stated the correct name of it and had the participant repeat it. For Treatment Condition Two, delayed repetition was used in which the clinician presented the picture and if the participant knew the name, he/she stated it. If/when the participant did not know the name of the item, he/she informed the clinician. Then, the clinician stated it and had the participant repeat it (Fuller et al., 2001; Rothi et al., 2005). In other words, there was a delay from the time the participant saw the picture until he/she repeated. The results revealed that half of the participants (i.e., three of six) showed a treatment effect. There was no generalization to the untrained items. However, results from the three month maintenance probes showed that the treatment effect was still evident. The investigators also conducted a post hoc review of the participant’s records and identified two other factors that appeared to play a role in the outcomes. The three participants who responded to the treatment both lived at home (versus in an institution) and were not on medications that could affect the learning process and brain plasticity (Fuller et al., 2001; Rothi et al., 2005).

This study (Fuller et al., 2001; Rothi et al., 2005) was the first to demonstrate successful application of errorless learning (i.e., treatment and maintenance effects) in a naming treatment in the PAD population. Furthermore, the influence of both the living status and the medications are important considerations for treatment planning. Perhaps another reason for the treatment effect was the use of semantic categories (both man-made and natural kinds) and their exemplars.

These three studies (Abrahams & Camp, 1993; Fuller et al., 2001; Ousset et al., 2002; Rothi et al., 2005) were theoretically motivated, well-designed, and offer different approaches for improving the naming performance in adults with PAD, including single subject and group design. More specifically, two of these treatment approaches combined cholinergic medication with training items from semantic categories, one in a lexical training with narratives and naming to definition (Ousset et al., 2002) and the other via an errorless learning paradigm (Fuller et al., 2001; Rothi et al., 2005). Generalization to untrained items essentially did not occur in any of these studies. However, the finding that treated behaviors were maintained for one participant two weeks after training (Abrahams & Camp, 1993) and for three participants three months after training (Rothi et al., 2005) is remarkable. This reflects the responsiveness to the training and that the effect was robust despite the MMSE scores. The lack of generalization to untrained items following anomia treatment in the PAD warrants further investigation.

According to Nadeau and Gonzalez Rothi (2004), the connectionist approach views anomia (that is caused by a semantic deficit) as a reflection of insufficiently engaged representations of features that are critical for making distinctions among concepts. When a network is damaged, a large amount of information still remains in the network, so the focus should be on refining the damaged network via semantic therapy. In particular, the network needs to be changed in terms of its connectivity so that there is more reliable engagement of the distinguishing features, while simultaneously there is relatively a disengagement of the shared features (Nadeau & Gonzalez Rothi, 2004). These findings suggest that participants with early PAD might therefore be appropriate for the semantically based training provided in our study which compared typical and atypical category exemplars.

CHAPTER 3
LITERATURE REVIEW 2: A BRIEF OVERVIEW OF THE ORGANIZATION OF SEMANTIC MEMORY

Since this study relied heavily on the predictions of theories about semantic category structure, it is important to provide a brief overview of some semantic theories and categorization. This is followed by a discussion of the two views which can be used to compare training with typical and atypical category exemplars. The family resemblance/prototype (Rosch, 1975) view provides support for training typical items, while support for training atypical items to achieve maximum generalization comes from the Complexity Account of Treatment Efficacy (Thompson et al., 2003), and a connectionist model of a computer simulation of acquired dyslexia (Plaut, 1996). Next, there is a brief description of the four studies that have compared these views in the context of a naming intervention with typical and atypical category exemplars for adults with aphasia (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006). A rationale for applying this training in adults with mild to moderate probable Alzheimer disease (PAD) is also be provided. This leads to a discussion of the current study.

A Brief Overview of Semantic Memory and Categorization

Semantic memory has been described as a hierarchical network consisting of conceptual information and knowledge, storing semantic representations of facts, knowledge for objects and concepts, words and their corresponding meanings, as well as associations (Au & Bowles, 1991; Bayles, Kasniak, & Tomoeda, 1987; Tulving, 1983).

Within the semantic system, the representations are organized in a hierarchical order and distributed throughout the association cortices of the brain (Marshall, 1988; McCarthy & Warrington, 1990).

It is important to briefly describe the networks that allow for this organization at the word level. The spreading activation theory (Collins & Loftus, 1975; Quillian, 1967) suggests that there are at least two levels of representations, semantic, which is organized based on semantic similarity of the meaning of the word, and lexical, which is organized based on the phonemic similarity (and to an extent, orthographic similarity). Furthermore, concepts have been hypothesized to have representations as nodes in a network that have features of the concept that are represented as associated connections to other nodes (Collins & Loftus, 1975; Quillian, 1967). These links are based on the importance or relevance of the information to the concept's meaning such that more important nodes have shorter links or distances between two concepts, and together the nodes make up the network (Collins & Loftus, 1975; Quillian, 1967). The similarity allows for information to spread along pathways in the network (Collins & Loftus, 1975; Quillian, 1967).

Support for the notion of a semantic network and the spreading activation theory (Collins & Loftus, 1975; Quillian, 1967) comes from the semantic priming effect, initially described by Meyer and Schvaneveldt (1971), in which participants had faster reaction times when the prime in a lexical decision task was associated with the target, but were slower when it was not associated. For example, responses to "butter" were faster when preceded by "bread" than when preceded by "window." It was theorized these faster responses (i.e., facilitation) occurred as a result of spreading activation due to shared semantic contexts. The decision was faster because there was an increase in the availability of the target as a result of the prime. More specifically, the memory system is connected to the node for butter, so when one is activated, the activation spreads to connected nodes (Meyer & Schvaneveldt, 1971). According to Martin (1992), semantic priming assumes that there should be automatic activation of an object's representation upon presentation of the object (picture presentation or word), and related concepts should be activated by the spread of activation, which increases accessibility to the name of the object.

Martin (1992) further suggested that these explanations of semantic priming are based on the assumption that the brain is a distributed neural network that instantiates semantic representations. For example, presentation of a picture results in activation of a corresponding neural network, which includes representations of the attributes of the item. Items that are closely related have many shared attributes and overlapping networks that represent them (Martin, 1992). This also explains how the semantic network responds to such information, and when activated above the threshold, semantic processing allows for selection of the target word instead of others (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000). Although there are different theories related to this topic, it is beyond the focus of our study.

The contents of mental representations have been used to account for a range of phenomena such as knowledge of correctly recognizing objects by their name (i.e., word's label) (Medin & Smith, 1984). There is a large body of research on semantic categories and how they are represented (Kiran & Thompson, 2003a). This literature review focuses on the superordinate classification of a category, which is a higher-level grouping allowing for concrete objects to be classified into various divisions (Rosch, 1975).

There are several benefits of categorization, for example, it is both economic and informative. In terms of economy, categories frankly provide a practical and adaptive means for classifying similar items in order to avoid the overwhelming cognitive task of considering all items as if they were unique (Anderson, 1991). Categorizing also allows for assumptions about information that is not explicitly provided, for example, that the item will share similarities with other items in that category (Anderson, 1991; Komatsu, 1992).

Historically, different views have been put forth to account for concepts and categorization. Within some theories of psychology and philosophy the nature of categories is described as Aristotelian (Rosch, 1973a, 1975): They are logical; have clear boundaries; and criteria-based membership. Thus, being a category member requires having a basic set of essential attributes or features such that each item is fully and equally a member. This approach is known as the classical view , which suggests that the representation of categories is achieved by having a set of defining features (Kiran & Thompson, 2003a).

The classical view was challenged by Wittgenstein (1953) who argued that the requirements of formal criteria were not necessary from either a logical or psychological point of view (Rosch & Mervis, 1975). Despite critique, the classical view was researched during the 1960s and considered to be adequate as an appropriate way to describe lexical concepts found in everyday living (Komatsu, 1992). However, by the 1970s, problems with the classical view surfaced, including a lack of evidence for “defining properties” of objects and concepts, identification of categories that did not have clear-cut boundaries, and evidence suggesting that there is inequality among category members, and resulted in its decline (Kiran & Thompson, 2003a; Komatsu, 1992; Medin & Smith, 1984; Posner & Keele, 1968; Rosch, 1973a, 1975). This led to the development of other theories, including the family resemblance/prototype view (Rosch, 1975; Rosch & Mervis, 1975), which has been described as being among the most widely cited psychological concepts (Hampton, 1995) and is discussed in more detail below.

Family Resemblance/Prototype View

Rosch (1973a; Rosch, 1975) argued that for many natural semantic categories, the Aristolean approach does not apply. “Family resemblance” is the term that was applied to highlight the structural principles governing category membership. The notion of family resemblance means that there is an internal structure *to the category*, such that it is organized around a category prototype (Rosch & Mervis, 1975). This allows for a very general approach to the relationships and for application to categories whether they share features with other category members or meet formal criteria for membership in the category. Some of the original work on this topic (Rosch, 1973a) revealed that the prototype of the category is the “best example” of a category and/or “the clearest case” of category membership. The prototype occupied the central location within the category, surrounded by other category members that differed in degree of similarity to the prototype, and, thus, were not equal in terms of category membership (Rosch, 1973a). Rosch (1975) reasoned that there were graded representations of what the category represented signified. Consequently, some items were better representatives of a category. This approach offered an alternative to the classical view, with both its need for clear-cut boundaries for categories and criteria of defining features for category membership. Rosch (Rosch, 1975; Rosch & Mervis, 1975) theorized that categories have an internal structure such that they are organized by a family resemblance to a prototypical category member, and it was the degree and type of resemblance to the category prototype that was the foundation of the category structure.

More specifically, the notion of categories having an internal structure of categories has been tested by asking healthy participants to use a seven-point scale to rate how well a category member fits their image of a given category's name (Rosch, 1975; Rosch & Mervis, 1975). The results showed that some items were more representative than others of that category and these ratings, because they shared features with other items in that category. Thus, they had a family resemblance. These were considered good examples, while others were considered to be poor examples because they did not. These ratings also have been found to be reliable predictors of category verification task performance in which the participants are presented with a statement (e.g., a ___ (i.e., an exemplar) is a ___ (i.e., a category)) and decide if it is true or false. Faster reaction times and higher ratings are found when an item shares features with the category prototype. For example, the exemplar robin is a good example because it is much closer to the prototype, compared to a penguin (a poor example); participants were faster and more accurate verifying that a robin was a bird than that a penguin was a bird (Rips, Shoben, & Smith, 1973; Rosch, 1973a, 1975; Rosch & Mervis, 1975). This is called the typicality effect and additional reports of it are discussed below. It has been suggested that the organization of semantic categories that is based on family resemblance to a prototype is relevant to semantic memory and retrieval (Rips et al., 1973; Rosch, 1973b, 1975).

One of the corollaries of the family resemblance/prototype theory is that the most prototypical members of a category have the most shared attributes with other items in that category and, consequently, represent the central tendency of the category (Rosch, 1975; Rosch & Mervis, 1975). Moreover, as a result of being a prototype defining the category, the prototypical members have the least family resemblance with other categories (Rosch, 1975; Rosch & Mervis, 1975). Furthermore, the more typical items in a category are, the more they share features with the prototype and with each other (e.g., robins, bluejays) while the more atypical items share fewer features with the category prototype and also have more in common with items from other categories (e.g., penguins, ostrichs). Therefore, categories have an internal organization somewhat like a target, such that the prototype is at the center of the bull's-eye, surrounded by typical category exemplars that have many overlapping features with each other and the prototype. Atypical items are more distant from the prototype and share fewer features among themselves and with the prototype (Rosch, 1975; Rosch & Mervis, 1975).

Some of the conclusions from the family resemblance/prototype view (Rosch, 1973a, 1975; Rosch & Mervis, 1975) are applicable to our study. First, there are clear and pervasive findings suggesting that semantic categories have internal structure that influence the processing of words. This is inconsistent with the classical view, but compatible with the spreading activation theory of semantic memory (Collins & Loftus, 1975; Quillian, 1967). Second, faster reaction times to pictures of objects than words, would suggest that the underlying nature of the semantic representation could be closer to pictures (i.e., more visual than verbal) (Rosch, 1973a, 1975; Rosch & Mervis, 1975).

A wide range of studies have investigated the predictions of the family resemblance view. The typicality effect, the finding of better performance (e.g., faster reaction times and better accuracy) for typical compared to atypical category members has been widely reported in non-brain damaged adults (Hampton, 1995; Kiran & Thompson, 2003a; Posner & Keele, 1968; Rips et al., 1973; Rosch, 1973a, 1975; Rosch & Mervis, 1975).

The various tasks tend to involve words, and these are discussed first, for example in category exemplar rating tasks (Rosch, 1975; Uyeda & Mandler, 1980) and priming paradigms involving category verification (Fujihara, Nageishi, Koyama, & Nakajima, 1998; Kiran & Thompson, 2003a; Rosch, 1975; Rosch & Mervis, 1975; Smith, Shoben, & Rips, 1974). Patterns have also been used in a task requiring learning items in a category (Posner & Keele, 1968). A description of the task is provided below, and this is followed by some studies that suggest that the use of pictures might be different from words.

In the priming studies that used the category verification task with words, the superordinate semantic category name was used as the prime, and the target was a typical or atypical category exemplar from that category. Participants were instructed to decide if the target item was a member of that category. The results revealed faster reaction times and more accurate responses for primes that had been rated as good examples, and these were typical items. Longer reaction times and lower accuracy scores were found for poor examples, and these were mostly atypical exemplars (Fujihara et al., 1998; Kiran & Thompson, 2003a; Rosch, 1975; Rosch & Mervis, 1975; Smith et al., 1974).

The typicality studies using categorization tasks offer a comparison of how typical and atypical items are processed in these tasks (Collins & Loftus, 1975). Typical items have distinguishing features that reduce a positive decision on the basis of the superordinate connection (Collins & Loftus, 1975; Quillian, 1967). Also, superodinate connections have different strengths which relates to accessibility, which is dependent upon use (Collins & Loftus, 1975; Quillian, 1967). Furthermore, there is a high correlation between typicality ratings and accessibility (Collins & Loftus, 1975; Quillian, 1967). Thus, the typicality effect provides more evidence that semantic similarity increases the speed of making a positive decision and reduces the time for making a negative decision. In contrast, increased reaction times are found for atypical items, because, based on the overlap of features, they represent only some of the category that they belong to and might overlap with other categories as well (Collins & Loftus, 1975; Kiran & Thompson, 2003a; Quillian, 1967; Rosch, 1975; Rosch & Mervis, 1975).

Further support for the typicality effect comes from a study using event-related potentials (i.e., ERPs) during a category verification task in healthy adults (Fujihara et al., 1998). Event related potentials measure electrical manifestations of particular psychological processes that occur in preparation for or in response to discrete events (Fabiani et al., 2000; Kutas & Hillyard, 1980). The N400 is a negative going deflection occurring at ~400 ms and occurs in response to anomalous information, specifically, semantic violations; for example it is larger when the prime is unrelated to the target compared to when it is related (Fabiani et al., 2000; Kutas & Hillyard, 1980). Fujihara et al. (1998) found that participants had different responses to typical and atypical words. The N400 were more negative after atypical words, than they were after typical words. A typicality effect was found (i.e., fastest and most accurate responses occurred after the typical words). This led to the conclusion that greater priming occurred during typical words compared to atypical words. This was interpreted to be consistent with the hypothesis that the prototype represents the concept of the category, which is the central tendency for each of the category members (Fujihara et al., 1998).

Pictures have also been used as stimuli in category tasks, and according to Snodgrass and McCullough (1986) pictures from a category look similar to each other compared those from different categories, for example, they cite the finding by Rosch (Rosch, 1977, 1978) that based on visual similarity ratings of items from categories of fruit, animals, and vehicles, the ratings were higher for the ones that looked similar to each other. Snodgrass and McCullough (1986) provided more evidence that the time to decide if a picture belongs to a category could be increased or decreased based on the similarity of the items in the category. The investigators suggested that people can categorize a picture and a word via semantic access, but with a picture, people can just use the visual stimulus or the prototype of the category. It was suggested that both visual and semantic strategies are used. Snodgrass and McCullough (1986) thus concluded that picture categorization is not an appropriate task for a comparison of speed of understanding of words and pictures. Although this study did not compare typical and atypical items, it might have referred to them in terms of similar (i.e., typical) and dissimilar (i.e., typical). There is at least one study that provided evidence for the notion that identification of objects first occurs at a level of abstraction that is not the greatest degree of general or specific "entry point", but instead varies based on typicality (Jolicoeur, Gluck, & Kosslyn, 1984). This suggested that identification of typical items occurred at the "basic level" compared to atypical items, which are subordinate to this level (Jolicoeur et al., 1984).

The term representativeness has also been used to refer to typicality such that high representativeness suggests typical items and low representativeness indicates atypical items (Grossman, Robinson, Biassou, White-Devine, & D'Esposito, 1998). Grossman et al. (1998) has reported that pictures were used with healthy older adults and they showed sensitivity to the degree of typicality (i.e., typicality effect). Another study using pictures found that older adults showed typicality effects with both category judgment tasks and exemplar judgment tasks (Cobb, 2005).

All of these studies examining the typicality effect have involved healthy, non-brain damaged adults. The findings have shown that typical items are processed faster and more accurately. Several of these and other studies were primarily focused on adults with brain damage, but it was important to reveal the pattern in the normal population prior to addressing the patient groups. These include adults with probable Alzheimer disease and other research with adults with aphasia. Some examples are discussed next.

Some studies have found typicality effects in adults with probable Alzheimer disease (PAD) using category verification tasks (Cobb, 2005; Grossman et al., 1998). Other variations of this task, for example ranking attributes or ordering the representativeness of category exemplars (Cronin-Golomb et al., 1992) and some also included rankings of dominance, which relates to the significance of the feature to the meaning of the item (Nebes et al., 1986; Nebes & Brady, 1990). In a category fluency task adults with PAD have shown normal variations in the typicality of the items produced (Ober, Dronkers, Koss, Delis, & Friedland, 1986). Grossman et al. (1998) found normal typicality effects on a category judgment task, but this was only in the subgroup of PAD participants who did not have a semantic impairment, but had naming deficits.

In contrast, Grossman et al. (1998) reported that the other subgroup of their PAD participants had a semantic deficit, demonstrated naming deficits and insensitivity to the representativeness. These participants did not differentiate among the levels (low, moderate, and high) and this was attributed to compromised ability to make similarity judgments (Grossman, 1981; 1998). Other PAD studies have shown no typicality effects or otherwise differences in processing atypical and typical item in this population, for example impaired performance on ranking attributes (Grober et al., 1985). Sailor et al. (2004) compared the responses of healthy older adults and PAD participants on a category generation task. The analyses revealed that compared to healthy older adults, the PAD group produced significantly fewer atypical items. This was the trend across the categories of fruit, vegetables, animals, and footwear. It was concluded that the PAD group had a slower search rate through semantic memory compared to the healthy older adults (Sailor et al., 2004). Also, as mentioned in the previous chapter, a property verification task revealed worse performance on category generation low typical items (Smith et al., 1995).

Although most of these studies indicated that the PAD population has deficits on atypical items, to the best of our knowledge, there is at least one study that showed better performance on atypical items (Cobb, 2005). The purpose of this study by Cobb (2005) was to attempt to understand if typical items were affected by the pathology of Alzheimer disease. Using the bottom-up deterioration hypothesis in PAD, Cobb (2005) theorized that the vulnerability of typical items compared to atypical items could be described as:

Generalized (equal); Strength (not as vulnerable); Distinctiveness (more vulnerable). Category judgment tasks and exemplar judgment tasks were used in which pictures were shown and the participants verified them. The results indicated that like the healthy older adults, the PAD group showed a typicality effect on the category judgments (i.e., faster and more accurate responses were found on the typical items). In contrast, on the other task of verifying the exemplars, the PAD group showed deficits on the *typical items* compared to healthy older adults. Cobb (2005) suggested that these findings provide some evidence that the typical items were more vulnerable to the deterioration, and thus provided support for the distinctiveness model. However, Cobb (2005) also cautioned about generalizing the results because of this preliminary study due to some issues related to methodology.

Research on the typicality effect has also been conducted in the aphasia population. In a category generation task, Grossman (1981) compared the number and typicality of the exemplars generated (i.e., typical and atypical) between a group of participants with nonfluent aphasia and another group with fluent aphasia. The majority of the items produced by the nonfluent group were typical items. Grossman (1981) suggested that this occurred because these participants used the central tendency of the category as a guide and compared the target item to it. In contrast, the fluent aphasia group produced a more even distribution of typical and atypical exemplars, starting out with more typical items (like the nonfluent group), but then producing items that were as close to the central part of the category and even from other categories. Moreover, participants with fluent aphasia actually produced many responses that were not members of the specified category. To account for these findings, Grossman (1981) suggested that these participants were aware of the central tendency of the category but had reduced mental representations related to the category's border. Thus, they violated category boundaries(Grossman et al., 1998). Grober et al. (1980) used a category verification task in two modalities (i.e., pictures and words) to determine if performance on these tasks could distinguish groups of patients with anterior (i.e., nonfluent) aphasia and posterior (i.e., fluent) aphasia. There were no differences between performance with pictures and words for any of the patient groups. Furthermore, across groups, the typicality effect was realized as both faster response times and better accuracy for the typical exemplars. However, while the nonfluent patients were fairly accurate in their classifications of atypical category exemplars, this was not the case for the fluent patients. When the item clearly belonged to a given category, the fluent group was accurate, but when the item was at the category boundary, they were less accurate. Grober et al. (1980) suggested this difficulty with atypical items was due to a disruption of the underlying semantic structure of semantic categories in fluent aphasia. Although this task was different from the category generation task used by Grossman (1981), the findings for the patients with fluent aphasia are compatible.

In a category verification study, Kiran and Thompson (2003a) compared the performance of two groups of adults with fluent and nonfluent aphasia to young and older adults without brain damage. Typicality effects were expected to occur in all of the groups. However, based on previous research (Grober et al., 1985; Grossman, 1981), the pattern of activation for the fluent aphasia group was expected to be different. Kiran and Thompson (2003a) used words from the categories of vegetables, birds, and fish typicality ratings that they developed based on the seven-point scale used by Rosch

(1975). Participants first read a category name and decided if the next word was a member of that category or not. Furthermore, the overall results were consistent with the predictions: typicality effects were found in all the groups, and the fluent group displayed a different pattern of performance from the other groups. The fluent group had the smallest effect of typicality and the greatest number of errors. Although this group was more accurate on the typical items compared to the atypical items, their reaction times for accepting the correct typical items were not significantly faster than the correct atypical items. Kiran and Thompson (2003a) interpreted this as two different manifestations of a semantic impairment from the fluent aphasia. The first was that weakened boundaries of the target category led to increased numbers of errors. The second difficulty was that there was a decrease in the ability to access the prototype and this caused a deficit when comparing the target with the prototype of the category. They concluded that the findings of longer reaction times and more errors with atypical exemplars suggested impoverished associations linking atypical exemplars with their category's prototype. Kiran and Thompson (2003a) theorized that these findings could offer more information about the nature of atypical category exemplars as having differential representation in semantic memory. Furthermore, the greater number of dissimilarities among atypical items in a category allows them to convey a broader range of variation across the category, and thus they were considered to be "more complex" (Kiran & Thompson, 2003a).

Thus, in the aphasia literature, there is at least one study suggesting that fluent aphasia affects the processing of atypical category exemplars in a category generation task (Grossman, 1981). This is understandable given that category generation requires searching semantic memory for an item within a category when only provided with the superordinate category label. Some category verification task studies have shown typicality effects in adults with nonfluent aphasia (Grober et al., 1980; Kiran & Thompson, 2003a; Stanczak et al., 2005), including transcortical motor aphasia (Stanczak et al., 2006), and fluent aphasia (Stanczak et al., 2005), including conduction aphasia (Stanczak et al., 2006). In others studies, the typicality effect was found but was somewhat disrupted in adults with fluent aphasia (Grober et al., 1980; Grossman, 1981; Kiran & Thompson, 2003a), as shown by more difficulty (e.g., slower or less accurate) on atypical items. Although the typicality effect indicated better performance with typical items, at least one study did not find that typical items were verified faster than atypical items that were correct (i.e., member of the given category) by the adults with fluent aphasia (Kiran & Thompson, 2003a).

In summary, the family resemblance/prototype view (Rosch, 1975; Rosch & Mervis, 1975) has had a strong influence on natural language research in category representation (Barr & Caplan, 1987). Based on the family resemblance/prototype hypothesis, there are several reasons for training typical category exemplars and they could have an advantage over training atypical category exemplars. Typical items have richer representations, and people have greater knowledge of them so they are used more, and therefore accessed faster (Heilman, 2005, personal communication). Evidence from adults with brain damage suggests that knowledge about atypical items is disrupted following strokes to the posterior portion of the left hemisphere (Grober et al., 1985; Grossman, 1981; Kiran & Thompson, 2003a), and also as a result of PAD (Sailor et al., 2004; Smith et al., 1995).

The Complexity Account of Treatment Efficacy

As described above, atypical items have been hypothesized to be more complex (Kiran & Thompson, 2003a). The Complexity Account of Treatment Efficacy (CATE) was developed by Thompson et al. (2003) based on their study in which adults with agrammatic aphasia were trained to use sentences that were syntactically complex. As a result, participants showed generalization to less complex sentences whose structures shared processes with trained items. Several studies by Thompson and colleagues (Thompson et al., 1998; Thompson et al., 1997; Thompson et al., 2003) have reported that adults with aphasia have shown generalization following treatments in which complex syntactic structures were trained. This was because when more complex items were trained, this included variables that are relevant to items that are simpler. This promoted greater degrees of access to untrained items compared to the training of simple items (Thompson et al., 1998; Thompson et al., 1997; Thompson et al., 2003). Thompson et al. (2003) concluded that these findings were compatible with other studies in the literature and suggested that enhancement of performance is found under the following two conditions: when there is related linguistic structure and when treatment moves from greater degrees of complexity to lesser degrees of complexity. In addition to acquired language disorders, complexity has been used by Gierut (2001) as a guide for stimuli to promote generalization with children who have phonological disorders. The findings also have been positive and indicate that with greater degrees of target complexity, there are greater degrees of gains in phonological skills; these behaviors have been shown to generalize to untreated sounds (Gierut, 2001). These principles are borrowed from motor skill learning in adults and conditions under which the behavior is practiced (Schmidt & Lee, 1999). Generalization to untrained items in a picture naming study has also been found to be successful when the focus has been on atypical items as a means for stimuli selection because they have a hierarchical complexity (Kiran & Thompson, 2003b). This and other behavioral studies are discussed below. However, it is important to first understand the initial work in this area by Plaut (1996) in which he compared the training of atypical and typical items in a connectionist model of acquired dyslexia.

Connectionist Model and Computer Simulations

Prior to describing the computer simulations by Plaut (1996), a brief overview of the connectionist or parallel distributed processing (PDP) model is provided. The PDP is a computational framework that allows for exploration of cognitive processes of normal and damaged systems. Simulations are conducted to learn more about the processes and to contribute to the development of potential rehabilitation programs (Plaut, 1996). Parallel processing is the term used to describe how individual processing units (similar to neurons), in a large array, are connected, but each performs a basic function and represents a component of the modeled entity (Nadeau, 2000; Plaut, 1996; Rumelhart, McClelland, & PDP Research Group, 1986). The units are linked by connections and the strength of these connections reflects the knowledge of the network (Nadeau, 2000; Nadeau & Gonzalez Rothi, 2004). Large patterns of activation, based on excitatory and inhibitory signals, represent a concept which is referred to as a distributed representation (Nadeau, 2000; Plaut, 1996; Rumelhart et al., 1986).

Plaut (1996) tested the family resemblance theory using a connectionist computer model in which he simulated acquired dyslexia in three experiments. Only the first two are discussed here. First, Plaut (1996) trained the model, then lesioned it to simulate the effects of stroke. Finally, he retrained the model to simulate therapy after stroke in order to better understand the degree and speed of recovery; how generalization occurs from treated to untreated items, and the way this generalization might be maximized via selection of the treatment items.

In Plaut's (1996) first experiment, he programmed the network and trained it to generate semantic information of about 40 words from five semantic categories (natural kinds and man-made) based on orthographic input. The training was done to allow mapping from the written words to their corresponding meanings. Then, Plaut (1996) damaged the network by random selection and removal of some of the units representing semantics and others representing orthography. Plaut (1996) randomly identified items for retraining: half of the words that were trained correctly and half that were not trained correctly. The remaining words were not retrained. Plaut (1996) reported that, compared to initial learning, retraining was faster because there was a re-establishment of consistency that was relevant among words and thus untreated words improved. Plaut (1996) suggested that location of the damage was important. When the semantic level was damaged, there was fast relearning and generalization was substantial. However, when the orthography level sustained damage, relearning was slower and generalization did not occur. Thus, there was better generalization at the semantic level because the words had relationships based on their meanings. This was based on the structure of the semantic organization of the set of words. Plaut (1996) hypothesized that the part of the system that was damaged could influence recovery. Furthermore, Plaut (1996) suggested that generalization had improved because the semantic categories from which the words were derived had a substantial amount of overlap in their semantic representation. This experiment provided support for using items from semantic categories in therapy because the consistency inherent in the items in the category allowed for efficient processing and generalization. These findings also offer support for the family resemblance/prototype hypothesis (Rosch, 1975; Rosch & Mervis, 1975) for achieving a treatment effect. However, in terms of maximizing generalization, Plaut (1996) provided evidence for the benefit of training atypical items, and this was discovered in Experiment 2.

Plaut (1996) theorized that generalization could be obtained if the nature of the semantic representation of the items was a good estimate of semantic structure relevant to the whole set of words in the training stimuli. Selection of words for training were based on the word's semantic representations (Plaut, 1996). Acknowledging that nouns are categorically organized, Plaut (1996) also used typicality, a semantic variable that is essential because it reflects the proximity to the central tendency of a category (Rosch, 1975). Plaut (1996) hypothesized that the extent of generalization from retraining is affected by the relative typicality of the treated words (i.e., typical versus atypical).

The training items that Plaut (1996) chose were 100 artificial "words" from one semantic category. These words were evenly divided into two sets in which the meaning was generated by varying the degree of featural overlap between a prototype item and other items in the category. Thus, the semantic features for the typical words were closer to the prototype of the category; whereas the features for the atypical words were further from the prototype. Training the network involved iterations that allowed for generation of the correct semantic representation for each word when the corresponding orthographic form was presented. Then, to lesion the network, Plaut (1996) randomly selected and deleted connections. Next he retrained 25 of the typical words and 25 atypical words in different networks. This allowed for evaluation of the impact of typicality on generalization to untrained typical items and untrained atypical items. The results were consistent with Plaut's (1996) predictions. The lesion rendered typical words more impaired compared to atypical words. Thus, Plaut (1996) suggested that typical words were more vulnerable to damage because they had more competitors that were close, due to the overlap. Thus, distinguishing typical words from other typical words (compared to distinguishing atypical words) required better accuracy by the network. In contrast, there were not as many competitors for atypical words, and distinguishing among atypical words was not as difficult as it was with typical words. However, as a result of retraining, trained typical words had significantly better performance compared to trained atypical words (Plaut, 1996).

Although these findings are relevant to recovery, Plaut (1996) argued that the more *critical* finding is the fate of the untreated words, because this related to choosing stimuli that could influence generalization, an important component of rehabilitation that clinicians could control. When the training was with *either* typical or atypical words, there was overall substantial improvement in the untreated typical words. However, atypical words only improved when the network was trained with atypical words. When atypical words were retrained there was more generalization to both untrained atypical and typical words compared to when typical words were trained. In contrast, when the network was trained with typical words, this had a negative impact on performance of untreated atypical words. To account for these results, Plaut (1996) hypothesized that typical words just focus on the *central tendency* of its category, without indicating the semantic features that can differ from the prototype Thus, the typical words offer a good estimate about the degree of variation in the semantic features that is shared by other typical words, but not the information about the variation for the atypical words. In contrast, atypical words indicate the *amount of variation* that can exist across members of a category while also approximating the category's central tendency. Thus, Plaut (1996) hypothesized that atypical words collectively estimate the semantic representations that span more of the features required by the full set of words compared to representations of typical words. Consequently, training atypical words should result in generalization to typical words, because the semantic representations of the atypical words included information about the degree of variability within the category structure as well as information about the central tendency of the category (Plaut, 1996). While Plaut (1996) recommended using this approach, he cautioned that the atypical items should not be extremely atypical. The semantic dimensions should be relevant and thus allow for the average effects to be near the category's central tendency. Plaut (1996) recommended that future research endeavors should involve training atypical items from several categories at the same time.

In summary, the PDP model has been described as operating in a manner consistent with the brain (Nadeau, 2000; Plaut, 1996; Rumelhart et al., 1986). Overall findings from Plaut (1996) provided a rationale for training items within a semantic category. Also, while training typical items resulted in better performance (i.e., acquisition or a training effect), training atypical yielded better generalization to untrained items. The results from Plaut (1996) are very positive and have great implications for clinical application. In particular, both category structure and atypical items are important in promoting generalization because they relate to the semantic representations of the items. However, it is important to remember that computer simulations involve isolated information and events, in which the experimenter programs all of the inputs and thus has complete control over the information in the model. For example, the second experiment used artificial words from one category and the typical and atypical items were created by the network. This differs from the wide range of variation in humans in terms of semantic memory, learning, and the effects of brain damage and behavioral manifestations of different disorders. Nevertheless, as discussed below, studies with humans have been conducted and some of the findings are consistent with Plaut (1996).

Similar Behavioral Studies

A series of experiments on category learning in non-brain damaged college students was conducted by Posner and Keele (1968). Participants looked at patterns with varying degrees of distortions. These distortions ranged from high (i.e., the distance from the central tendency of the category was greater) to low (i.e., closer to the central tendency). The participants pressed a button to specify their selection for the given pattern, and received feedback about accuracy. This was repeated over several trials to allow for learning. Results showed that the participants had better generalization from learned patterns to new ones after training with highly variable patterns compared to when the training emphasized low variability. To account for this, Posner and Keele (1968) suggested that variability is an important part of the learning process. Although this study did not involve naming or a patient population, and did not use the term typical and atypical, the patterns with higher variability might be considered atypical, whereas those with much lower variability might be considered typical. Furthermore, the theoretical underpinnings are along the same lines, and can be considered support for training atypical items (Plaut, 1996).

More recently, three different groups of investigators (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006) examined and compared the effects of training typical and atypical category exemplars, and thus compared the two opposing views of typicality discussed above (Plaut, 1996; Rosch, 1975; Rosch & Mervis, 1975). Each of these therapy studies involved adults with aphasia in the context of a single-subject design targeting confrontation picture naming, and showed consistent evidence of treatment effects. However, the results were mixed, especially in terms of generalization. A brief description of these three studies is provided below. This is followed by a brief critique pertaining to both their strengths and limitations, highlighting how they offer unique contributions to the literature and have implications for the current study.

Kiran and Thompson (2003b) predicted that when atypical items were trained, this process would result in emphasizing both how much the features in the category vary *as well as* the prototype's features. This, in turn, would result in a strengthening of the associations among the typical items, increased access to these typical items, and finally generalization to untrained typical items. In contrast, they argued that training typical category exemplars would have more limited effects, because typical items have limited variation, focusing only on the central tendency of the category (Plaut, 1996). Kiran and Thompson (2003b) argued that, compared to typical items, the atypical items were more complex because their semantic features have a greater degree of dissimilarity with each other, and thus they offer more diversified information about the whole category, as a collective unit (Plaut, 1996; Thompson et al., 1998; Thompson et al., 2003).

To test these predictions, Kiran and Thompson (2003b) provided therapy for four adults with fluent aphasia on a naming task. Baseline naming performance was assessed using 48 items consisting of 24 exemplars from each of two semantic categories (birds and vegetables). The study design allowed for each participant to be separately trained on one category at a time, starting with either eight atypical category exemplars or eight typical items from each category. After the participant met the criteria the next category was trained with the opposite typicality. The participants attended two hour therapy sessions twice a week until either criteria of 7/8 correct was achieved over two consecutive sessions or after a maximum of 20 training sessions had been completed (Kiran & Thompson, 2003b). During training, participants performed a variety of semantic tasks tapping both comprehension and production skills: picture naming, category sorting, feature selection, answering yes/no questions; and picture naming again.

Participant 1 was first trained on typical birds for seven weeks and this led to a treatment effect (i.e., acquisition of the trained items) but there was no generalization to untrained intermediate or atypical items. Subsequent training of atypical birds resulted in improved performance on those items. At that time, there were no changes on performance of the control category (vegetables). Then Participant 1 was trained on atypical vegetables, and in 8 weeks, acquisition criterion was observed as was generalization to untrained intermediate and typical birds. Participant 2 was first trained with atypical birds and in 11 weeks, met criteria and showed generalized naming on intermediate and typical items. At that time, performance on the vegetable category remained the same. Treatment of birds led to no change in performance on the untreated category, vegetables, so the participant was subsequently trained on atypical vegetables. This resulted in a treatment effect and generalization to untrained intermediate and typical items. Participant 3 was first trained with typical vegetables and although criteria was met, there was no generalization to intermediate or atypical vegetables. Next, Participant 3 was trained on atypical vegetables that resulted in improved performance on the trained atypical vegetables; no further treatment was provided with the bird category because of the extended time for training on the first (28 weeks). Participant 4 was trained only on atypical items and showed a pattern similar to Participant 2: a treatment effect after 6 weeks of training on vegetables and after nine weeks on birds, and generalization to both untrained intermediate and atypical items.

The patients in the Kiran and Thompson (2003b) study acquired atypical items faster acquisition than typical items. Furthermore, naming of untrained typical and intermediate items improved following training of atypical items, thus showing generalization. However, when the typical items were trained, there was no generalization to exemplars that were atypical or of intermediate typicality. Kiran and Thompson (2003b) suggested that the training for the atypical items emphasized how the semantic features in the category varied; whereas the training for the typical items was limited to a small number of features that were shared among the other typical items. Moreover, maintenance of the treatment effects was found 6-10 weeks after the training.

Kiran and Thompson (2003b) conducted an error analysis to gain some insights into the effects of the treatment. Before the intervention, the errors were primarily general, as characterized by superordinate labels, no responses, and neologisms. This was attributed to the participants failing to retrieve the specified semantic and/or phonological details of the target name. However, after the treatment, the participants showed more accurate naming on trained items and on untrained items, reflecting better semantic and phonological access to the representation. The shift in the type of errors, from general at baseline to semantic and phonemic after the training was attributed to a greater impact of excitation at the level of semantics and phonology that had occurred when the participants attempted to name the items. However, this interfered with the ability to accurately select the target name from others that were also activated. Kiran and Thompson (2003b) suggested that although the treatment had a positive impact that was illustrated in enhanced spread of activation to targets within a category that were related, the intervention was not completely successful in eliminating the interference caused by many category exemplars being activated at the semantic/phonemic level during naming. Kiran and Thompson (2003b) suggested that there was improved performance because the treatment was based on semantics and highlighted the underlying elements of semantic representations and processing. The investigators suggested that this was compatible with the literature using the semantic feature approach (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000; Drew & Thompson, 1999). Kiran and Thompson (2003b) concluded that their findings provided more support for improving generalization to untrained items based on using more complex training items, and thus providing more evidence for the Complexity Approach to Treatment Efficacy (Thompson et al., 1998; Thompson et al., 2003) as well as for training atypical items to achieve better generalization (Plaut, 1996).

In this study, Kiran and Thompson (2003b) were the first to show that the computer simulations of acquired dyslexia by Plaut (1996) could be successfully applied to adults with fluent aphasia. The findings illustrated the value of training atypical category exemplars in the context of a training highlighting semantic category structure. This study was well designed, including the development of typicality ratings and semantic attributes, intentionally selecting low frequency items to avoid effects of word frequency, the use of both comprehension and production tasks to activate the entire word representation, and a training regimen that focused on just one category at a time. Two of the training tasks, feature selection and yes/no questions provided comprehensive information about the items, including the written form. However, specific information was provided by the clinician, and the participant had to select it. It is possible that the training might have been even more robust if the participants were required to generate the information. More importantly, there appears to be a lack of generalization probes for untrained atypical items and untrained typical items that have the same and the opposite typicality of the items that were trained (i.e., trained atypical and trained typical items). Our study included a measurement of untrained atypical and typical items for the same and opposite typicality of the trained items (i.e., that were trained with atypical and typical items; e.g., Training on typical items and untrained typical items), and a more active role for the participants on the training tasks. These are discussed later.

Mayer et al. (2004) attempted to replicate the findings of Kiran and Thompson (2003b) in three participants with severe aphasia and to determine if other semantic categories might yield potent typicality effects. Each participant had a different type of aphasia. Participant 1 had global aphasia. Participant 2 was considered borderline fluent, (i.e., fairly intact auditory comprehension and reading ability, with fluent speech and neologisms including some word approximations). Participant 3 had mixed nonfluent aphasia (i.e., intact auditory comprehension for basic information, and although nonverbal, he spontaneously gestured and used facial expressions and drawings).

A total of six categories were used by Mayer et al. (2004): animals, clothing, furniture, sports, tools, and vehicles. Each participant was trained separately on 10 atypical items from one category, and then on 10 typical items from another category.

Each category was treated for nine weeks, for a total of 18 weeks. Two other categories were used as treatment probes (before and after treatment) and generalization probes. To maximize support in a structured way using phonological and semantic stimulation, training utilized a response-contingent hierarchy (Bandur & Shewan, 2001). The clinician presented a picture and guided the participant through a semantic and phonological cueing hierarchy (starting with the least effective and moving to the most effective). This was used until the participant named the item. Then, the participant was asked to repeat the correct name of the item five times (Pring, Hamilton, Harwood, & MacBride, 1993), before completing a confrontation naming task. The results indicated that there was a treatment effect for each participant, but this did not generalize to any untrained items that were probed. Mayer et al (2004) reported that one participant initially showed progress and typicality effects that were described as subtle: generalized naming to untrained typical sports items after training on atypical sports items. However, because there were unanticipated improvements in not just in trained categories but also in untrained categories, Mayer et al. (2004) indicated that this could not be considered generalization.

Mayer et al. (2004) identified three factors to explain the lack of generalization. First, they considered it possible that the severity of the aphasia might have interfered with the cognitive processes needed for rehabilitation. Second, Mayer et al. (2004) suggested that the training tasks were structured to compensate for the severity of the deficits, and, thus, perhaps did not have the metalinguistic requirements of the tasks used by Kiran and Thompson (2003b). Third, Mayer et al. (2004) discussed Plaut's (1996) assumption pertaining to the retraining and the mapping; this might not have been applicable to their participants. According to Plaut (1996) generalization might require a good estimate of the relevant semantic information of the training set overall and this needs to be in alignment with the consistency of the mapping that the damaged part of the network is responsible for carrying out. Mayer et al. (2004) reasoned that the severity of the brain damage in their participants negatively affected the integrity of their lexical-semantic network, rendering it incapable of retrieving the semantic features in order to demonstrate generalization.

This study by Mayer et al. (2004) was carefully designed and included various nuances that are important in therapy, including development of typicality norms, selecting six categories that were based on participants' interests when possible. In addition they included repetition of the target as a training task, to minimize phonological access deficits.

Stanczak et al. (2005) also sought to understand the relationship among performance on a category verification task (i.e., a measure of online typicality), training typical and atypical category exemplars for naming production, and generalization. Two participants with anomic aphasia, one fluent and the other nonfluent participated in the study. First, prior to the baseline picture naming probes, the participants completed a category verification task (with stimuli from the Kiran and Thompson (2003b) study). The purpose of this was to determine if they would show typicality effects (i.e., faster reaction times and more accurate responses on typical versus atypical items) in response to the written category primes and their written exemplar targets. Then the participants completed baseline probes. This was followed by a naming therapy, in which each participant was simultaneously trained on two categories, birds and vegetables, one with typical category exemplars and the other with atypical. The training tasks were naming the target picture, verification of semantic attributes related to it, and naming the picture again. The results of the category verification task that was completed at the onset of the study indicated that the participants were faster to verify typical items compared to atypical items (i.e., this was described as “online typicality effects”). The naming treatment had differential effects. The participant with fluent aphasia only demonstrated acquisition of trained typical birds, but not atypical vegetables. In contrast, the participant with nonfluent aphasia demonstrated treatment effects for the atypical birds only (not typical vegetables). There was no generalization to untrained items in either participant. The finding of a typicality effect on the category verification task was interpreted by Stanczak et al. (2005) as activation occurring in a bottom-up fashion (i.e., the semantic retrieval was automatic). The differences in these two participants’ responses to the typical and atypical training was attributed to their lesion site and fluency (anterior and nonfluent versus posterior and fluent) interacting with the semantic processes involved with typical versus atypical items. Specifically, it was hypothesized that the controlled semantic processing in the fluent participant was relatively intact, allowing for better performance with the typical words, despite their similarity with other items and higher degrees of semantic interference. However, learning the atypical items was considered to be more difficult. In contrast, the nonfluent participant, Stanczak et al. (2005), theorized that the demands for atypical items were less than the typical items (i.e., fewer competitors), and thus yielded better performance. This study by Stanczak et al. (2005) provided an intriguing interpretation of the results by linking them to neural substrates.

This is feasible. However, one limitation was that the investigators did not elaborate about the difficulty related to atypical words.

Very recently, Stanczak, Waters, and Caplan (2006) further extended this research by applying the same protocol (Stanczak et al., 2005) to two different participants with anomia from different etiologies. The purpose of this study was to determine if generalization would be affected by the type of deficit and the type of treatment provided. Participant 1 had conduction aphasia, and her deficits were primarily phonological; her training was with atypical vegetables and typical birds. These were trained over a 20 week period. Participant 2 had transcortical motor aphasia and had phonological and semantic impairments; his training was on atypical birds and typical vegetables. The results indicated that compared to atypical items, typical items were learned faster by both participants. Although both participants showed significant treatment effects and better learning of birds compared to vegetables, the patterns of responsiveness to the training varied for each participant. Participant 1, who had a primary phonological impairment, displayed greater learning when trained with typical items compared to atypical items. In contrast, Participant 2, who exhibited both semantic and phonological deficits, did not initially show a difference between learning typical versus atypical items, but did show better learning of atypical items as the treatment continued. Participant 2, who had been trained on atypical birds and typical vegetables, also had statistically significant generalization to untrained typical birds, and only marginal generalization to atypical vegetables.

To account for these different findings, Stanczak et al. (2006) offered interpretations which again were based on neural substrates. It was not the phonological impairment in Participant 1 that resulted in the lack of learning atypical items, but the greater degree of semantic difficulty inherent in atypical items. However, for participants with damage to the left prefrontal area, as in Participant 2 (with both semantic and phonological impairments), the typical items initially should be learned better because of the high degree of similarity in the features but, then over time, this overlap leads to confusion. Stanczak et al. (2006) hypothesized that learning atypical items, which have greater variation, should result in better performance because the degree of competition is less. Despite the better performance on birds, the investigators refuted the possibility that this was a category specific learning (i.e., that it was easier to learn birds compared to vegetables), and provided a rationale for the effects of typicality on learning (i.e. both participants were responsive to being trained with typical items but only one responded to the training with atypical items). Stanczak et al. (2006) concluded that while their findings were broadly consistent with Plaut (1996) and Thompson et al. (2003), the differences from findings by Kiran and Thompson (2003b) might be attributable to differences in the training (i.e., training two categories at that same time) or the participants (degree of semantic impairment).

These four studies (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006) were theoretically motivated and well planned, including programming for generalization. The training tasks included at least some information about the target item in terms of semantics and the phonological form, as well as confrontation naming. Some potential limitations of these studies, with the exception of Mayer et al. (2004) included the use of only two animate categories without a third control category (although the second category was used as a control) and a lack of information about the clinician's response to a failed naming attempt by the participant (Kiran & Thompson, 2003b; Stanczak et al., 2005, 2006). The clinician's response to the participant's naming failure is important because it could affect the outcome. This is because it provides information about the response so that the patient can better understand the information and can learn to self-evaluate (Bandur & Shewan, 2001).

Also, in some of the studies it was not clear if generalization probes to untrained items that had the same, different, or both types of typicality as the items that were trained (e.g., train atypical items, probe untrained typical items in the trained category). It appears that the generalization probes were not measured on both in some of the studies.

Other broader methodological concerns about these four studies include the lack of opportunity to or ability to generate semantic information about the item in a more active manner (i.e., beyond yes/no questions). Although yes/no tasks, sorting and other similar tasks are used in the literature and have value, they require the participant to play a less active role in the therapy. Mayer et al. (2004) and Stanczak et al. (2006) identified concerns related to this in terms of the severity of their patient population and how it might have affected generalization. Another possibility that Mayer et al. (2004) identified was that the semantic category structure in their patients was impaired such that they were unable to benefit from the training which was built around the semantic representations of the category (Plaut, 1996). Another important consideration about achieving a generalization effect has been raised by Murray and Clark (2005): it is not known if it is the combination of training typical and atypical category exemplars in the context of semantic training requiring task requiring processing of semantic representations (Kiran & Thompson, 2003b), or simply the use of atypical and typical category exemplars. We attempted to address some of these issues in the methodology of the current study.

The Current Study

Rationale for Applying the Training to the PAD Population

Brain damage can cause a language impairment, and object naming is a sensitive marker for this (Benson, 1979). There is some evidence that the semantic system is preserved in adults with nonfluent and fluent aphasia based on similar performance on a lexical decision task compared to non-brain damaged adults (Gerratt & Jones, 1987). However, several of the studies discussed above examining adults with aphasia suggested there was a disruption or disorganization of the semantic system in adults with fluent aphasia (Grober et al., 1985; Grossman, 1981; Kiran & Thompson, 2003a). There are other reports of a semantic level impairment in adults with aphasia based on impaired performance on property verification and category verification tasks (Koemeda-Lutz, Cohen, & Meier, 1987); and lexical-semantic discrimination tasks (Chieffi, Carlomagno, Silveri, & Gainotti, 1989; Gainotti, 1981). More specifically, in transcortical sensory aphasia and anomic aphasia, disturbed semantic processing has been identified as the underlying mechanism causing the language deficits (Raymer et al., 1997; Raymer & Gonzalez Rothi, 2001). It is possible that these deficits involving the semantic category structure might prevent generalization to untrained items and perhaps account for such findings in the literature (Mayer et al., 2004; Stanczak et al., 2005, 2006).

The current study proposed that individuals with early PAD provide an alternative population to study the effects of typical and atypical exemplar training from semantic categories. The rationale for this is based on the disease process, which results in diffuse damage (as opposed to focal damage from a stroke) and evidence that to some degree, the semantic system's structure is preserved in PAD (Albert & Milberg, 1989; Benson & Geschwind, 1985; Hartman, 1991; Nebes, 1992; Ober et al., 1995; Schwartz et al., 1996). This assertion is supported by findings that show knowledge about semantic categories is preserved in PAD (Flicker et al., 1987; Huff et al., 1986; Martin & Fedio, 1983; Nebes et al., 1986; Salmon, Butters et al., 1999; Schwartz et al., 1979; Warrington, 1975).

The goal of our exploratory study was to determine if choosing remediation targets for anomia based on typicality (i.e., atypical and typical) would result in generalization to untrained items. This involved a comparison of the family resemblance/prototype (Rosch, 1975) favoring typical category exemplars and the Complexity Account of Treatment Efficacy (Thompson et al., 2003), as well as the connectionist model of a computer simulation of acquired dyslexia (Plaut, 1996) favoring training atypical category exemplars for achieving generalization. We provided both lexical phonological information about the item via repetition as well as a semantic feature training with atypical and typical items for adults with PAD. In designing our study, relevant findings were culled from the PAD literature, particularly the three naming treatment studies (Abrahams & Camp, 1993; Fuller et al., 2001; Ousset et al., 2002; Rothi et al., 2005), the five aphasia studies comparing typical and atypical category exemplars (Kiran & Thompson, 2003b; Mayer et al., 2004; Plaut, 1996; Stanczak et al., 2005, 2006), and other relevant findings from the aphasia treatment literature.

According to Nadeau and Gonzalez Rothi (2004), the connectionist approach views anomia (that is caused by a semantic deficit) as a reflection of insufficiently engaged representations of features that are critical for making distinctions among concepts. When a network is damaged, a large amount of information still remains in the network, so the focus should be on refining the damaged network via semantic therapy. In particular, the network needs to be changed in terms of its connectivity so that there is more reliable engagement of distinguishing features simultaneously while there is relatively a disengagement of the shared features (Nadeau & Gonzalez Rothi, 2004). These findings suggest that participants with early PAD might therefore be appropriate for the semantically based training provided in our study which compared typical and atypical category exemplars.

As noted previously from the aphasia literature, a facilitation effect refers to improved performance on trained items from pretest to posttest in a single session (Howard, 1985; Patterson et al., 1983). Our study used a single session for different reasons, including the lack of picture naming treatment studies in the PAD literature, and to determine if the lexical-semantic system affected by PAD could be stimulated in a short period. This information would be useful to know prior to development of an intervention for naming deficits in adults with PAD.

For the purpose of this study, the term “response generalization” is adapted from Thompson (1989) to indicate that a participant has improved naming of untrained items based on training of other items. Thompson (1989) suggested that for response generalization to occur, the training and the probes should allow for sampling of responses that have a similarity to the trained items, either within a defined set or across it. The methodology for our study allowed for a measure of generalization to untrained atypical and untrained typical items regardless of the type of training that category received. (e.g., in the Train-Typical vehicles condition, there were untrained typical vehicles and untrained atypical vehicles). The aphasia literature suggests criteria to measure generalization; some examples include an increase in performance from baseline by 40% (Kiran & Thompson, 2003b) or 50-80% (Thompson, 1989). Another approach has been an increase of baseline performance as measured by three or more probe items (Boyle, 2004). In our exploratory study, we chose to measure generalization based on an increase from pretest that was statistically significant at the alpha .05 level.

The typicality ratings from Rosch (1975) provided categories and their exemplars for this study. Using 24 items per semantic category allowed for an adequate number of training and generalization items (Kiran & Thompson, 2003b; Rothi et al., 2005). Two categories were used for the training (seven atypical and seven typical items) and the other was a control category (Mayer et al., 2004; Rothi et al., 2005). While three of the four aphasia treatment studies comparing typical and atypical category exemplar training (Kiran & Thompson, 2003b; Stanczak et al., 2005, 2006) used the categories of birds and vegetables, two of them reported generalization . just one used a combination of natural kinds and man-made artifacts (Mayer et al., 2004).

Our study involved three artifact categories: vehicles, clothing and tools. Although at least one PAD study has suggested that there was no significant difference in performance on natural kinds and artifacts (Tippett, Grossman, & Farah, 1996), others have found there is evidence that adults with PAD have shown lower levels of performance on natural kinds (Chertkow, Bub, & Caplan, 1992; Montanes, Goldblum, & Boller, 1995; Warrington, 1975), or that the profile changes with the progression of the disease (Gonnerman et al., 1997). To avoid this controversy, the current study used only artifact categories.

It has been suggested that providing feedback is an important part of the learning process (Bandur & Shewan, 2001). Repetition has been used as a training task, and found to be effective in obtaining a treatment effect (Fuller et al., 2001; Mayer et al., 2004; Patterson et al., 1983; Pring et al., 1993; Rothi et al., 2005). Hickin et al. (2002) suggested that presenting a picture stimuli and repeating the name of the word actually elicits semantic processing. Our study combined these three components. This was followed by the semantic training requiring generation of semantic information about the target item. The rationale for this comes from studies that suggest that the naming deficit in PAD was both due to semantic specification and lexical access deficits (Bowles et al., 1987; Williamson et al., 1998). Our study used repetition so that if the PAD participants have a deficit in both the semantic system and lexical access, the form of the word was provided and being reinforced, and the semantic training targeted the semantic part of the deficit.

Evidence of a treatment effect has been reported after training with a combination of semantic and lexical (i.e., word form) tasks (Drew & Thompson, 1999; Le Dorze et al., 1994; Wiegel-Crump & Koenigsknecht, 1973). Two other anomia treatment approaches have also been found to result in treatment and generalization effects in adults with aphasia are semantic distinction training (Hillis, 1998); and semantic feature analysis (SFA) (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000). Both of these approaches combined specific training of semantic features with presentation of the phonological form. As noted above, Kiran and Thompson (2003b) applied a similar process and suggested it was an important aspect of the study. The three other typicality studies (Mayer et al., 2004; Stanczak et al., 2005, 2006) also used a more simplified version of this.

Our study attempted to achieve some of the benefits of SFA (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000), for example, to facilitate a threshold level of activation of the semantic network that surrounds the targeted item. Furthermore, like SFA, we aimed for systematic activation of the distinguishing features of the item rather than shared features (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000). This process is based on the spreading activation theory, which posits that selection of the target item among competitors is based on the highest activation achieved (Collins & Loftus, 1975; Dell & O'Seaghdha, 1992; Levelt, Roelofs, & Meyer, 1999). Once the concept is selected, phonological information is activated so that the target item is produced (Collins & Loftus, 1975; Dell & O'Seaghdha, 1992; Levelt et al., 1999). Finally, SFA has also been shown to result in both a treatment effect and generalization to untrained items (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000). More specifically, our training protocol was a modified version of SFA (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000). Our PAD participants answered questions about the target item and then named it. We also included an aspect of the semantic distinction training (Hillis, 1998), by asking them about how an item differs from others like it. When the participant was unable to respond correctly, the clinician provided the information and had the participant repeat it, based on Coelho et al. (2000).

Other studies also required active production of information, including the use of antonyms and synonyms (McNeil et al., 1998), personalized cueing (Freed, Marshall, & Nippold, 1995; Lowell, Beeson, & Holland, 1995; Marshall, Freed, & Karow, 2001); and the participant using circumlocution (i.e., describing the item) (Francis et al., 2002). Some of these studies also reported generalization to untrained items (Francis et al., 2002; Lowell et al., 1995). In conclusion, we used information from the three PAD naming treatment studies (Abrahams & Camp, 1993; Fuller et al., 2001; Ousset et al., 2002; Rothi et al., 2005), studies comparing training atypical and typical category exemplars (Kiran & Thompson, 2003b; Mayer et al., 2004; Plaut, 1996; Stanczak et al., 2005, 2006) and other aphasia studies that showed treatment and generalization effects (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000; Davis & Pring, 1991; Francis et al., 2002; Hickin, Best, Herbert, Howard, & Osborne, 2001; Hillis, 1998; Howard et al., 1985; Pring et al., 1993; Raymer, Thompson, Jacobs, & le Grand, 1993). Many of these studies also provided rationale for predictions about our research questions, which are discussed below.

Research Questions and Predictions

The first four research questions focused on performance (i.e., accuracy and reaction times) on picture naming. While the first two research questions addressed the effects of repetition and the training (i.e., facilitation effects), the third question addressed changes in performance on untrained items (i.e., generalization or practice effects). The fourth question also addressed generalization, but to another task, category generation. These questions are listed below with their corresponding predictions. The type of training (i.e., training type) always refers to the typicality of trained items in a particular trained category, and are always referred to as Train-Atypical, to indicate that atypical items were trained, and Train-Typical, to indicate that typical items were trained.

<u>Research Question 1:</u> Does repetition of the control items (i.e., not with semantic training) impact performance at posttest?

Prediction: There will be better performance (higher accuracy and faster reaction times) on repeated items compared to items that were not repeated (Fuller et al., 2001; Patterson et al., 1983; Pring et al., 1993; Rothi et al., 2005)

Research Question 2. What are the effects of semantic training on trained items?

Research Question 2a: Do accuracy scores and the reaction times for trained items change from pretest to posttest?

Prediction: Accuracy and reaction times for trained items will show a facilitation effect after training: Accuracy scores will be higher and reaction times will be faster on the trained items at the posttest (Drew & Thompson, 1999; Howard, 1985; Le Dorze et al., 1994).

Research Question 2b: Does the type of training (i.e., Train-Atypical or Train-Typical) impact the performance on trained items at posttest?

Prediction 2b.1: Performance will improve after both types of training, but there will be a greater improvement after training typical items (Plaut, 1996).

Prediction 2b.2: Performance will only improve after training typical items (Rosch, 1975).

Research Question 3a: Will there be an overall change in performance on untrained items from trained categories?

Prediction 3a: Accuracy will increase and reaction times will decrease in untrained items in trained categories (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000; Hillis, 1998; Plaut, 1996; Wiegel-Crump & Koenigsknecht, 1973).

Research Question 3b: Does the type of training (i.e., Train-Atypical/Train-Typical) affect performance on untrained items in trained categories (i.e., generalization)?

Prediction 3b.1: Training atypical items (Train-Atypical) will result in greater generalization as shown by higher accuracy scores and faster reaction times (Kiran & Thompson, 2003b; Plaut, 1996; Stanczak et al., 2006; Thompson et al., 2003).

Prediction 3b.2: Training typical exemplars will result in greater generalization: higher accuracy scores and faster reaction times for typical exemplars within the same category (Rosch, 1975).

Prediction 3b.3: Training typical exemplars will result in greater generalization to both untrained typical and atypical items, because typical items have richer representations and share more features with all other category members (Heilman, personal communication, 2005).

Research Question 3c: Is there an interaction between time and item typicality (i.e., atypical items and typical items)?

Prediction 3c: Untrained typical items from trained categories will show more improvement than atypical items from trained categories after training.

Research Question 3d: Is there an interaction among time, training type, and item type?

Prediction 3d.1: Training atypical items will result in generalization to untrained typical items (Kiran & Thompson, 2003b; Plaut, 1996; Stanczak et al., 2006) and to untrained atypical items (Plaut, 1996; Stanczak et al., 2006).

Prediction 3d.2: Training typical items will have a positive impact on untrained typical items from trained categories and a negative impact on the performance with atypical items from trained categories (Plaut, 1996).

Research Question 4: Is there generalization to category generation?

<u>Prediction 4:</u> There will be an increase in the number of items named in the trained categories, and this will be greater for the trained categories compared to the untrained categories (Ousset et al., 2002).

CHAPTER 4
METHODS

The purpose of this exploratory study was to extend previous research on the effects of training typical and atypical category exemplars on generalization (Kiran & Thompson, 2003b; Mayer et al., 2004; Plaut, 1996; Stanczak et al., 2005, 2006) while comparing two views of typicality (Plaut, 1996; Rosch, 1975; Rosch & Mervis, 1975). Participants with early probable Alzheimer disease participated in this picture naming training study. The primary goal was to determine if remediation targets could be chosen based on typicality, to maximize generalization to untrained items (Kiran & Thompson, 2003b; Plaut, 1996; Stanczak et al., 2006). This was achieved by training atypical (i.e., Train-Atypical) and typical (i.e., Train-Typical) category exemplars from two semantic categories and leaving untrained atypical and typical items in each category for both the same item typicality that was trained (e.g., Train-Atypical, untrained atypical items), and the opposite (e.g., Train-Atypical, untrained typical items). Using a repeated measures (i.e., within subject) design, our study provided each participant with a modified semantic feature based training (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000) for each category. This was completed in a single session and the results were analyzed for a facilitation effect (Hickin et al., 2002; Howard, 1985; Howard et al., 1985; Patterson et al., 1983) and generalization to untrained items. While the outcome measures (i.e., dependent variables) for the picture naming were accuracy and response time means (in milliseconds), the number of items generated was the outcome measure for the category generation task.

Participants

Following approval from the UF Institutional Review Board (Appendix A), the participants with PAD were recruited and enrolled into the study. Recruitment of participants was accomplished with the help of the University of Florida Memory and Cognitive Disorder Clinics. Additional recruitment efforts included presenting information and flyers about the study at the Alzheimer's Association, and Al'z Place, an adult day care facility for individuals with Alzheimer disease and other dementias. Participants were compensated $20 for their time.

Diagnosis of the participants with mild-moderate PAD was made by a team of medical professionals including neurologists and neuropsychologists based on criteria from the National Institute of Neurological and Communicative Disorders & Stroke (NINCDS), and the Alzheimer's Disease and Related Disorders Association (ADRDA) (McKhann et al., 1984). Individuals were not eligible for the study if they had any history of brain or head injury, psychiatric hospitalization, alcohol or drug abuse; any chronic medical or psychiatric conditions, or any developmental learning disability. Participants were pre-screened at the UF Memory Disorders Clinics for PAD. Some of the results from the neuropsychological testing are shown in Table B-1, Appendix B.

A phone screen was used to confirm eligibility and to collect demographic information. Twelve adults (7 women and 5 men) with newly diagnosed early PAD completed the study. The average age was 77.6 (SD = 9.6) and the average number of years of education was 15.8 (SD = 3.8). The mean score for the Mini Mental Status Exam (Folstein et al., 1975) was 22.6 (SD = 3.4) with a range of 17-27. More specific demographic information and performance scores for each participant are shown in Table 4-1.

Table 4-1. Participant Demographics and Scores on Mini Mental Status Exam and Reading Subtest

ID #	Age	Years of Education	Gender	Mini Mental Status Exam	Reading Subtest WRAT-3 (percentile)
1001	72	17	F	21	87
1002	77	13	F	24	50
1003	76	10	M	17	8
1004	92	24	M	24	91
1005	82	12	F	19	10
1006	78	12	F	22	50
1007	55	20	M	27	90
1008	84	15	F	25	75
1009	69	18	M	20	50
1010	86	16	F	19	63
1011	75	16	M	27	81
1012	85	16	F	26	87
Mean	77.6	15.8		22.6	61.8

Notes: Mini Mental Status Exam (Folstein et al., 1975); WRAT-3=Wide Range Achievement Test-3 Reading subtest (Wilkinson, 1993).

Experimental Stimuli

There were six training lists (1-6) that were counter balanced using a Latin Square for training conditions (i.e., Train-Atypical, Train-Typical, Control) and semantic category (i.e., vehicles, clothing, and tools). As depicted in Table 4-2, each list had two training conditions and one control condition. Participants were assigned to these lists sequentially, as they were enrolled into the study. For example, the first participant (#1001) had List 1, and this process continued with the remaining participants (e.g., participant eight (#1008) had list 2).

Table 4-2. Lists of Training Conditions and Corresponding Training Conditions

List #	Vehicles	Clothing	Tools
1	Train-Typical	Control	Train-Atypical
2	Train-Atypical	Train-Typical	Control
3	Control	Train-Atypical	Train-Typical
4	Train-Typical	Train-Atypical	Control
5	Control	Train-Typical	Train-Atypical
6	Train-Atypical	Control	Train-Typical

The pre and posttesting picture naming stimuli consisted of 72 black and white line drawings from three categories: vehicles, clothing, and tools. (Table C-1, in Appendix C, shows the list of 72 items used in the study). Since there are several variables being addressed, it is important to first clarify the terminology for the items in the pre and posttest naming tasks, then the details related to the training are discussed. All items were designated as either typical or atypical, based on the ratings in the Rosch norms (Rosch, 1975). Items were chosen from these three categories based on the seven-point typicality ratings from Rosch (1975), such that half the items in each category were atypical and the other half were typical. Typicality ratings for items used in this study ranged from 1.02 (the most typical) to 5.36 (the most atypical). For each category, the typical and atypical items were selected at about a midpoint range so that half would be typical and the other half would be atypical. The mean typicality score for the atypical items 3.5 (SE = .12), and the mean typicality score for the typical items was 1.6 (SE = .08). An independent t-test indicated that these groups of items were significantly different, $t(70)=12.967$, p = .00.

The picture stimuli were gleaned from a variety of sources, with the majority from the Florida Semantic Battery (Raymer et al., 1990). Name agreement was determined by piloting the naming task on 10 University of Florida (UF) undergraduates in the Language over the Lifespan Lab.

Each semantic category was designated as either the control category (i.e., no items received semantic training), Train-Typical (i.e., only typical items received semantic training), or Train Atypical (i.e., only atypical items received semantic training). There were several reasons for the choice and number of categories. All artifact categories were chosen to avoid the possible confound of mixing biological and artifact categories. Furthermore, it was important to prevent any overlap from other semantic categories (e.g., vehicles and sports). Also, previous research on picture naming has included these categories (Grossman, 1981; Mayer et al., 2004), although other studies of typicality effects have only included natural kinds (Fuller et al., 2001; Kiran & Thompson, 2003b; Rothi et al., 2005). There were 24 exemplars in each category (Kiran and Thompson, 2003b) 12 typical and 12 atypical.

From these 72 items, the semantic training stimuli were selected. For each category, seven atypical items were designated as the training set for when the category appeared in the Train-Atypical condition, and 7 typical items were designated as the training set for when that category appeared in Train-Typical condition. Thus, in a category that appeared in the Train-Typical condition, for example, the 7 typical items in the training set would receive semantic training, leaving 5 untrained typical items and 12 untrained atypical items from that category to analyze for generalization. Untrained items were those items that received no semantic training, but were members of the trained category. Therefore, for any given item, it was necessary to specify whether it was typical or atypical, trained or untrained, and what type of training, if any, that category received. The control items were untrained items in the untrained category. Control items were named at both pretest and posttest. Because all items from the trained categories were repeated, half of the items from the control category were also repeated to determine what effect simple repetition without semantic training had on lexical access. An example based on list 1 is found in Table 4-3.

Table 4-3. Example from List 1: Train-Typical Vehicles and Train-Atypical Tools.

Trained	Untrained (same as trained)	Untrained (opposite of trained)
7 Typical Vehicles	5 Typical Vehicles	12 Atypical Vehicles
7 Atypical Tools	5 Atypical Tools	12 Typical Tools

Note: The control category was comprised of 24 items of clothing.

The list of phrases or cues for the semantic training was developed by polling volunteers (ranging from a Ph.D. investigator, a certified speech-language pathologist, and several undergraduate research assistants) who are members of the UF Language over the Lifespan Lab in the Department of Communication Sciences and Disorders. These cues were compiled and selected based on the most salient ones.

In a series of post hoc analyses of the stimulus words from these pictures, we examined a possible influence from four variables on the atypical and typical items: typicality (Rosch, 1975), which we expected to show a significant difference; written word frequency (Frances & Kucera, 1982) and familiarity (Wilson, 1987), and phoneme length. We used an independent sample t-test to compare the complete list of 72 atypical and typical items on the four variables. For typicality and phoneme length, ratings were available for each item. The independent t-test for the typicality ratings had a *p*- value of .000, indicating a significant difference in terms of typicality between the atypical and typical items. The independent t-test for phoneme length was not significant, $t(70)=2.528$, $p = .69$, indicating there was no difference between the atypical and typical items. In other words, the phoneme length for the atypical and typical items did not influence the results.

The independent t-test comparing typical and atypical items from the Complete List was not significant for frequency, $t(70) =-1.389$, $p =.169$, suggesting that there was no difference in the frequency ratings between the atypical and typical items. However, although the independent t-test for familiarity was significant, $t(38) =-3.625$, $p = .00$, this analysis was only computed on 40/72 items (i.e., many of the items did not have familiarity ratings). Thus, the results must be interpreted with caution. Table 4-4 depicts the means of the four variables discussed above.

Table 4-4. Means for Complete List of Atypical and Typical Items.

Train Type	Typicality Score	Frequency	Familiarity	Phoneme Length
Atypical	3.6	15.7	480	4.8
Typical	1.7	25.1	551	4.9

Notes: Typicality (Rosch, 1975); Freuency (Frances and Kucera, 1982); Familiarity (Wilson, 1987).

Procedure

Prior to participation in the study, written informed consent was obtained from each participant. As needed, the clinician allowed time for the participant to read the informed consent form. Then, the clinician answered any questions and assisted the participant in interpreting the information.

Screening

Two measures were used for the screening: The Mini-Mental Status Exam (Folstein et al., 1975) and the reading subtest of the Wide Range Achievement Test (WRAT-3) (Wilkinson, 1993). The clinician administered the MMSE using the standard procedure (Folstein et al., 1975). Next, the clinician presented a computerized version of the reading subtest (Wilkinson, 1993) in which the participant read aloud the individual words from the screen. These responses were digitally recorded on an Olympus DM-10 digital recorder (Olympus, 2002) and response times (RTs) were collected via DirectRT (Jarvis, 2004) for later scoring and analysis.

Pretesting

The pretest consisted of two tasks, Category Generation (Benton, 1968) and Picture Naming. For the Category Generation task, the clinician provided instructions for the participant to name as many items as possible in one minute for a given category (i.e., vehicles, clothing, and tools). The categories were presented in a pre-determined random order. The clinician wrote the responses verbatim on the pretest form, and responses were also digitally recorded for checking reliability.

The Picture Naming task consisted of the complete set of 72 line drawings randomly presented on the computer via DirectRT (Jarvis, 2004). At the onset of the task, the clinician read aloud the set of instructions that appeared on the screen and confirmed comprehension of the task. The clinician controlled the pace of the session by using the mouse button presses on a wireless mouse. Each picture stayed on the screen until the clinician advanced to the next picture (i.e., the picture did not disappear or time out). There were four practice items from the category of fruit. After the participant attempted to name a presented picture, the clinician informed the participant if he/she was correct or not, and then asked the participant to repeat the correct name of the item three times. If the participant produced a correct alternative name of the item (e.g., “jet” for “airplane”) their response was accepted as correct, and this word was repeated (Appendix C includes a list of acceptable alternative names for items). The clinician then advanced to the next trial. As stated above, twelve items from the control category were not repeated (half were typical and half were atypical). A prompt appeared on the computer screen to alert the clinician as to whether this was a repeated item or not. This procedure was used for all of the items in the two trained categories, and for half of the control items (i.e., repeat or no repeat, without semantic training).

More specifically, after the participant named an item, the word “repeat” appeared on the lower center section of the screen. Next, the clinician provided verbal feedback and directed the participant to repeat the name of it three times (i.e., “That’s correct, now you say it three times ___, ___, ___” or It’s supposed to be a ___, now you say it three times, ___, ___, ___”) (Kendall, personal communication, 2005; Fuller et al. 2001; Rothi et al. 2005). When an unrepeated item was presented, the participant named the item, then the words “Go on” appeared in place of the word “Repeat,” and the clinician advanced to the next item. Thus, on these control items the clinician did not provide feedback about the participant’s response, and the item was not repeated. The clinician continued this process with each of the remaining pictures until all 72 items had been named. At this point, the participant took a short break while the clinician set up for the semantic training. All of the naming responses were digitally recorded for later transcription and analysis, and response times were collected via DirectR (Jarvis, 2004).

Semantic Training

This was a guided, feature generation task, loosely based on semantic feature analysis (Boyle, 2004; Boyle & Coelho, 1995; Coelho et al., 2000) and semantic distinction training (Hillis, 1998). The computer presented the pictures from one of the training categories (i.e., seven items) in random order and the clinician asked the participant to answer each of the following questions about the item: 1. “What is its function?”; 2. “Who uses it?”; and 3. “How is it different from other __?” (i.e., things like it). After the participant answered the three questions, the clinician asked the participant to name the item. Whenever a participant was unable to correctly answer a question or provided ambiguous information, the clinician provided cues from a pre-determined typed list. Whenever the participant was unable to name the item, the clinician stated it and asked the participant to repeat it. The clinician continued this process with the remaining six training items for that category. Each participant was trained on the two categories from their designated list. The trained categories (e.g., Train-Atypical Tools, and Train-Typical Vehicles) were randomly presented in blocks, but the order of the items presented within each category was random. This allowed each participant to be trained on each category and the corresponding training type (i.e., Train-Atypical, Train-Typical) for a total of six practice sets (i.e., three presentations for each category). Then, the participant took a break while the clinician prepared for the posttesting. The responses from the training were also digitally recorded.

Posttesting

The posttesting consisted of the same two tasks and instructions as the pre-testing: Category Generation and Picture Naming. During the Category Generation task, the participant was given the same instructions and the same random order for generating as many items as possible for each of the three categories (vehicles, clothing, and tools). The clinician wrote down participants' responses verbatim and also digitally recorded them in order to check reliability.

For the Picture Naming task, the same 72 item picture set was presented by computer running Direct-RT (Jarvis, 2004). Items were presented in a different random order than at pretest. In this version, the participant only named the pictures aloud, without feedback or repetition.

Scoring

The scoring for the screening tasks was completed as follows. The Mini Mental Status Exam was scored using the standard scoring outlined by Folstein et al. (1975). Digital recordings of the reading subtest from the WRAT-3 (Wilkinson, 1993) were used to code the data for accuracy of production by a trained research assistant. The clinician scored the responses and calculated raw and standard scores and converted them to the percentiles, based on instructions from the WRAT-3 Manual (Wilkinson, 1993). It is important to note that normative data for adults with PAD is not provided in the WRAT-3 Manual (Wilkinson, 1993), and that most of our PAD participants were beyond the age ranges provided (i.e., 55-64.11; and 65-74.11). Thus, these findings must be interpreted with caution.

The pre and posttest picture naming performance for accuracy and response times was collected by Direct RT (Jarvis, 2004). Accuracy and response times were tracked by the clinician's mouse clicks (i.e., left for correct and right for incorrect) based on the participant's response and coded these as ones and zeros. The list of acceptable alternate names is displayed in Appendix C. The clinician used the same procedure for the response times (i.e., left for usable and right for unusable). For the response times, only fluent, correct responses were computed in the analyses (e.g., response times were excluded if the participant said, "uh airplane", but the response would be scored as accurate/correct). Furthermore, response times for each participant were adjusted for outliers by replacing any value that was greater than three standard deviations above the mean, with the mean value.

The responses for the Category Generation task were transcribed independently by a trained research assistant. The clinician and research assistant then compared these results to the clinician's original on-line written documentation from the session and discrepancies were resolved by referring back to the corresponding voice files until agreement was reached. Informal reliability measures were carried out with ~ 50% of the data, using the voice files. Agreement was ~98% and disagreements about scoring of responses were resolved by consensus during the Language over the Lifespan Lab meetings.

Responses from the Picture Naming task were also transcribed by a trained research assistant. The clinician used these transcripts to verify coding of 100% of the picture naming scoring.

Statistical Analyses

Our study used a repeated measures design. The analyses compared pre and post-test category generation and picture naming performance using both accuracy and response times where possible. Analyses evaluated the effects of repetition and practice from data in the control condition, the effects of semantic training, and generalization effects to untrained items in trained categories. The effects of training atypical and typical category exemplars were also examined.

The pre and posttest picture naming data were analyzed using t-tests, repeated measures analysis of variances (ANOVAs) and follow-up post hoc t-tests where necessary. It is important to point out that the pretest naming data reflects the participants' initial attempt at naming the pictures; this was followed by the repetition and semantic training. Thus, the effect of repetition and training can only be evaluated by comparing pre and posttest performance. Due to an error in the programming related to the control items, four participants repeated all items. Consequently, data from only eight of the twelve participants could be included in the analysis of repetition effects.

Since inspection of the data for some of the analyses showed pronounced differences at pretest, difference scores were calculated from the pre and posttest data and then analyzed. The dependent variables for picture naming were means from the accuracy and response time data. The independent variables varied according to the questions, but included time (pre /posttest), repetition (no repeat/repeat), training type (i.e., the Train-Atypical/Train-Typical), item typicality (i.e., item type, atypical/typical). For consistency, whenever possible the term atypical is always discussed before typical, this convention is alphabetical and used in order to reduce confusion with the terminology (there is no intentional bias).

The dependent variables for the pre and posttest category generation data were the number of items generated per category, and the independent variables were time and training type (i.e., Control category, Train-Atypical, Train-Typical). The data was analyzed with repeated measures ANOVAs.

Research Question 1: Control Items

Does repetition of the control items (i.e., not with semantic training) impact performance at posttest? This question addressed the possibility of a repetition or practice effect. Only accuracy and response time means from the untrained items in untrained categories (i.e., the control condition) made up the dependent variable. Half of these were in the repetition condition and half were not (i.e., not repeated). The independent variables were time (pre/post) and repetition (yes/no). A (2) Time (pre/post) X (2) Repetition (yes/no) repeated measures ANOVA was used to answer this question.

Research Question 2: Trained Items in Trained Categories

This two-part question addressed the possibility of a facilitation effect in semantically trained items (i.e., Train-Atypical and Train-Typical). The data consisted of scores from the 14 trained items (i.e., the seven atypical and seven typical items from the two trained categories). This within category measure was assessed as a comparison of pretest and posttest performance on the confrontation naming task

Research Question 2a: Do accuracy scores and the response times for trained items change from pretest to posttest? The independent variable was time (i.e., change from pretest to post test).

Research Question 2b: Does the type of training (i.e., Train-Atypical or Train-Typical) impact the performance at posttest? The independent variables were time and training type.

To answer both of these questions, (2) Time (pre/post) x (2) Training type (Train-Atypical/Train-Atypical) repeated measures ANOVAs were computed.

Research Question 3: Untrained Items in Trained Categories

Question 3 had several subcomponents, but overall addressed the possibility of a primary generalization effect. If there was a facilitation effect in question 2, the purpose of questions 3a, b, c, and d was to determine if the facilitation effect generalized to untrained items within the same trained category.

Research Question 3a: Is there generalization to untrained items in the trained category? The independent variable was time (pre/post).

Research Question 3b: Does the type of training (i.e., Train-Atypical/Train-Typical) affect generalization to untrained items in trained categories? The independent variables were training type and time.

Research Question 3c: Does the typicality of the untrained item (i.e., atypical items and typical items) determine whether it will benefit from generalization? The independent variables were time and item type.

Research Question 3d: Do time, training type, and item type interact to promote generalization limited to a particular item type following a particular training type? The independent variables were time, training type, and item type.

The analyses for Question 3 used the 34 untrained items in the two trained categories (i.e., 17 untrained items in each trained category). For the accuracy means, we used a (2) Time x (2) Training Type (Train-Atypical/Train-Typical) x (2) Item Type (Atypical/Typical) repeated measures ANOVA.

The response time means were analyzed differently. For Question 3a, a t-test was used to compare pre and posttest performance. To answer Questions b, c, and d difference scores were computed between pre- and posttest response times and analyzed via a (2) Training Type x (2) Item type repeated measures ANOVA.

Research Question 4: Category Generation

Research Question 4: Is there generalization from picture naming to category generation? Does the type of training the category receive affect whether there is generalization to category generation. The independent variables were time and training type. The dependent variable was total number of items generated per category and these were analyzed with a (3) Training Type (Train-Atypical, Train-Typical and Control) x (2) Time repeated measures ANOVA.

CHAPTER 5
RESULTS

In this study we compared the performance of adults with probable Alzheimer disease (PAD) on picture naming and category generation tasks before and after a semantic training using atypical and typical category exemplars. The accuracy means and response times (in milliseconds, ms) from pre and posttest picture naming, and number of items generated per category on a category generation task were analyzed with t-tests and repeated measures analysis of variances (ANOVAs). To control for some of the differences in the pretest scores across conditions, difference scores were computed for several of the questions and then analyzed to compare direction and magnitude of change. An a priori decision was made to use a .05 level of significance for all of the analyses. The two training conditions are always referred to as Train-Atypical (i.e., training atypical items) and Train-Typical (i.e., training typical items).

Research Question 1: Control Items

Question 1 asked if repetition of the control items (i.e., not with semantic training) impacted the performance of picture naming at posttest. The dependent variables were the accuracy and response times, and the independent variables were time and the repetition and no repetition conditions. The data for these analyses were from the control categories (i.e., untrained items in untrained categories) in which half of the items were repeated (i.e., repetition condition) and the other half were not (i.e., no repetition condition).

A (2) Time (pre/posttest) x (2) Repetition (yes/no) repeated measures ANOVA was used to examine the data from eight participants. The results showed that there were no significant main effects or interactions for the accuracy means. Although the items in the repetition condition were somewhat higher (Mean (M) = .89, SE = .04) compared to items in the no repetition condition (M = .80; SE = .04), this was not significant, $F(1,7) = 3.691$; $p = .096$; $\eta^2 = .35$. There were no other main effects or interactions with the accuracy means for the control items. Thus, repetition did not have a significant effect on the accuracy means.

The response time means showed a similar pattern from the (2) Time x (2) Repetition repeated measures ANOVA. There were no significant main effects to suggest that repetition alone impacted performance from pretest to posttest. There were decreases in the response time means, indicating faster responses, but these were not significant. One was for time, $F(1,7) = 3.702$; $p = .096$; $\eta^2 = .346$; M_{Pre} = 2456 ms (SE = 521); M_{Post} = 1858 ms (SE = 243). The other was for repetition, $F(1,7) = 4.416$; $p = .074$; $\eta^2 = .387$; M_{REP} = 1979 ms (SE =321 ms), M_{NoRep} = 2334 ms (SE = 440 ms). The interaction was not significant, $F(1,7) = 1.566$; $p = .251$; $\eta^2 = .183$.

The items in the control condition did not receive any semantic training, but half of them were repeated at pretest. Although the accuracy in the control condition means were slightly higher, this was not significant. The response times decreased somewhat, but not significantly. Thus, repeating items in the control condition did not result in statistically significant changes in accuracy or response times.

Research Question 2: Semantically Trained Items

Questions 2a and b asked if performance on the semantically trained items (i.e., Train-Atypical and Train-Typical) changed from pre to posttest, and, if so, whether there was a difference between the two types of training (i.e., Train-Atypical versus Train-Typical). The dependent variables for both questions were accuracy and response times from the 14 trained items (i.e., seven atypical and seven typical items) from the 12 participants. Table 5-1 provides an overview of the accuracy and response time means. These are discussed separately below.

Table 5-1. Pre and Post Group Accuracy and Response Time Means for Trained Items

Accuracy			Response times		
Pretest	Posttest	*p* value	Pretest	Posttest	*p* value
.76 (.04)	.86 (.03)	.004	1719 ms (183)	1362 ms (96)	$p < .06$

Note: The numbers in parentheses are standard errors (*SE*).

To determine the effects of semantic training on performance, a (2) Time (pre/post) x (2) Training type (Train-Atypical/Train-Typical) repeated measures ANOVA was used to compare pre and posttest accuracy means. These showed a significant main effect of time, $F(1,11) = 12.780$; $p < .01$; $\eta^2 = .54$. The trained items were named significantly more accurately at posttest compared to the pretest ($M_{Pre} = .76$, $SE = .04$; $M_{Post} = .86$, $SE = .03$). There were no other main effects or interactions for the accuracy means. Although trained items were more accurate after the training, there was no difference between training typical versus atypical items. Table 5-2 shows the results for the accuracy and response times.

Table 5-2. Group Accuracy and Response Time Means for Trained Items (Train-Atypical; Train-Typical)

	Accuracy		Response Times		
	Pretest	Posttest	Pretest	Posttest	Difference Scores*
Train-Atypical	.71 (.07)	.84 (.05)	1506 ms (168)	1397 ms (154)	-109 (217)
Train-Typical	.81 (.06)	.87 (.05)	1931 ms (243)	1326 ms (101)	-605 (186)

Notes: (*SE*).* Difference scores for the response time means were significant ($p = .04$).

Response times were examined to determine if semantic training impacted the speed of lexical access. A t-test of pre and posttest performance showed that response times were 357 ms faster after the training; these results approached significance, $t(11) = 2.12$; $p = .057$; $M_{\text{Pre}} = 1719$, $SE = 183$ ms; $M_{\text{Post}} = 1362$ ms, $SE = 96$ ms.

Since the pretest response times for the two training conditions varied, difference scores were calculated for Question 2b to determine if the type of training (i.e., Train-Atypical versus Train-Typical) affected the response times. The difference scores were computed by subtracting the posttest response times from pretest response times for each training type. Next, a paired t-test was computed, which showed that the results were significant, $t(11) = 2.28$; $p = .04$. As shown in Table 5-2, there was a larger difference for the Train-Typical items (M= -605, SE = 217 ms), which were the slowest at pretest, but fastest at posttest compared to Train-Atypical items (M = -109, SE = 186 ms).

Thus, the semantic training had a facilitation effect on lexical access in individuals with mild-moderate probable Alzheimer disease (PAD), increasing accuracy on picture naming. Although the change in accuracy for trained items was similar regardless of training type, response times for Train-Typical items benefited more from semantic training.

Research Question 3: Untrained Items in Trained Categories

Question 3 addressed various aspects of generalization to the 17 untrained items in the 2 trained categories. For each trained category, these 17 items comprised 5 untrained items with the same typicality as the items trained and 12 items of the opposite typicality. For example, if Typical Vehicles were trained (i.e., Train-Typical), untrained items would consist of 5 untrained typical vehicles and 12 untrained atypical vehicles. The dependent variables were accuracy and response times. The different components of

Question 3 were addressed using different repeated measures ANOVAs. The accuracy scores are discussed together, answering Questions 3,a, b, c, and d, and is followed by a similar procedure for the response times.

The purpose of Question 3a was to determine if there was generalization to untrained items in the trained category. Question 3b aimed to determine if the type of training (Train-Atypical versus Train-Typical) affected generalization within a category at posttest. The independent variables were time and training type. Questions 3c, d, and e asked about the possibility of specific interactions between time and training type, time and item type (i.e., item typicality, atypical or typical); and among time, training type, and item type. The independent variables were time, training type, and item type, and the dependent variable was accuracy means. To address these questions systematically, a (2) Time x (2) Training type (Train-Atypical/Train-Typical) x (2) Item type (Atypical/Typical) repeated measures ANOVA was computed. Table 5-3 shows the results.

Table 5-3. Analysis of Variance for Accuracy Means for Untrained Items in Trained Categories.

Effect	df	F	η^2	*p* value
Time	1	5.811	.346	.035*
Train	1	4.968	.311	.048*
Item Type	1	1.902	.147	.195
Time x Train	1	2.681	.196	.130
Time x Item Type	1	6.597	.375	.026*
Train x Item Type	1	3.717	.253	.080
Time x Train x Item Type	1	5.987	.352	.032*

Notes: * = significant at .05 alpha level.

The accuracy results for Question 3a showed that there was a significant main effect of time, $F(1,11) = 5.811$; $p < .04$; $\eta^2 = .35$. As depicted in Figure 5-1, the accuracy means were significantly higher at posttest compared to pretest (M_{Pre} .71, versus M_{Post} =.77; *SEs* = .04).

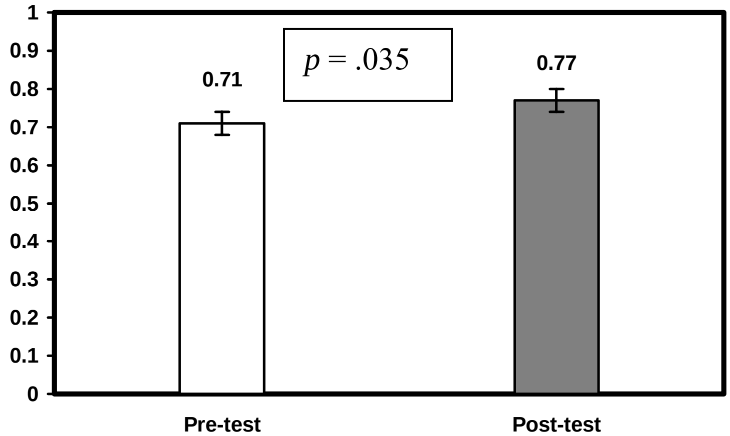

Figure 5-1. Accuracy Means for Untrained Items in Trained Categories, Main Effect of Time (p = .035; the numbers in the figure are proportions).

There was also a main effect of training type, $F(1,11) = 4.968$; p =.048; $\eta^2 = .31$, but this was secondary to the three-way interaction, discussed below. There was no interaction between time and training type that the category had received. However, a significant two-way interaction between item type and time, $F(1,11) = 6.597$; $p = .03$; η^2 =.38 was found for Question 3d. This revealed that, at pretest, the accuracy means for the atypical items were at .76 (*SE* = .06) but the typical items were only .66 (*SE* = .03). At posttest, the atypical items were at .78 (*SE* =.047) and the typical items were at .77 (*SE* = .04). Thus, the typical items showed a larger increase in accuracy than atypical items. A three-way interaction among training type, item type and time was also significant $F(1,11) = 5.987$; $p = .03$; $\eta^2 = .35$. As can be seen in Figure 5-2, significant increases in accuracy only occurred for untrained typical items in categories that received training on typical items (i.e., untrained typical items in the Train-Typical condition).

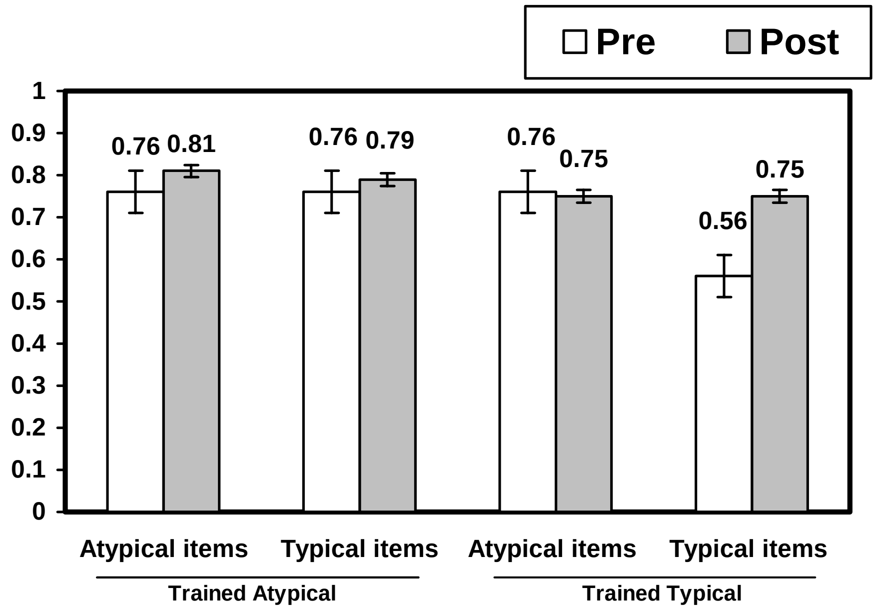

Figure 5-2. Untrained Items in Trained Categories: Three-Way Interaction Among Time (Pretest versus Posttest), Training Type (Trained Atypical versus Trained Typical) and Item Type (Untrained Atypical Items and Untrained Typical Items); $p = .03$; the numbers in the figure are proportions. Only typical items in categories that were Trained Typical increased in accuracy.

Due to variability in the pretest accuracy data, the 3-way interaction was explored using difference scores to examine change due to training. We computed pre to post difference scores for each of the four conditions. Then, we used these four difference scores in a (2) Training type x (2) Item type repeated measures ANOVA. This showed a main effect of item typicality, $F(1,11) = 6.597$; $p = .03$; $\eta^2 = .38$, indicating a larger difference for typical items ($M = .113$) compared to atypical items ($M = .02$). A significant interaction between training type and item typicality was also found $F(1,11) = 5.987$; $p = .03$; $\eta^2 = .35$. This confirmed that the biggest difference for the accuracy means ($M =.20$) was in the typical items from categories that were trained with typical items.

Table 5-4 shows the mean response times for the untrained items in trained categories for the 12 participants.

Table 5-4. Group Response Time Means (ms) for Untrained Items in Trained Categories.

	Train-Atypical		Train-Typical	
	Pre	Post	Pre	Post
Atypical Items	1957 ms (235)	1575 ms (180)	1668 ms (156)	1255 ms (60)
Typical Items	1837 ms (194)	1756 ms (212)	1551 ms (101)	1535 ms (199)

Notes: (*SE* response times are in ms).

The pre and posttest response time means for Question 3a were analyzed with a t-test to determine if there was generalization to untrained items in trained categories. These results were not significant, $t(7) = 1.77$; $p = .12$; ($M_{Pre} = 1596$ ms, $SE = 154$ ms; $M_{Post} = 1430$, $SE = 156$ ms. It is important to note that for this analysis, only eight participants had usable response times for both pre and posttest, and these were the only data included in the analysis. As mentioned above, the criteria for usable response times were a correct response for the item name and a fluent response that was not preceded by other words, such as "It's a __". Consequently, there was missing data in four of the cells at pretest, and thus a reduction in the number of available data points for the response time analyses. This could have influenced the results.

Response time means for Question 3b, c, and d were analyzed with pre to post test difference scores calculated for each item type within each Training Type (e.g., for untrained typical items and untrained atypical items in each trained category). While Question 3b asked if there was a main effect of training type (e.g., Train-Typical), Question 3c focused on a main effect of item type (e.g., atypical items). Question 3c addressed the possibility of a two way interaction between Training type and Item type. To answer these questions, a (2) Training type x (2) Item type repeated measures ANOVA was computed. There were no significant main effects: for the Training type, $F(1,7) = .170$; $p = .69$; $\eta^2 = .02$; or for Item type, $F(1,7) = 2.294$; $p = .\ 17$; $\eta^2 = .25$. There was no interaction between Training type and Item type, F (1,7) = .72; p = .43; η^2 = .09. This lack of significance could be due to missing data as mentioned above or due to a lack of effect of training and item type on response times.

Overall, participants showed significant improvements in accuracy on the untrained items in trained categories, and this was evident in the main effect of time: there was higher accuracy after the training. Also, there were two interactions. The interaction between time and item type showed that there was more improvement for the untrained typical items compared to the untrained atypical items. From the three-way interaction among time, training type and item type, it was evident that at posttest, accuracy improved significantly in the untrained typical items from the Train-Typical category. Thus, at posttest, our participants had improved naming accuracy on untrained items in trained categories; especially on the typical items, in general, and more specifically, when they were in the category that had been trained with typical items. However, there was no concomitant generalization of training to response times.

Research Question 4: Category Generation

Is there generalization to category generation? Question 4 asked if there was generalization to another task, category generation, and if the type of training influenced the outcome. While the dependent variable was total number of items generated per category, the independent variables were time and training type. A (3) Training type (Train-Atypical/Train-Typical/Control) x (2) Time (per/post) repeated measures ANOVA was calculated. There was a significant main effect of time, $F(1,11) = 9.968$; $p < .01$; η^2 = .48. Figure 4-3 shows that a greater number of items were generated at posttest compared to at pretest (M_{Pre} = 7.16, SE = .72, versus M_{Post} = 8.61, SE = .83). There were no other significant findings. Training type had no significant main effect, $F(2,10) = .578$; p =

.578; $\eta^2 = .104$, nor was the interaction between training type and time significant, $F(2,10) = .186$; $p = .833$; $\eta^2 = .036$. These results suggested that the type of training did not influence the category generation performance at posttest. Figure 5-3 shows the number of items generated at pretest and posttest for the picture naming training conditions (i.e., Train-Atypical, Train-Typical, and Controls).

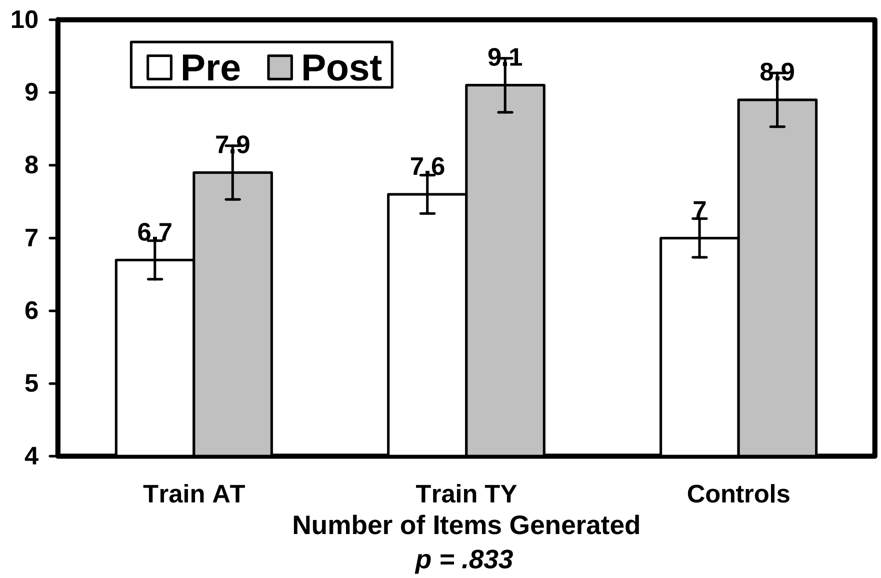

Figure 5-3. Category Generation, Number of Items Generated at Pre & Posttest for Trained & Control Items. Train-AT=Train-Atypical, Train-TY =Train Typical. *p*= .833.

Summary of Results

In summary, the group of adults with PAD performed picture naming and category generation tasks after participating in a semantic training by answering questions about typical and atypical items and then naming them. The participants repeated the correct name of most of the items after an initial attempt to name them during the pretest. Although the control items showed slight increases in accuracy and faster response times, these changes were not significant. The accuracy scores for trained items showed a facilitation effect from the semantic training (i.e., combined Train-Atypical and Train-

Typical conditions). Although the responses were faster for the trained items, this only approached significance. However, the difference scores for the response times did show a significant reduction for the Train-Typical items compared to the Train-Atypical items. The accuracy scores showed generalized naming to untrained items in trained categories, and a three-way interaction suggested that the participants also demonstrated better performance on the typical items, especially when they were in the Train-Typical condition, but this was only significant for the accuracy scores. In the category generation task, the PAD group also demonstrated a significant improvement in the number of the items generated from pre to posttest, but the type of training did not have an impact. These results are reviewed and interpreted in the next chapter.

CHAPTER 6
DISCUSSION

In this exploratory study, 12 adults recently diagnosed with early probable Alzheimer disease (PAD) individually completed a three phase session involving pretesting, semantic training, and postesting. The first phase involved category generation and picture naming. After an initial attempt at naming the picture, participants received performance feedback and repeated the name of all of the items, except for half of the items in the control condition. In the second phase, 14 items were trained, 7 atypical items (i.e., Train-Atypical) from one semantic category and seven typical items (i.e., Train-Typical) from another semantic category. Participants were required to answer questions about the target item and then name it. In the third phase, category generation and picture naming tasks were used again, except there was no feedback or repetition. Outcome measures were accuracy scores and response times for the picture naming and number of items produced for the category generation task.

The research questions addressed the effects of repetition, training, and generalization, as well as typicality. The two views were compared, one favoring typical items, based on the family resemblance view (Rosch, 1975) and the other favoring atypical items, based on the Complexity Account of Treatment Efficacy (Thompson et al., 2003) and a connectionist model of acquired dyslexia (Plaut, 1996), as well as adults with aphasia (Kiran & Thompson, 2003b; Stanczak et al., 2006) who had better generalization after training with atypical items. Interpretations of our data are related to the predictions from these views and theories of semantic representations in PAD

(Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997), as well as PAD anomia therapy studies (Abrahams & Camp, 1993; Fuller et al., 2001; Rothi et al., 2005) and some of the anomia treatment studies from the aphasia literature. It is important to note that there were many differences in these studies, including issues related to the patient population (i.e., PAD versus aphasia; aphasia classification; severity, and corresponding language deficits) and methodology (i.e., group versus single subject design; accuracy and response time measures versus just accuracy).

Summary of Findings

In this exploratory study, adults with PAD showed that repeating items in the control condition led to slight changes (i.e., increases in the accuracy scores and decreases in the response times) but these were not significant. However, the participants showed a facilitation effect in this single semantic training using typical and atypical semantic category exemplars in which they were required to generate information about the target item and than name it. Thus, at posttest, they showed a facilitation effect characterized by both significantly improved accuracy scores and faster response times on picture naming. Although accuracy and response times for both Train-Atypical and Train-Typical items changed, only the change in response times for the Train-Typical condition was significant. In terms of generalization to untrained items in trained categories, changes in accuracy scores were significant; specifically, there was significant improvement in the accuracy scores for the untrained typical items that were in the Train-Typical category. For the category generation task, the participants produced significantly more items in the category generation at posttest; however, this was not due to semantic training, since all categories were affected.

It is important to note that several comparisons showed relatively high effect sizes but non-significant results. This is likely due both to the small sample size and to the high variability among participants, which gave this study little power to detect subtle differences. Therefore, while the results are suggestive, a larger sample size is recommended for future studies. In the next section, implications of the findings are discussed.

Implications for Anomia Treatment in PAD

The findings from our group of adults with PAD are consistent with the family resemblance view (Rosch, 1975) and the hypothesis that training typical items results in generalization (Heilman, personal communication, 2005). Furthermore, the PAD group's performance is compatible with some (Plaut, 1996; Stanczak et al., 2006) but not all of the treatment effects from typicality training studies with aphasic speakers and some accounts of PAD (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997; Martin, 1992; Sailor et al., 2004; Smith et al., 1995).

Previous research has suggested possible reasons to account for better performance with typical items including the timeframe related to the initial phase of the training (Stanczak et al., 2006). These are discussed briefly and compared with our findings. This is followed by an alternative account of why our PAD group did better with the typical items.

To account for the performance of one of their participants with transcortical motor aphasia and semantic and phonological deficits, Stanczak et al. (2006) proposed that for such participants, typical and atypical items might be best treated in sequence. This was because over time, training the typical items may lead to confusion due to the competition among related typical items, at which point the atypical items, which provide more variation and thus have fewer competitors, are learned better compared to the typical items (Stanczak et al., 2006). A similar phenomenon was also noticed in the training of a connectionist network. Plaut (1996) reported that his computer simulation of acquired dyslexia had more difficulty learning to distinguish among highly typical category exemplars although ultimately the model learned the typical items better than the atypical items. The results of the current study support this hypothesis to some extent. We found that at pretest, responses to typical items were more accurate but much slower than responses to atypical items, which is consistent with there being some difficulty distinguishing between items. It is certainly possible that continued intense training of typical items could lead to confusion among them (Stanczak et al., 2006). This has also been described as a semantic blocking effect or interference: repeatedly accessed words with similar semantics become more difficult to access the more frequently they must be named (Schnur, Schwartz, Brecher, & Hodgson, 2006). At that point, the greater interitem distinguishability among the atypical items might benefit our participants (Stanczak et al., 2006). However, this is an empirical question that requires investigation using a more standard treatment study methodology.

It is important to note that these typicality training studies were conducted in systems that had damage due to focal lesions, either acquired dyslexia in a connectionist model (Plaut, 1996) or aphasia in adults (Kiran & Thompson, 2003b; Mayer et al., 2004; Stanczak et al., 2005, 2006). Gonnerman et al. (1997) reminds us that PAD does not result in exclusive damage to an area, instead, it is a diffuse pattern of damage. Thus, it might be unreasonable to expect similar outcomes from training typical and atypical category exemplars in adults with PAD. Indeed, the results of the current study are consistent with previous studies examining the effects of PAD on semantic tasks, and can be accounted for by existing hypotheses about the effects of PAD on the semantic system (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997).

The primary findings of the current study were that training effects were strongest among trained typical items, manifesting as faster response times, and only typical items from categories in which typical items were named more accurately, a generalization effect. Further, the training encouraged participants to answer questions that emphasized distinguishing features of the target items, as suggested by Hillis (1998) and Nadeau and Rothi (2004). Our findings are consistent with those reported in Sailor et al. (2004) who found that individuals with PAD generated fewer atypical than typical items. Smith et al. (1995) reported that a group of adults with PAD performed poorly when verifying features of low typical (and low dominance, a similar construct related to importance), suggesting that information about the low typical items had deteriorated. However, Smith et al. (1995) found that both shared and distinguishing features were affected.

The differential effects of PAD on the accessibility of shared versus distinguishing features can account for the findings in this study. Earlier, typical items were described as sharing more features with others in the category compared to atypical items (Rosch, 1975; Rosch & Mervis, 1975). When sets of features co-occur in many items in a category, they are frequently activated together, and, consequently develop strong connections with each other (Collins & Loftus, 1975; Hebb, 1949). This highly interconnected network provides strong redundancy in the channels through which activation can flow among these connected features, which allows alternate pathways for activation of features, even when there is damage to some of these connections. Thus, when there is damage to connections among the shared features, the behavioral manifestations are not severe because activation can still be achieved by the existing conditions, although it might not be as efficient (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997; McRae, de Sa, & Seidenberg, 1997). The difficulty with items whose representations consist primarily of shared features, however, comes from trying to distinguish one from the other. Distinguishing features are, by definition outside the set core of shared features of a category and serve to distinguish one item from another. Distinguishing features occur in only one or possibly a few items within a category, and, thus, have relatively fewer connections that link them to the shared feature set. Consequently, they are not highly interconnected with other features and have no redundant or alternative pathways for activation to compensate for damage (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997). As a result, damage to the connections to distinguishing features increases the possibility that it will be inaccessible, whereas similar damage to connections between shared features is more likely to be compensated for (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997; Martin, 1992; Sailor et al., 2004; Smith et al., 1995).

In the earlier discussion of the semantic representations of atypical items, they were described as having a lower proportion of shared features and higher proportion of distinguishing features than typical items. As just discussed, this would make atypical items particularly vulnerable to the diffuse damage of PAD. Furthermore, due to the relative lack of connections to distinguishing features and to the lack of connections between distinguishing features, atypical items might be more difficult to remediate.

In the current study, we suspect that the semantic training used, as well as the emphasis on training distinguishing features, facilitated access to other typical items by raising the overall activation levels of the core semantic features in the category. Due to the constant stimulation of the connections between the shared features, it likely strengthened their connectivity by raising their activation thresholds (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997). Moreover, we also emphasized distinguishing features for the typical items, allowing participants to better discriminate among them.

Atypical items might not have benefited as much as the typical items for a number of reasons. According to the hypothesis, the semantic representations of atypical items contain fewer of the shared features of the category, and there is no guarantee that the shared features included in the representation of atypical item 1 are the same shared features included in the representation of atypical item 2. Thus, the degree of feature overlap among atypical items might have been minimal, so there could be no benefit of spreading activation from other trained items in the category. Alternatively, it is possible that the connections to the distinguishing features were not merely just below threshold activation but were damaged and inaccessible, possibly due to infrequency of use as suggested in Smith et al. (1995).

This discussion has shown that the findings of the current study are consistent with previous findings on the performance of individuals with PAD on semantic tasks, and can be well accounted for by existing theories of semantic representations and their degradation in PAD (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997; Martin, 1992; Sailor et al., 2004; Smith et al., 1995). Extending these theories to the creation of a possible intervention for anomia in PAD has led to promising results supporting the use of a combination of semantic treatment with word repetition for this population. Whether atypical items will be remediable in this population, or whether there may even be an advantage for atypical items after further treatment is an open question to be addressed in future studies.

Implications for Methodology

There are several methodological factors that might have influenced the outcomes of our study relative to other studies investigating the effects of training typical versus atypical items. As discussed above, some of these include the population being trained (e.g., connectionist networks, adults with different types of aphasia or PAD) and type and severity of the damage (Mayer et al., 2004; Stanczak et al., 2005, 2006), as well as the typicality of the items (Kiran & Thompson, 2003b; Stanczak et al., 2005, 2006) and the timeframe (Stanczak et al., 2006). Other differences include the type of study (e.g., facilitation versus treatment) and the study design (single-subject, versus group), as well as the procedure and training tasks. The group design used here is very different from the majority of PAD and aphasia studies discussed above. There are numerous advantages of the single-subject design, such as, the careful monitoring of treatment outcomes in relationship to the targeted behavior (Kearns, 2000). Group data does not capture the nuances of individual subject performance as well as a single subject design. However, the goal of the current study was to determine if a facilitation effect (Howard, 1985) could be elicited after this single training, which was a question more suited to group design. Knowing that a facilitation effect is possible in individuals with PAD should encourage researchers to begin testing more in-depth treatments.

The combination of the repetition (i.e., lexical information) and the semantic training appeared to be beneficial for our PAD group. Repetition over a short period was not effective by itself. Yet a treatment effect in the PAD anomia study by Fuller et al. (2001) and Rothi et al. (2005) was found after many training sessions with repetition in an errorless learning paradigm. Similarly, Abrahams and Camp (1993) reported that the spaced retrieval technique, which also involved repetition, resulted in a treatment effect. The possibility of generation effects (Slamecka & Graf, 1978) has been raised. According to Mitchell et al. (1986) the generation effect has been attributed to persistence of activation of previously accessed information within semantic memory. Generation effects have been documented in some PAD studies (Fleischman et al., 1997; Multhaup & Balota, 1997) but not in others (Dick, Kean, & Sands, 1989; Mitchell et al., 1986). Our study found somewhat increased accuracy and faster response times in the unrepeated control items, but these were not significant. These findings support the use of interventions for anomia in PAD that include both a repetition element and a semantic element to address the two common loci of word finding difficulties in this population.

The semantic training task could have also been a factor influencing the performance of our PAD group. The questions were specific to the particular categories and effective but they might not have been as powerful as the feature selection task used for the fluent participants in the Kiran and Thompson (2003b) study. These participants saw a board containing a large number of semantic features pertaining to the category. While the set of defining features were shared by all of the typical category exemplars from that category, another set was comprised of characteristic features, such that some were more relevant to typical items while other features covered a broader spectrum of information pertinent to atypical items. Consequently, when participants were asked to choose appropriate features they were exposed to a large set of features appropriate to an item. The task demands included choosing not only the features that were appropriate to the item, but also rejecting those that were not relevant to the item. It is possible that this, combined with the other training tasks, and the atypical items themselves, allowed the essential differences between training atypical and typical items to be revealed and have their impact on items in a measurable way, in favor of atypical items. This variation of a semantic training task has only been used by Kiran and Thompson (2003b), and it differs considerably from the other semantic training methods used in the other typicality training studies (Mayer et al., 2004; Plaut, 1996; Stanczak et al., 2005, 2006). It is also different from having to self-generate characteristics of an item, as in our study.

We suspect that our training method did not provide as robust a comparison between items to allow participants to hone in on the broad distinctions between atypical and typical items (Kiran & Thompson, 2003b). Instead, our method may have put the focus on the individual items themselves. The PAD population might benefit from using the feature selection task, but modified to present fewer pieces of information to avoid overstimulation and confusion from the simultaneously presented printed information.

Another factor that might have influenced the performance of our PAD group is the choice of semantic categories. At least two studies have shown generalization following training with natural kinds (vegetables and birds) (Kiran & Thompson, 2003b; Stanczak et al., 2006). These studies demonstrated evidence for better generalization after training atypical items. However, we used artifact categories for training, which has been shown to have more functional information than perceptual features found in natural kinds

(Warrington & Shallice, 1984). Thus, artifacts might have very different internal structures (e.g., for vehicles, there are many subtypes, e.g., 2-wheel, 4-wheel, floating, flying) than the more hierarchically organized natural kind categories, like birds and vegetables (Devlin et al., 1998; Gonnerman et al., 1997). Clearly, future studies should contrast the effects of training natural kinds and artifacts.

It is difficult to know if training two categories simultaneously with two different training types (i.e., Train-Atypical, Train-Typical) provided an advantage or not. Both Stanczak et al. (2006) and Mayer et al. (2004) provided training for two categories at the same time and found a treatment effect. Mayer et al (2004) used a combination of artifacts and natural kinds; they attributed the lack of generalization of training in their participants to the severity of their aphasias. Stanczak et al. (2006) trained birds and vegetables simultaneously, in blocks, and found treatment effects, but only one of two participants showed generalization (i.e., to untrained typical birds after being trained on atypical birds and marginally generalized naming of untrained atypical vegetables after typical vegetables were trained). However, Kiran and Thompson (2003b) showed successful outcomes (i.e., treatment effect and generalization after being trained on atypical items) with adults with fluent aphasia when only training one category (i.e., birds or vegetables) at a time with atypical or with typical exemplars. Considering these variable results across studies, it is difficult to determine if the number and type of trained categories accounted for any differences in the outcomes.

A lack of power was certainly an issue, especially for the items in the control condition (i.e., untrained items in untrained categories). The effects of the small sample size were likely exaggerated by variability among the participants in variables such as age, education, and severity of impairment (despite being recently diagnosed). This variability is causally related to the challenges of identifying and recruiting the participants with mild-moderate PAD and completing data collection (e.g., the 12 participants with PAD took about four months). There might also have been effects related to other factors such as mental fatigue, learning curves, and practice schedules. Although the clinician offered opportunities for breaks between the pretesting and training, and before the posttesting, it is possible that more time was needed for these breaks in order to fully benefit from the training. Anecdotally, several of the participants indicated that they were tired at the posttest, yet they still enthusiastically expressed interest in and completed their participation in the study without incident. The wide range of differences among our study and the other typicality studies warrant caution in generalizing the findings. However, despite these methodological challenges, the data has provided meaningful information.

Future Studies

The findings from our study provide a good foundation for further development of this protocol into an anomia treatment program for PAD. Various issues have been raised and will be considered in our future anomia treatment studies. First, generalization should be programmed prior to starting the study (McNeil et al., 1998). We also recommend continuing the use of generalization probes for untrained atypical items and untrained typical items so that the type of training can be compared with the item type. Again, performance should be compared across conditions, but with a category in which there is no repetition. This would allow for a comparison between repetition with semantic training versus just semantic training.

In addition, sessions should be scheduled for the screening, and pre and posttesting as well as a follow-up maintenance session. The testing should continue to use the Mini Mental Status Exam (Folstein et al., 1975) but perhaps with more strict criteria (18-24). Additional cognitive-linguistic assessments should be used: the Boston Naming Test (Kaplan et al. 1983, 2001), for a standardized measure of confrontation picture naming, the Hopkins Verbal Learning Test (Brandt, 1991) to measure memory and new learning, and perhaps the Western Aphasia Battery (Kertesz, 1982) for an overall standardized language screening. A pretest picture naming with the 72 items should also be used as a screening to determine eligibility for the study perhaps with a cutoff score to identify treatment items based on failed naming attempts at pretest. At the onset of the study, a category verification task could be conducted (Stanczak et al., 2005, 2006). Perhaps a few categories and their corresponding exemplars that are not part of the naming task could be used.

The training should be extended into a treatment in which the participant attends several sessions per week, for several weeks, for example three one hour sessions per week for three weeks. To insure that the participant can correctly name the item, the treatment protocol could start with the errorless naming (i.e., immediate repetition) paradigm (Fuller et al., 2001; Rothi et al., 2005). This could be followed by the current training (i.e., answer the three questions and name it). Additional suggestions are to include categories that are meaningful to the participant (Mayer et al., 2004). Natural kinds should also be included to further examine some of the hypotheses related to artifacts (Warrington & Shallice, 1984) and intercorrelated features compared to distinguishing features in PAD (Altmann et al., 2001; Devlin et al., 1998; Gonnerman et al., 1997). A modified semantic feature task based on Kiran and Thompson (2003b) should be considered, to emphasize the broad differences between the typical and atypical items in the categories. Semantic feature distinction with the target items that the participant had difficulty with could also be incorporated into the intervention (Hillis, 1998) to compare and contrast the features. Family members or caregivers could be involved and trained to assist with a home program. Perhaps with additional training and over a longer period of time, some of these unresolved questions will be answered and the current protocol will be transformed into a more potent vehicle for improving naming deficits in adults with PAD.

In summary, in a period of less than three hours, we found that adults with early PAD respond well and quickly to a semantic training program. Following training, the PAD participants showed an advantage for typical category exemplars, first in faster response times (i.e., a facilitation effect from trained items) and then in accuracy scores (i.e., generalization to untrained items categories trained with typical exemplars). These preliminary findings suggest training typical items produced the best results, although this advantage of typical over atypical items may change over time (Stanczak et al., 2006). Overall, our findings not only provide support for the family resemblance (Rosch, 1975) view, and the prediction by Heilman (personal communication, 2005) that typical items would show better generalization, but also suggest that in PAD, Hebbian learning (Hebb, 1949) is possible. Another important finding is that that there was no generalization to untrained categories, which is highly important for clinicians involved in semantic training of all populations. These findings are extremely encouraging for the development of principled strategies for choosing items to encourage generalization in the remediation of anomia, as well as for the development of lexical-semantic treatment paradigms for individuals with early PAD.

APPENDIX A
INFORMED CONSENT FORM

Informed Consent

Choosing Remediation Targets for Naming Deficits: Effects of Typicality

Please read this consent document carefully before you decide to participate in this study.

Purpose of the research project
The purpose of this study is to test two potential treatments for the naming impairment in Alzheimer disease. Both treatments are methodologically identical, varying only in the type of stimuli used.

What you will be asked to do in the study
In this study, you will be asked to complete two tasks for the screening phase by answering general knowledge questions, and reading words aloud. For the pre-testing there will be two tasks. First, the examiner will name each of the three semantic categories and ask you to name as many items as you can that belong to each. Second, the examiner will present pictures and ask you to name them, and then to repeat the correct name of the picture. For the training phase the examiner will present the training pictures and ask you three questions about the item. Then you will be asked to name the picture. For the post-testing, you will be asked to do the same tasks as you did for the pre-testing (name as many items as you can in three categories, and name pictures but without repeating the correct name of the picture). Responses on this task will be digitally recorded and transcribed for later analysis.

During the time that you are participating in the study, we will offer brief breaks, and you can ask for a longer break at any time.

All testing will be carried out by the principal investigator, Claudia A. Morelli, Dr. Altmann, or trained research assistants.

Time required
We anticipate it will take 2-3 hours for the participants with Alzheimer disease to complete the protocol and about 2 hours for the healthy participants.

Risks and Benefits
Participating in this study may help people who have difficulties with word finding. The risk is minimal, but it is possible that participants will become fatigued or bored, during the course of the testing and training. These will be minimized by offering breaks as needed.

Compensation
You will be compensated for participating in this study ($20.00 for people with Alzheimer disease, and $15.00 for older adults).

Confidentiality
Your identity will be kept confidential to the extent provided by law. Your information will be assigned a code number. The list connecting your name to this number will be kept in a locked file in the Language over the Lifespan lab. When the study is completed and the data have been analyzed, the list will be destroyed. Your name will not be used in any report.

Voluntary participation
Your participation in this study is completely voluntary. There is no penalty for not participating.

Right to withdraw from the study
You have the right to withdraw from the study at any time.

Whom to contact if you have questions about the study
Claudia A. Morelli, M.S., CCC-SLP, Ph.D. Candidate, Communication Sciences and Disorders Department, P.O. Box 117420, 336 Dauer Hall, University of Florida, Gainesville, FL 32611-7420. cmorelli@ufl.edu 352-392-2113 X232.

OR

Lori J. Altmann, Ph.D., Communication Sciences and Disorders Department, P.O. Box 117420, 336 Dauer Hall, University of Florida, Gainesville, FL 32611-7420
laltmann@ufl.edu. 352-392-2113 x279

Whom to contact about your rights as a research participant.
UFIRB Office, box 112250, University of Florida, Gainesville, FL 32611-2250 352-392-0433.

Agreement
I have read the procedure described above. I voluntarily agree to participate in the procedure and have received a copy of this description.

Participant: ______________________________________ **Date:**__________

Principal Investigator: ________________________________ **Date:**__________

APPENDIX B
TABLES WITH INFORMATION ABOUT THE PARTICIPANTS

Participant Demographics and Selected Neuropsychological Test Results

Table B-1. Participant Demographics and Selected Neuropschological Testing Results from the PAD Diagnosis.

ID #	Age	Years of Education	Gender	MMSE Total Score (Max 30)	MMSE Recall (Max 3 items)	Boston Naming Test	Hopkins Verbal Learning Test: Free Recall (Max 36)	Hopkins Verbal Learning Test: Delayed Recall (Max 12)
1001	72	17	F	N/A	N/A	N/A	N/A	N/A
1002	77	13	F	20/30	N/A	N/A	N/A	N/A
1003	76	10	M	18/30	0/3	10/30	9/36	0/12
1004	92	24	M	19/30	0/3	N/A	9/36	0/12
1005	82	12	F	19/30	0/3	11/19	4/36	0/12
1006	78	12	F	22/30	0/3	14/15	15/36	0/12
1007	55	20	M	25/30	0/3	15/15	20/36	0/12
1008	84	15	F	23/30	0/3	15/15	16/36	0/12
1009	69	18	M	24/30	2/3	15/15	12/36	N/A
1010	86	16	F	21/30	0/3	11/15	10/36	0/12
1011	75	16	M	N/A	N/A	N/A	N/A	N/A
1012	85	16	F	24/30	2/3	24/30	12/36	0/12

Notes: N/A: Scores not available; MMSE= Mini Mental Status Exam (Folstein et al., 1975); Hopkins Verbal Learning Test (Brandt, 1991); Boston Naming Test (Kaplan et al., 1983, 2001).

Demographics and Performance Before and After Semantic Training

Table B-2. Participant Demographics and Pre and Posttest Category Generation and Picture Naming Tasks (i.e., Before and After the Semantic Training, Day of Participation in the Study).

ID Number	Age	Years of Education	Gender	Pretest Category Generation	Posttest Category Generation	Pretest Picture Naming (Proportion)	Pretest Picture Naming (Max=72)	Posttest Picture Naming (Proportion)	Posttest Picture Naming (Max=72)
1001	72	17	F	25	25	.74	53/72	.83	60/72
1002	77	13	F	23	23	.81	58/72	.83	60/72
1003	76	10	M	15	15	.71	51/72	.78	56/72
1004	92	24	M	14	14	.51	37/72	.57	41/72
1005	82	12	F	18	17	.60	43/72	.58	42/72
1006	78	12	F	30	36	.81	58/72	.81	58/72
1007	55	20	M	24	34	.89	64/72	.94	68/72
1008	84	15	F	27	37	.83	60/72	.86	60/72
1009	69	18	M	20	26	.81	58/72	.83	60/72
1010	86	16	F	23	33	.76	55/72	.82	59/72
1011	75	16	M	24	31	.83	60/72	.96	69/72
1012	85	16	F	19	22	.60	43/72	.72	52/72
Group Mean	77.6	15.8		21.8	26.1	.74	53/72	.79	57/72

Notes: Category Generation is the total across the 3 categories (vehicles, clothing and tools); Picture Naming is proportion correct, and Total Correct, Maximum=72.

APPENDIX C
TABLES WITH INFORMATION ABOUT THE STIMULI

Master List of Stimulus Items Used in the Study

Table C-1. Master List of Stimulus Items Used in the Study. *Trained Items are Italicized*

Number	Item Name	Typicality Rank	Typicality Score	Written Word Frequency	Familiarity Rating	Phoneme Length
TY 1	Automobile	1	1.02	74	456	8
2	station wagon	2	1.14	0	-	11
3	*truck*	3	1.17	80	620	4
4	bus	5.5	1.27	42	-	3
5	taxi	5.5	1.27	19	-	5
6	jeep	7	1.35	16	564	3
7	motorcycle	9	1.65	0	-	8
8	van	11	1.95	2	542	3
9	*train*	14	2.15	86	548	4
10	*bicycle*	16	2.51	7	-	6
11	*airplane*	18	2.64	21	-	6
12	*boat*	20	2.75	123	584	3
AT 13	ship	22	2.82	126	553	3
14	*scooter*	23	3.24	1	468	5
15	tractor	24	3.3	31	518	6
16	*wagon*	25	3.31	72	443	5
17	*wheelchair*	29	3.68	0	-	6
18	tank	31	3.84	30	511	4
19	*rowboat*	33	3.92	0	-	5
20	*tricycle*	35	4	0	436	7
21	canoe	36	4.01	8	441	4
22	raft	37	4.37	5	483	4
23	*submarine*	38	4.51	35	450	7
24	*blimp*	42	4.81	1	-	5
TY 25	*pants*	1	1.12	9	575	5
26	*shirt*	2.5	1.14	29	612	3
27	*dress*	2.5	1.14	63	588	4
28	skirt	4	1.21	22	551	4
29	blouse	5	1.27	2	562	4
30	*suit*	6	1.49	64	543	3
31	jacket	8	1.68	39	596	5
32	*coat*	9	1.88	52	610	3
33	*sweater*	10	1.89	18	-	5
34	*socks*	16	2.08	7	-	4
35	parka	17	2.19	0	-	5
36	pajamas	18	2.25	3	-	7
AT 37	bathing suit	22.5	2.44	0	-	8
38	*bath robe*	24	2.65	3	-	6
39	*shoes*	27	2.73	44	-	3
40	*vest*	30	2.81	4	472	4
41	boots	33	3.42	1	-	4

Table C-1. Continued

Number	Item Name	Typicality Rank	Typicality Score	Written Word Frequency	Familiarity Rating	Phoneme Length
42	sandals	34	3.56	5	-	6
43	*tie*	35	3.71	27	559	2
44	*belt*	37	3.93	36	550	4
45	scarf	38	3.96	4	-	5
46	*mittens*	39	3.99	2	-	5
47	*hat*	40	4.08	71	580	3
48	gloves	43	4.53	6	-	5
TY 49	*saw*	1	1.04	8	552	2
50	*hammer*	2	1.34	6	515	4
51	ruler	3	1.48	13	571	4
52	*screwdriver*	4	1.56	1	544	9
53	*drill*	5	1.59	21	473	4
54	*nails*	6	1.67	20	-	4
55	tape measure	7	1.69	0	-	7
56	sawhorse	8	1.77	1	-	6
57	*level*	11	1.82	50	504	4
58	toolbox	14	2.12	1	-	7
59	T-square	15	2.22	0	-	7
60	*chisel*	16	2.26	5	469	4
AT 61	pliers	22	2.59	1	499	5
62	*wrench*	23	2.6	1	-	4
63	*ladder*	24	2.64	19	507	4
64	*vise*	25	2.76	1	368	3
65	*screws*	26	2.77	10	-	5
66	*awl*	33	3.09	0	257	2
67	*crowbar*	34.5	3.12	0	-	6
68	bolts	40	3.63	1	-	5
69	paintbrush	45	3.81	1	-	8
70	stapler	48	4.21	0	-	6
71	*axe*	53	4.53	19	461	3
72	scissors	59	5.36	1	559	5

Notes: TY=Typical items; AT=Typical items. Typicality Rank and Score ; Written Word Frequency ; Familiarity .

Alternate Names for Stimulus Items Used in the Study

Table C-2. Master List of Stimulus Items Used in the Study. Alternate names for stimuli counted as correct.

Number	Item Name	Alternate Names Counted as Correct
TY 1	Automobile	Car, sedan
2	station wagon	Volvo
3	*truck*	Pick-up, semi, 18 wheeler; semi-truck, tractor trailer
4	bus	
5	taxi	Taxi cab, cab
6	jeep	
7	motorcycle	
8	van	
9	*train*	
10	*bicycle*	Bike
11	*airplane*	Plane, jet
12	*boat*	Tugboat
AT 13	ship	Cruiseliner, oceanliner, steamship
14	*scooter*	Moped
15	tractor	
16	*wagon*	
17	*wheelchair*	
18	tank	
19	*rowboat*	
20	*tricycle*	
21	canoe	
22	raft	
23	*submarine*	
24	*blimp*	Derigible, zeppelin
TY 25	*pants*	
26	*shirt*	Dress shirt
27	*dress*	
28	skirt	
29	blouse	
30	*suit*	
31	jacket	
32	*coat*	Winter coat
33	*sweater*	Turtleneck sweater, turtleneck
34	*socks*	
35	parka	
36	pajamas	
AT 37	bathing suit	Swimsuit
38	*bath robe*	Robe
39	*shoes*	
40	*vest*	
41	boots	
42	sandals	
43	*tie*	Necktie
44	*belt*	
45	scarf	
46	*mittens*	
47	*hat*	Fedora
48	gloves	

Table C-2. Continued

Number	Item Name	Alternate Names Counted as Correct
TY 49	*saw*	
50	*hammer*	
51	ruler	
52	*screwdriver*	
53	*drill*	Electric drill
54	*nails*	
55	tape measure	Measuring tape
56	sawhorse	
57	*level*	
58	toolbox	Toolkit
59	T-square	
60	*chisel*	
AT 61	pliers	
62	*wrench*	
63	*ladder*	
64	*vise*	
65	*screws*	
66	*awl*	
67	*crowbar*	
68	bolts	
69	paintbrush	
70	stapler	
71	*axe*	
72	scissors	

Notes: TY=Typical items; AT=Typical items.

LIST OF REFERENCES

Abeysinghe, S. C., Bayles, K. A., & Trosset, M. W. (1990). Semantic memory deterioration in Alzheimer's subjects: Evidence from word association, definition, and associate ranking tasks. *Journal of Speech and Hearing Research, 33*(3), 574-582.

Abrahams, J. P., & Camp, C. J. (1993). Maintenance and generalization of object naming training in anomia associated with degenerative dementia. *Clinical Gerontologist, 12*(3), 57-72.

Alathari, L., Trinh Ngo, C., & Dopkins, S. (2004). Loss of distinctive features and a broader pattern of priming in Alzheimer's Disease. *Neuropsychology, 18*(4), 603-612.

Albert, M. S., & Milberg, W. (1989). Semantic processing in patients with Alzheimer's disease. *Brain and Language, 37*(1), 163-171.

Altmann, L. J., Kempler, D., & Andersen, E. S. (2001). Speech errors in Alzheimer's disease: Reevaluating morphosyntactic preservation. *Journal of Speech, Language, and Hearing Research, 44*(5), 1069-1082.

Anderson, J. (1991). The adaptive nature of human categorization. *Psychological Review, 98*(3), 409-429.

Appell, J., Kertesz, A., & Fisman, M. (1982). A study of language functioning in Alzheimer patients. *Brain and Language, 17*(1), 73-91.

ASHA. (American Speech-Language-Hearing Association). (In press). Roles of speech-language pathologists in the identification, diagnosis, and treatment of individuals with cognitive-communication disorders: Postion statement. *ASHA Supplement, 25*.

Au, R., & Bowles, N. (1991). Memory influences on language in normal aging. In D. Ripich (Ed.), *Handbook of geriatric communication disorders* (pp. 293-305). Austin: Pro-Ed, Inc.

Bandur, D. L., & Shewan, C. M. (2001). Language-Oriented Treatment: A psycholinguistic approach to aphasia. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (pp. 629-662). Philadelphia: Lippincott, Williams & Wilkins.

Barbarotto, R., Capitani, E., Jori, T., Laiacona, M., & Molinari, S. (1998). Picture naming and progression of Alzheimer's disease: An analysis of error types. *Neuropsychologia, 36*(5), 397-405.

Barker, M. G., & Lawson, J. S. (1968). Nominal aphasia in dementia. *British Journal of Psychiatry, 114*(516), 1351-1356.

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory and Cognition, 15*(5), 397-418.

Bayles, K. A. (1982). Language function in senile dementia. *Brain and Language, 16*(2), 265-280.

Bayles, K. A., Kasniak, A. W., & Tomoeda, C. K. (1987). *Communication and cognition in normal aging and dementia*. Austin, TX: Pro-Ed, Inc.

Bayles, K. A., & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and Language, 19*(1), 98-114.

Bayles, K. A., Tomoeda, C. K., & Cruz, R. F. (1999). Performance of Alzheimer's disease patients in judging word relatedness. *Journal of the International Neuropsychological Society, 5*(7), 668-675.

Bell, E. E., Chenery, H. J., & Ingram, J. C. (2001). Semantic priming in Alzheimer's dementia: Evidence for dissociation of automatic and attentional processes. *Brain and Language, 76*(2), 130-144.

Benson, D. F. (1979). Neurologic correlates of anomia. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in Neurolinguistics* (Vol. 4). New York: Academic Press.

Benson, D. F., & Geschwind, N. (1985). Aphasia and related disorders: A clinical approach. In M. M. Mesulam (Ed.), *Principles of Behavioral Neurology* (pp. 193-238). Philadelphia: F.A. Davis.

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia, 6*, 53–60.

Bowles, N. L., Obler, L. K., & Albert, M. L. (1987). Naming errors in healthy aging and dementia of the Alzheimer type. *Cortex, 23*(3), 519-524.

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology, 13*(3), 236-249.

Boyle, M., & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*, 94-98.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist, 5*, 125-142.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. What do various measures measure? *Brain, 113*, 397-417.

Chertkow, H., Bub, D., & Caplan, D. (1992). Constraining theories of semantic memory processing: Evidence from dementia. *Cognitive Neuropsychology, 9*(4), 327-365.

Chertkow, H., Bub, D. N., & Seidenberg, M. S. (1989). Priming and semantic memory loss in Alzheimer's disease. *Brain and Language, 36*(3), 420-446.

Chieffi, S., Carlomagno, S., Silveri, M. C., & Gainotti, G. (1989). The influence of semantic and perceptual factors on lexical comprehension in aphasic and right brain-damaged patients. *Cortex, 25*(4), 591-598.

Chui, H. C. (1989). Dementia. A review emphasizing clinicopathologic correlation and brain-behavior relationships. *Archive of Neurology, 46*(7), 806-814.

Cobb, R. E. (2005). *The bottom-up deterioration hypothesis of the semantic impairment associated with Alzheimer's disease: The effect of typicality on object verification.* University Microfilms International.

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology, 14*(2), 133-142.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Cronin-Golomb, A., Keane, M. M., Kokodis, A., Corkin, S., & Growdon, J. H. (1992). Category knowledge in Alzheimer's disease: Normal organization and a general retrieval deficit. *Psychology and Aging, 7*(3), 359-366.

Davis, A., & Pring, T. (1991). Therapy for word-finding deficits: More on the effects of semantic and phonological approaches to treatment with dysphasic patients. *Neuropsychological Rehabilitation, 1*(2), 135-145.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*(1-3), 287-314.

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience, 10*(1), 77-94.

Dick, M. B., Kean, M. L., & Sands, D. (1989). Memory for internally generated words in Alzheimer-type dementia: Breakdown in encoding and semantic memory. *Brain & Cognition, 9*(1), 88-108.

Drew, R. L., & Thompson, C. K. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech, Language, and Hearing Research, 42*(4), 972-989.

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related brain potentials methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed.). Cambridge: Cambridge University Press.

Fleischman, D. A., Gabrieli, J. D., Rinaldi, J. A., Reminger, S. L., Grinnell, E. R., Lange, K. L., et al. (1997). Word-stem completion priming for perceptually and conceptually encoded words in patients with Alzheimer's disease. *Neuropsychologia, 35*(1), 25-35.

Flicker, C., Ferris, S. H., Crook, T., & Bartus, R. T. (1987). Implications of memory and language dysfunction in the naming deficit of senile dementia. *Brain and Language, 31*(2), 187-200.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

Ford, J. M., Askari, N., Mathalon, D. H., Menon, V., Gabrieli, J. D., Tinklenberg, J. R., et al. (2001). Event-related brain potential evidence of spared knowledge in Alzheimer's disease. *Psychology and Aging, 16*(1), 161-176.

Frances, N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Francis, D. R., Clark, N., & Humphreys, G. W. (2002). Circumlocution-induced naming (CIN): A treatment for effecting generalisation in anomia?, *Aphasiology* (Vol. 16, pp. 243-259): Psychology Press (T&F).

Freed, D. B., Marshall, R. C., & Nippold, M. A. (1995). Comparison of personalized cueing and provided cueing on the facilitation of verbal labeling by aphasic subjects. *Journal of Speech and Hearing Research, 38*(5), 1081-1090.

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology, 29*(1), 65-75.

Fuller, R. H., Kendall, D. L., Nadeau, S. E., Spevack, A. A., Heilman, K., & Gonzalez Rothi, L. J. (2001). Cognitive-cholinergic therapy of anomia in Alzheimer's disease. *Neurology, 56*(Supplement 3), A57.

Gainotti, G. (1981). The relationship between type of naming error and semantic-lexical discrimination in aphasic patients. *Cortex, 17*(3), 401-409.

Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language, 93*(1), 79-94.

Gerratt, B. R., & Jones, D. (1987). Aphasic performance on a lexical decision task: Multiple meanings and word frequency. *Brain and Language, 30*(1), 106-115.

Gierut, J. A. (2001). Complexity in Phonological Treatment: Clinical Factors. *Language, Speech, & Hearing Services in Schools, 32*(4), 229.

Giffard, B., Desgranges, B., Nore-Mary, F., Lalevee, C., Beaunieux, H., de la Sayette, V., et al. (2002). The dynamic time course of semantic memory impairment in Alzheimer's disease: Clues from hyperpriming and hypopriming effects. *Brain, 125*(Pt 9), 2044-2057.

Giffard, B., Desgranges, B., Nore-Mary, F., Lalevee, C., de la Sayette, V., Pasquier, F., et al. (2001). The nature of semantic memory deficits in Alzheimer's disease: New insights from hyperpriming effects. *Brain, 124*(Pt 8), 1522-1532.

Gonnerman, L. M., Andersen, E. S., Devlin, J. T., Kempler, D., & Seidenberg, M. S. (1997). Double dissociation of semantic categories in Alzheimer's disease. *Brain and Language, 57*(2), 254-279.

Grober, E., Buschke, H., Kawas, C., & Fuld, P. (1985). Impaired ranking of semantic attributes in dementia. *Brain and Language, 26*(2), 276-286.

Grober, E., Perecman, E., Kellar, L., & Brown, J. (1980). Lexical knowledge in anterior and posterior aphasics. *Brain and Language, 10*(2), 318-330.

Grossman, M. (1981). A bird is a bird is a bird: Making reference within and without superordinate categories. *Brain and Language, 12*(2), 313-331.

Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M., et al. (2004). What's in a name: Voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain, 127*, 628-649.

Grossman, M., Robinson, K., Biassou, N., White-Devine, T., & D'Esposito, M. (1998). Semantic memory in Alzheimer's disease: Representativeness, ontologic category, and material. *Neuropsychology, 12*(1), 34-42.

Hampton, J. A. (1995). Testing the Prototype Theory of concepts. *Journal of Memory & Language, 34*(5), 686-708.

Hartman, M. (1991). The use of semantic knowledge in Alzheimer's disease: Evidence for impairments of attention. *Neuropsychologia, 29*(3), 213-228.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: John Wiley & Sons, Inc.

Heilman, K. M. (2005). *Aphasic disorders associated with dementia.* Paper presented at the 31st Annual Course in Behavioral Neurology and Neurolopsychology, Lake Buena Vista, FL.

Henderson, V. W., & Finch, C. E. (1989). The neurobiology of Alzheimer's disease. *Journal of Neurosurgery, 70*(3), 335-353.

Henderson, V. W., Mack, W., Freed, D. M., Kempler, D., & Andersen, E. S. (1990). Naming consistency in Alzheimer's disease. *Brain and Language, 39*(4), 530-538.

Hickin, J., Best, W., Herbert, R., Howard, D., & Osborne, F. (2001). Treatment of word retrieval in aphasia: Generalization to conversational speech, *International Journal of Language & Communication Disorders* (Vol. 36, pp. 13): Royal College of Speech & Language Therapists.

Hickin, J., Best, W., Herbert, R., Howard, D., & Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology, 16*(10/11), 981-999.

Hillis, A. E. (1998). Treatment of naming disorders: New issues regarding old therapies. *Journal of the International Neuropsychological Society, 4*, 648-660.

Hodges, J. R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33*(4), 441-459.

Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and knowing in dementia of the Alzheimer's type. *Brain and Language, 54*(2), 302-325.

Hodges, J. R., Salmon, D. P., & Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain, 114*, 1547-1558.

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia, 30*(4), 301-314.

Howard, D. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology, 2*(1), 49-80.

Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V., & Morton, J. (1985). Treatment of word retrieval deficits in aphasia. A comparison of two therapy methods. *Brain, 108*(4), 817-829.

Huff, F. J., Corkin, S., & Growdon, J. H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and Language, 28*(2), 235-249.

Huff, F. J., Mack, L., Mahlmann, J., & Greenberg, S. (1988). A comparison of lexical-semantic impairments in left hemisphere stroke and Alzheimer's disease. *Brain and Language, 34*(2), 262-278.

Jarvis, B. (2004). DirectRT. New York: Empirisoft.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16*(2), 243-275.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). Boston Naming Test. Philadelphia: Lea and Febiger.

Katzman, R., & Fox, P. J. (1999). The world-wide impact of dementia: Projections of prevalence and costs. In R. Mayeux & Y. Christen (Eds.), *Epidemiology of Alzheimer's disease: From gene to prevention* (pp. 1-17). Heidelberg: Springer-Verlag.

Kawas, C. H., & Katzman, R. (1999). Epidemiology of dementia and Alzheimer disease. In R. D. Terry, R. Katzman, K. L. Bick & S. S. Sangram (Eds.), *Alzheimer Disease* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.

Kearns, K. P. (1989). Generalization issues in the treatment of communication disorders. In L. V. McReynolds & J. E. Spradlin (Eds.), *Methodologies for studying generalization* (pp. 13-30). Toronto, Canada: B.C. Decker.

Kearns, K. P. (2000). Single-subject experimental designs in aphasia. In S. E. Nadeau, L. J. Gonzalez Rothi & B. Crosson (Eds.), *Aphasia and language* (pp. 421-441). New York: Guilford.

Kertesz, A. (1982). The Western Aphasia Battery. New York: Grune & Stratton.

Kiran, S., & Thompson, C. K. (2003a). Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain and Language, 85*(3), 441-450.

Kiran, S., & Thompson, C. K. (2003b). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research, 46*(3), 608-622.

Kirshner, H. S., Webb, W. G., & Kelly, M. P. (1984). The naming disorder of dementia. *Neuropsychologia, 22*(1), 23-30.

Koemeda-Lutz, M., Cohen, R., & Meier, E. (1987). Organization of and access to semantic memory in aphasia. *Brain and Language, 30*(2), 321-337.

Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin, 112*(3), 500-526.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Le Dorze, G., Boulay, N., Gaudreau, J., & Brassard, C. (1994). The contrasting effects of a semantic versus a formal-semantic technique for the facilitation of naming in a case of anomia. *Aphasiology, 8*(2), 127-141.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Science, 22*(1), 1-38.

Lipinska, B., & Backman, L. (1997). Encoding-retrieval interactions in mild Alzheimer's disease: The role of access to categorical information. *Brain and Cognition, 34*(2), 274-286.

Lowell, S., Beeson, P. M., & Holland, A. L. (1995). The efficacy of a semantic cueing procedure on naming performance of adults with aphasia. *American Journal of Speech-Language Pathology, 4*, 109-114.

Lukatela, K., Malloy, P., Jenkins, M., & Cohen, R. (1998). The naming deficit in early Alzheimer's and vascular dementia. *Neuropsychology, 12*(4), 565-572.

Margolin, D. I., Pate, D. S., & Friedrich, F. J. (1996). Lexical priming by pictures and words in normal aging and in dementia of the Alzheimer's type. *Brain and Language, 54*(2), 275-301.

Margolin, D. I., Pate, D. S., Friedrich, F. J., & Elia, E. (1990). Dysnomia in dementia and in stroke patients: Different underlying cognitive deficits. *Journal of Clinical and Experimental Neuropsychology, 12*(4), 597-612.

Marshall, J. C. (1988). Cognitive neuropsychology. Sensation and semantics. *Nature, 334*(6181), 378.

Marshall, R. C., Freed, D. B., & Karow, C. M. (2001). Learning of subordinate category names by aphasic subjects: A comparison of deep and surface-level training methods. *Aphasiology, 15*(6), 585-598.

Martin, A. (1992). Degraded knowledge representations in patients with Alzheimer's disease: Implications for models of semantic and repetition priming. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 220-232). New York: Guilford.

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language, 19*(1), 124-141.

Martin, N., Laine, M., & Harley, T. A. (2002). How can connectionist cognitive models of language inform models of language rehabilitation? In A. E. Hillis (Ed.), *The handbook of adult language disorders: Integrating cognitive neuropsychology, neurology, and rehabilitation* (pp. 375-396). New York: Psychology Press.

Mayer, J. F., Murray, L. L., & Karcher, L. (2004, June, 2004). *Treatment of anomia in severe aphasia.* Paper presented at the Clinical Aphasiology Conference, Park City, UT.

McCarthy, R., & Warrington, E. K. (1990). The dissolution of semantics. *Nature, 343*(6259), 599.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*(7), 939-944.

McNeil, M. R., Doyle, P. J., Spencer, K., Goda, A. J., Flores, D., & Small, S. (1998). Effects of training multiple form classes on acquisition, generalization and maintenance of word retrieval in a single subject. *Aphasiology, 12*(7/8), 575-585.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*(2), 99-130.

McReynolds, L. V. (1989). Generalization issues in the treatment of communication disorders. In L. V. McReynolds & J. E. Spradlin (Eds.), *Generalization strategies in the treatment of communication disorders* (pp. 1-12). Toronto, Canada: B.C. Decker.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*, 113-138.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234.

Mitchell, D. B., Hunt, R. R., & Schmitt, F. A. (1986). The generation effect and reality monitoring: Evidence from dementia and normal aging. *Journal of Gerontology, 41*(1), 79-84.

Montanes, P., Goldblum, M. C., & Boller, F. (1995). The naming impairment of living and nonliving items in Alzheimer's disease. *Journal of the International Neuropsychological Society, 1*(1), 39-48.

Multhaup, K. S., & Balota, D. A. (1997). Generation effects and source memory in healthy older adults and in adults with dementia of the Alzheimer type. *Neuropsychology, 11*(3), 382-391.

Murray, L. L., & Clark, H. M. (2005). *Neurogenic disorders of language: Theory driven clinical practice*. Clifton Park, NY: Thompson Delmar Learning.

Nadeau, S. E. (2000). Connectionist models and language. In S. E. Nadeau, L. J. Rothi & B. Crosson (Eds.), *Aphasia and language: Theory to practice* (pp. 299-347). New York: Guilford Press.

Nadeau, S. E., & Gonzalez Rothi, L. J. (2004). Rehabilitation of language disorders. In J. Ponsford (Ed.), *Cognitive and Behavioral Rehabilitation: From Neurobiology to Clinical Practice* (pp. 129-174). New York: The Guilford Press.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin, 106*(3), 377-394.

Nebes, R. D. (1992). Semantic memory dysfunction in Alzheimer's disease: Disruption of semantic knowledge or information processing limitation? In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 233-240). New York: Guilford Press.

Nebes, R. D., Boller, F., & Holland, A. (1986). Use of semantic context by patients with Alzheimer's disease. *Psychology and Aging, 1*(3), 261-269.

Nebes, R. D., & Brady, C. B. (1990). Preserved organization of semantic attributes in Alzheimer's disease. *Psychology and Aging, 5*(4), 574-579.

Nebes, R. D., Brady, C. B., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 11*(2), 219-230.

Nebes, R. D., Martin, D. C., & Horn, L. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of Abnormal Psychology, 93*(3), 321-330.

Neils, J., Brennan, M. M., Cole, M., Boller, F., & Gerdeman, B. (1988). The use of phonemic cueing with Alzheimer's disease patients. *Neuropsychologia, 26*(2), 351-354.

Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex, 21*(4), 595-606.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*(10/11), 935-979.

Ober, B. A. (1999). Semantic memory of Alzheimer's disease: Loss of knowledge or deficits in retrieval? Introduction from the symposium organizer. *Journal of the International Neuropsychological Society, 5*(7), 623-625.

Ober, B. A., Dronkers, N. F., Koss, E., Delis, D. C., & Friedland, R. P. (1986). Retrieval from semantic memory in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology, 8*(1), 75-92.

Ober, B. A., & Shenaut, G. K. (1995). Semantic priming in Alzheimer's disease: Meta-analysis and theoretical evaluation. In P. A. Allen & T. R. Bashore (Eds.), *Age differences in word and language processing* (pp. 247-271). Amsterdam, Netherlands: Elsevier.

Ober, B. A., & Shenaut, G. K. (1999). Well-organized conceptual domains in Alzheimer's disease. *Journal of the International Neuropsychological Society, 5*(7), 676-684.

Ober, B. A., Shenaut, G. K., & Reed, B. R. (1995). Assessment of associative relations in Alzheimer's disease: Evidence for preservation of semantic memory. *Aging and Cognition, 2*(4), 254-267.

Olympus. (2002). DM-10 [Digital Voice Recorder]: Olympus Corporation.

Ousset, P. J., Viallard, G., Puel, M., Celsis, P., Demonet, J. F., & Cardebat, D. (2002). Lexical therapy and episodic word learning in dementia of the Alzheimer type. *Brain and Language, 80*(1), 14-20.

Patterson, K., Purell, C., & Morton, J. (1983). The facilitation of word retrieval in aphasia. In C. Code & D. J. Muller (Eds.), *Aphasia therapy*. London: Edward Arnold.

Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L., & DeKosky, S. T. (2001). Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology, 56*(9), 1133-1142.

Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language, 52*(1), 25-82.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*(3), 353-363.

Pring, T., Hamilton, A., Harwood, A., & MacBride, L. (1993). Generalization of naming after picture/word matching: Only items appearing in therapy benefit. *Aphasiology, 7*(4), 383-394.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science, 12*(5), 410-430.

Raymer, A. M., Foundas, A. L., Maher, L. M., Greenwald, M. L., Morris, M., Rothi, L. J., et al. (1997). Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia. *Brain and Language, 58*(1), 137-156.

Raymer, A. M., & Gonzalez Rothi, L. J. (2001). The semantic system. In S. E. Nadeau & L. J. Gonzalez Rothi (Eds.), *Language and Aphasia*.

Raymer, A. M., Maher, L. M., Greenwald, M. L., Morris, M., Gonzalez-Rothi, L., & Heilman, K. M. (1990). The Florida Semantic Battery: Experimental Edition. Gainesville, Florida.

Raymer, A. M., Thompson, C. K., Jacobs, B., & le Grand, H. R. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology, 7*(1), 27-53.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning & Verbal Behavior, Vol. 12*(1), 1-20.

Rosch, E. (1973a). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

Rosch, E. (1973b). On the internal structure of perceptual and semantic categories. *Journal of Experimental Psychology: General, 104*(192-233).

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192-233.

Rosch, E. (1977). Human categorization. In N. Warren (Ed.), *Advances in cross-cultural psychology* (Vol. 1, pp. 1-72). London: Academic Press Ltd.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Llyod (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.

Rosenberg, R. N. (2005). Translational research on the way to effective therapy for Alzheimer disease. *Archives of General Psychiatry, 62*(11), 1186-1192.

Rothi, L. J., Fuller, R. H., Kendall, D. L., Nadeau, S., Leon, S., Wu, S., et al. (2005). *Naming treatment using an errorless learning technique: A possible adjuvant to donepezil for language dysfunction in patients with Alzheimer's disease.* Submitted for publication.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, Mass.: MIT Press.

Sailor, K., Antoine, M., Diaz, M., Kuslansky, G., & Kluger, A. (2004). The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology, 18*(2), 306-314.

Salmon, D. P., Butters, N., & Chan, A. S. (1999). The deterioration of semantic memory in Alzheimer's disease. *Canadian Journal of Experimental Psychology, 53*(1), 108-116.

Salmon, D. P., Heindel, W. C., & Lange, K. L. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society, 5*(7), 692-703.

Salmon, D. P., Shimamura, A. P., Butters, N., & Smith, S. (1988). Lexical and semantic priming deficits in patients with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 10*(4), 477-494.

Schmidt, A. M., & Lee, T. D. (1999). *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory & Language, 54*(2), 199-227.

Schwartz, M. F., Marin, O. S., & Saffran, E. M. (1979). Dissociations of language function in dementia: A case study. *Brain and Language, 7*(3), 277-306.

Schwartz, T. J., Kutas, M., Butters, N., Paulsen, J. S., & Salmon, D. P. (1996). Electrophysiological insights into the nature of the semantic deficit in Alzheimer's disease. *Neuropsychologia, 34*(8), 827-841.

Sclan, S. G., & Kanowski, S. (2001). Alzheimer disease: Stage-related interventions. *Lippincott's Case Management, 6*(2), 48-63.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental and Psychology: Human Learning and Memory, 4*, 592-604.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, Vol. 81*(3), 214-241.

Smith, S., Faust, M., Beeman, M., Kennedy, L., & Perry, D. (1995). A property level analysis of lexical semantic representation in Alzheimer's disease. *Brain and Language, 49*(3), 263-279.

Smith, S. R., Murdoch, B. E., & Chenery, H. J. (1989). Semantic abilities in dementia of the Alzheimer type. 1. Lexical semantics. *Brain and Language, 36*(2), 314-324.

Snodgrass, J. G., & McCullough, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(1), 147-154.

Stanczak, L., Waters, G., & Caplan, D. (2005, May 31-June 4, 2005). *Typicality-based differences in treatment of naming deficits.* Paper presented at the Clinical Aphasiology Conference, Sanibel Island, FL.

Stanczak, L., Waters, G., & Caplan, D. (2006). Typicality-based learning and generalisation in aphasia: Two case studies of anomia treatment. *Aphasiology, 20*(2-4), 374-383.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavioral Analysis, 10*, 349-367.

Thal, L. J. (1999). Clinical trials in Alzheimer disease. In R. D. Terry, R. Katzman, K. L. Bick & S. S. Sangram (Eds.), *Alzheimer Disease* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.

Thompson, C. K. (1989). Generalization issues in the treatment of communication disorders. In L. V. McReynolds & J. E. Spradlin (Eds.), *Generalization in the treatment of aphasia* (pp. 82-115). Toronto; Philadelphia: Decker.

Thompson, C. K., Ballard, K. J., & Shapiro, L. P. (1998). The role of syntactic complexity in training wh-movement structures in agrammatic aphasia: Optimal order for promoting generalization. *Journal of the International Neuropsychological Society, 4*(6), 661-674.

Thompson, C. K., Shapiro, L. P., Ballard, K. J., Jacobs, B. J., Schneider, S. S., & Tait, M. E. (1997). Training and generalized production of wh- and NP-movement structures in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research, 40*(2), 228-244.

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research, 46*(3), 591-607.

Thompson-Schill, S. L., Gabrieli, J. D., & Fleischman, D. A. (1999). Effects of structural similarity and name frequency on picture naming in Alzheimer's disease. *Journal of the International Neuropsychological Society, 5*(7), 659-667.

Tippett, L. J., Grossman, M., & Farah, M. J. (1996). The semantic memory impairment of Alzheimer's disease: Category-specific? *Cortex, 32*(1), 143-153.

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Oxford University Press.

Uyeda, K. M., & Mandler, F. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods and Instrumentation, 12*(6), 587-595.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*(4), 635-657.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*, 829-854.

Watson, M. E., Welsh-Bohmer, K. A., Hoffman, J. M., Lowe, V., & Rubin, D. C. (1999). The neural basis of naming impairments in Alzheimer's disease revealed through positron emission tomography. *Archives of Clinical Neuropsychology, 14*(4), 347-357.

Wiegel-Crump, C., & Koenigsknecht, R. A. (1973). Tapping the lexical store of the adult aphasic: Analysis of the improvement made in word retrieval skills. *Cortex, 9*(4), 411-418.

Wilkinson, G. S. (1993). Wide Range Achievement Test-Third Edition (WRAT-3): Jastak Associates, A Division of Wide Range, Inc.

Williams, B. W., Mack, W., & Henderson, V. W. (1989). Boston Naming Test in Alzheimer's disease. *Neuropsychologia, 27*(8), 1073-1079.

Williamson, D. J., Adair, J. C., Raymer, A. M., & Heilman, K. M. (1998). Object and action naming in Alzheimer's disease. *Cortex, 34*(4), 601-610.

Wilson, M. (1987). *MRC Psycholinguistic Database: Machine Usable Dictionary*. Retrieved May, 2006, 2006, from http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm

Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.

BIOGRAPHICAL SKETCH

Claudia A. Morelli received her Bachelor of Arts degree from Emmanuel College, Boston, MA. She also has a Master of Science degree from the University of Rhode Island, Kingston, Rhode Island, and is a certified speech-language pathologist. Claudia's interest in providing treatment for adults with neurogenic communication disorders motivated her desire to complete the doctoral program at the University of Florida. During that time, Claudia also had the opportunity to teach undergraduates for seven consecutive semesters, and from these experiences, developed a great appreciation for teaching. She has been trained in supportive and collegial academic and research environments driven by a core set of values focusing on the betterment of students, clients and their families, and ultimately the profession. In Fall, 2006, Dr. Morelli will be an assistant faculty member in the Department of Speech-Language Pathology/Audiology at Loyola College in Maryland.