Graduate Theses and Dissertations   Graduate School

January 2013

Does the Use of Personally Relevant Stimuli in Semantic Complexity Training Facilitate Improved Functional Communication Performance Compared to Non-Personally Relevant Stimulus Items among Adults with Chronic Aphasia?

Stephanie Karidas
University of South Florida, skaridas@mail.usf.edu

Does the Use of Personally Relevant Stimuli in Semantic Complexity Training

Facilitate Improved Functional Communication Performance Compared to

Non-Personally Relevant Stimulus Items among Adults with Chronic Aphasia?

by

Stephanie Karidas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Communication Sciences and Disorders
College of Behavioral and Community Sciences
University of South Florida

Co-Major Professor: Jacqueline Hinckley, Ph.D.
Co-Major Professor: Stefan Frisch, Ph.D.
Yael Arbel, Ph.D.
Cathy McEvoy, Ph.D.

Date of Approval:
July 10, 2013

Dedication

This dissertation is lovingly dedicated to my family, especially to my loving parents, Panagiotis and Gundula Karidas. Thank you for all your love, support and sacrifice throughout my life.

Acknowledgment

I am most grateful to the members of my committee, Dr. Jacqueline Hinckley, Dr. Stefan Frisch, Dr. Yael Arbel, and Dr. Cathy McEvoy for their time, encouragement, and expertise throughout this project. I would also like to thank the chair of our department, Dr. Theresa Chisolm, who has been so very supportive of my academic pursuits. I want to further thank Peggy Ott and Vivian Maldonado for always being there for me and answering all my questions.

There are many people in my life who have made my success both possible and rewarding. My parents, my sisters, my brother, my uncle, and my in-laws. Thank you now and always. Areti Marangaki, Stella and Panos Valavanis – you have become like my family and have never stopped believing in me and my dreams. Thank you for all your love. My former fellow students and now colleagues Natalie Douglas and Rachel Goff – thank you for all the laughs and team spirit. We have come a long way together and I am more than grateful to have embraced their friendship and support throughout this program. To all my dear friends - I will always remember the support and encouragement I received from you.

Table of Contents

List of Tables

List of Figures

Abstract

This study investigated the influence of semantic complexity treatment in individuals with fluent aphasia on discourse performance. Semantic treatment is an effective way to improve semantically based word retrieval problems in aphasia. Treatment focused on the semantic application of the Complexity Account of Treatment Efficacy (CATE) (Thompson, Shapiro, Kiran, & Sobecks, 2003) promotes training of complex items resulting in generalization to less complex, untrained items. In addition, research has shown that the personal relevance of treatment material can increase treatment efficacy. This study investigated the effect of semantic treatment of atypical personally relevant items among individuals with aphasia on discourse performance.

Two treatment phases were applied to examine the influence of personally relevant and non-relevant treatment material on discourse performance. In addition, generalization from trained atypical items to untrained typical items was investigated. Methods and procedures were partially replicated from Kiran, Sandberg, & Sebastian (2011) examining semantic complexity within goal-derived (ad hoc) categories. Three participants with fluent aphasia were trained on three semantic tasks including category sorting, semantic feature generation/selection, and Yes/No feature questions. A generative naming task was used for probe data collection every second session. Stimuli consisted of atypical items only.

The hypothesis that semantic complexity training of personally relevant items from ad hoc categories will produce greater generalization to associated, untrained items than training of non-relevant items and consequently increase discourse performance was not supported. The findings revealed a failure to replicate the magnitude and type of improvements previously reported for the typicality effect in generative naming. Clinical significance was found for personally relevant and non-relevant discourse performance. However, no consistent pattern was found within and across participants. In addition, effect size for generalization from trained atypical to untrained typical items was not significant.

Limitations of this study lead to future directions to further specify participation selection, such as cognitive abilities, procedural changes, and the inclusion of discourse performance as an outcome measure. Overall, the results of this study provide weak support for replicating semantic treatment of atypical exemplars in ad-hoc categories and hence demonstrate the critical role of replication across labs to identify key issues in the candidacy, procedures, and outcome measurement of any developing treatment.

Chapter One:

Introduction and Review of the Literature

Word retrieval deficits are the "hallmark impairment of aphasia" (Fridriksson, Holland, Beeson, & Marrow, 2005; p.99) and create a substantial obstacle in inter- & intra- personal communication (McNeil, Doyle, Spencer, Jackson Goda, Flores, & Small, 1997). The implementation of efficient treatment methods addressing word retrieval deficits is therefore critical.

Anomia, the impaired access to one's vocabulary, also known as 'dysnomia' (Goodglass, 1993; Coelho, McHugh, & Boyle, 2000), is a common characteristic of all aphasia types and is most prevalent in the fluent types, i.e. Wernicke's, anomic and transcortical sensory aphasia (Kiran & Bassetto, 2008; Raymer, 2005; Kertesz, 2006). Clinical forms of anomia present with some type of paraphasia which is the "unintended error of word or sounds choice" (Goodglass, 1993, p. 78). Paraphasias can be subdivided into verbal (semantic), phonemic (literal), phonosemantic blends, and neologistic symptoms. Paraphasias may be rooted in an impaired storage, retrieval, encoding/decoding, and/or selection- based process in the language system (Davis, 2000; Coelho et al., 2000).

Consequently, a word retrieval deficit can be a deficit at a semantic or phonological/phonemic level. The distinction between these different language structures is extremely important, because the correct diagnosis will provide specific process-based treatment goals. Greenwald, Raymer, Richardson, & Rothi (1995)

1

differentiate between an input modality-specific naming impairment where individuals have difficulties retrieving semantic features from visually presented items, and an output mode impairment such as oral naming difficulties (phonological output). These two naming difficulties are approached through different naming tasks, i.e. visually presented items will facilitate semantic retrieval of lexical items, whereas cueing tasks are able to support the retrieval of semantic features, as well as the phonemic portion of words.

In general, the underlying process of speech production consists of three broad areas: 1) C*onceptualization*, 2) *Formulation*, and 3) *Encoding* (Levelt, 1989). *Conceptualization* involves the process of "determining what to say", which is then *formulated* into a linguistic form and finally executed as a motor function. The first and last areas are not directly involved in word retrieval and will therefore not be further discussed for the purposes of this paper. Levelt (1989) described *Formulation* as a two component level, consisting of lexicalization (selection of a word) and syntactic planning (forming a sentence with selected words) (Harley, 2001). A two-step interactive psycholinguistic model of lexical retrieval (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997) describes the basis of word retrieval deficits in more detail. In this model, speech errors of two different types, lexical (verbal paraphasias) and sublexical (phonemic paraphasias and neologism) are simulated. According to Dell et al. (1997), individuals with Wernicke's aphasia characteristically produce lexical and sublexical errors, whereas individuals with anomic aphasia are more likely to produce semantic errors resulting in circumlocutions and word searching. Conduction aphasia

exposes more phonemic paraphasias which are characterized by self-awareness as these individuals self-interrupt and correct repeatedly (Dell et al., 1997). Nonetheless, these selective impairments in word retrieval processes primarily occur during spontaneous speech. Picture naming evokes both lexical and sublexical errors in all fluent aphasia types. The difference between free conversation (spontaneous speech) and picture naming is the result of the different demands of the tasks involved. During picture naming, the individual with aphasia is constrained to retrieve a specific word, whereas free conversation allows the use of rather high frequency, i.e. more familiar words, that are easier to retrieve, and allows the use of alternative words in case of a retrieval failure. Unconstrained access to high frequency words and sub-word constituents results in jargon filled speech and neologism (Goodglass, 1993).

Given these different sources of lexical retrieval difficulty, naming disorders should not be treated homogenously, because individuals display different retrieval difficulties and will benefit the most from treatment tasks addressing the individual system breakdown. Model-based therapeutic techniques provide a useful tool to engage and reactivate specific retrieval processes (Nettleton & Lesser, 1991; Drew & Thompson, 1999). Improvements in word retrieval can also improve the ability of a person with aphasia to participate in conversation, thus potentially reducing social isolation and depression (Cruice, Worrall, & Hickson, 2006; Code, 2003). Language therapy can increase the patient's social network through the relearned ability to communicate and can help prevent a depressive state of mind through communicative and social isolation (Hinckley & Packard, 2001; Cruice et al., 2003).

Language treatments specific to word retrieval have focused on simple training regimens, such as emphasizing more basic and rudimentary structures of language, rather than complex items or structures of language. Word retrieval treatment outcomes through a more sophisticated approach have recently resulted in more success with greater generalizable language performance (Holland, Fromm, DeRyter, & Stein, 1996). This approach is known as the Complexity Account of Treatment Efficacy (CATE) (Thompson, Shapiro, Kiran, & Sobeck, 2003) which offers a new and more effective approach to semantic treatment for anomia. The complexity account suggests that treating atypical (more complex) items in a category leads to greater gains in lexical access across the category. For this study, we will focus on semantic treatment gains achieved by training atypical stimulus items with an additional variable, the personal relevance of the semantic category of the stimuli.

Category Typicality and Semantic Complexity

Various models of semantic network processing have been applied to the study of word-retrieval deficits in aphasia. A very fundamental model for this study's purpose is the Prototype/ Family Resemblance Model (Rosch & Mervis, 1975). The creators of this model state that category members share a set of family resemblances and that the knowledge of semantic features of things is stored in category prototypes (Rogers, in press). The model predicts that the semantic attributes of specific prototypes are being retrieved and compared to words that are entering the system, i.e. lexical items are being matched to categorical prototypes based on their feature similarity. Therefore, a semantic

hierarchy can be derived from this model in which each item is found within a graded structure of feature similarity, also defined as *typicality*. Typicality plays a significant role for the complexity approach to treatment of anomia, because typicality is based on a hierarchical organization of semantic features. This hierarchy is also known as a graded structure of categories which represents two types of items within the category: typical and atypical items. These types of typicality-rankings are based on a rather simple concept, namely how many features (attributes) of a category prototype are in fact represented as defining attributes of a specific item. The category *bird* for example has *Robin* and *Sparrow* defined as a rather typical item of this category, because it represents more idealized features, i.e. prototypical items, such as *is small, hops, flies* which they share with quite a few other items of the same category. Shared attributes make an item in the semantic network less 'different' and these items have therefore fewer distinctive features that make them 'more special'. Atypical examples on the other hand are more 'unique' in a sense that they carry fewer prototypical features, such as *Ostrich*. In exchange, these atypical items carry more distinctive features that are shared by fewer examples of the category, e.g. *has long neck* and *has long legs*.

Generally, typical items carry core features that have a greater frequency of appearance within the category itself. In contrast to atypical items, typical items carry more prototypical features and share features with other members of the category. Atypical items on the other hand carry fewer prototypical features

and are defined by more distinctive features, because they are defined by features that are less prevalent and diffuse within the category (Kiran, 2007).

An important contribution to our understanding of the cognitive representation of semantic categories and processing of typical vs. atypical items in the semantic network has been made by Rosch (1975). The author proposes that the human semantic network reacts faster in processing and retrieving items that are typical because typical items carry less complex semantic attributes. Conversely, atypical items have a slower reaction time, primarily because their features are more complex and carry more weight, i.e. a greater number of features need to be processed. Consequently, matching items to specific categories results in a faster retrieval time for typical items compared to atypical items. For example, the decision time for a 'chair' as a category member of 'furniture' is faster than deciding whether a 'rug' belongs to the same category (Grober, Perecman, Kellar, & Brown, 1980).

Acknowledging an existing typicality effect in the semantic network is important for the treatment of brain damaged individuals. A great contribution here has been made by Plaut (1996) who conducted a computerized experiment on the typicality effect for nouns. In this study, a computer network was trained to recognize a set of artificial typical and atypical words. Plaut provided the words with 'semantic features' in the form of binary codes, with typical ones receiving fewer distinctive features than atypical ones to represent a graded structure. The network was then purposely damaged to simulate brain damage, and was subsequently re-trained again. As a result, Plaut found a generalization effect

when atypical words were (re-)trained in the network, but not when typical items were (re-)trained. The training of atypical items forced the system to encounter more features due to the graded structure. Since typical items share features with atypical items, a generalization effect from atypical to typical items occurs. Hence, an effective generalization effect through complexity was provided.

Consistent with Plaut's artificial network study, many behavioral studies have achieved similar results with individuals with aphasia, such as Grober, et al. (1980), Grossman (1981), Kiran & Thompson (2003b), and Kiran, Sandberg, & Abbott (2010). An overview of the most recent studies will be provided later on.

In order to address the complexity effect on semantic retrieval training in aphasia, it is necessary to define to what extent the semantic memory system is impaired in this specific population. Individuals with anterior aphasia (non-fluent) show fewer deficits in accurately naming atypical category members and unrelated non-members of a category. Conversely, individuals with posterior (fluent) aphasia show greater deficits in the same application of naming tasks, which leads us to the conclusion that the semantic memory system in fluent aphasia is more greatly damaged compared to non-fluent individuals with aphasia.

Typicality of lexical items has been investigated by many researchers and "typical examples receive preferential processing relative to other examples in the category and this phenomenon has been labeled the typicality effect" (Kiran & Thompson, 2003a, p. 441). We will therefore refer to the difference of typicality

ratings as the 'typicality effect' (i.e. typical items are being accessed faster compared to atypical items in the semantic network).

Semantic Complexity in Adults with Aphasia

As already addressed in the beginning of this paper, a very early and, for the purpose of the study's subject, very fundamental study on lexical knowledge in individuals with anterior and posterior aphasia was conducted by Grober et al. (1980). The experiment focused on latency and accuracy in categorizing typical versus atypical members of a category. Participants were asked to decide whether a presented item belonged to a specific category or not. Results revealed that reaction times for both groups, with anterior and posterior lesion, were faster for judging typical items of a category compared to atypical items. However, the data showed that participants with anterior aphasia were more accurate in deciding whether an item was atypical or merely a semantically related nonmember of the presented category (Grober et al., 1980). These results coincide with results provided by Grossman (1980, 1981) who found that non-fluent aphasia patients are more precise in recalling highly representative semantic features of a category whereas individuals with fluent aphasia "not only were less sensitive to central instances of superordinates, but they often went well beyond what normals often considered a reasonable extension of the word" (Grossman, 1981, p. 316). As mentioned earlier, these semantic feature characteristics were confirmed in a subsequent study (Grossman, 1980) and it was stated that there is greater ease of access of typical items in people with non-fluent aphasia. For example, these participants were able to recall 'robin'

and 'sparrow' as members of the category 'birds', but they were unable to define the membership for more atypical items of the same category, such as 'goose' or 'swan' (Grossman, 1980). More evidence was provided by Plaut's computerized experiment which provided novel evidence for a greater generalization effect when atypical items were trained compared to typical items. Results of that study revealed that individuals with fluent aphasia (Wernicke's) did not show any difference in reaction time between typical and atypical items during a categorization verification task (Plaut, 1996).

As a partial replication of the above described study by Grober et al. (1980), Kiran & Thompson (2003a) investigated the typicality effect of category exemplars on category verification in individuals with Broca's and Wernicke's aphasia, including healthy young and elderly participants as a control group. Similar to Grober et al.'s (1980) study, stimuli consisted of three animate categories: birds, vegetables, and fish. Each category contained 15 typical and 15 atypical examples, as well as 30 nonmembers. Participants were asked to press a Yes/No button after *"Is (x) a member of (y)?"* with 'x' being the target word, and 'y' being the superordinate category label (e.g. "Is 'robin' a bird?" as a typical member of the category *bird,* and "Is 'penguin' a bird?" as an atypical example).

The authors calculated the percent advantage for typical examples compared to atypical examples across all groups. Data revealed similar results as found in Grober et al.'s (1980) study, confirming that reaction time does differ depending on the type of aphasia. The data revealed that individuals with

Wernicke's aphasia had a slower reaction time and greater error proportion compared to all other groups, including individuals with Broca's aphasia. Individuals with Broca's aphasia performed slower and less accurately than both control groups, young and elderly participants. The typicality effect was evident among the young, elderly, and individuals with Broca's aphasia, but not for the Wernicke's group. A general result was, however, that all four groups were slower and less accurate in classifying atypical items into a category compared to typical items.

In a study published shortly after this verification experiment, Kiran & Thompson (2003b) investigated semantic complexity by controlling for typicality of category exemplars when examining the generalization effect to untrained items (Kiran & Thompson, 2003b). Four individuals with fluent aphasia underwent naming treatments for typical and atypical items within two categories. The following categories were selected for treatment, *bird* and *fish*, as well as additional distracter categories (fruits, musical instruments, and animals). Semantic features were divided into four categories: physical, functional, characteristic, and contextual. However, distracter features from other categories (sports, transportation, animals, insects, flowers, weapons) were also included in the testing.

Overall, a complexity effect was found. Specifically, generalization for untrained items was only observed when atypical items were trained, i.e. improved naming of typical and intermediate items of the same category was

observed following atypical treatment. No generalization to naming of atypical and intermediate items was observed when training focused on typical items.

Further investigation of the generalization of the complexity effect by the same authors has led to the conclusion that naming therapy focusing on more complex items results in greater language improvement and improves the everyday quality of life for the individuals. Kiran (2008) provided evidence of generalization for training of atypical, inanimate categories on five individuals with aphasia (two fluent and three non-fluent individuals with aphasia). Treatment consisted of picture naming, sorting pictures by category, identification of semantic features, and yes/no answers about semantic features. Two inanimate categories, *clothing* and *furniture*, were used as training and testing stimuli. Each category contained the same number of trained typical and atypical items. Interestingly, training of atypical clothing and furniture items resulted in generalization to untrained typical items for two participants. Training of typical furniture items did not lead to generalization to untrained atypical items (one participant). Two out of five participants dropped out of the study due to their own wish to terminate the study. These participants completed one treatment phase each.

A slightly different approach, conceptual generalization of trained abstract versus concrete words, was investigated by Kiran, Sandberg, & Abbott (2010). The authors propose that abstract concepts are defined by concrete and other concepts, e.g. an abstract word such as 'prayer' will activate the retrieval of concrete words related to a category like 'church' as the overall concept, such as

'candle' and 'bible' (Kiran et al., 2010). Concrete words on the other hand (e.g. 'candle') can further activate more concrete words, but will not activate abstract words. Hence, abstract words are more complex in their semantic representation compared to concrete words. To investigate the complexity effect of these conceptual representations, generalization was investigated through simple judgments regarding the frequency and imageability of words. Four individuals with fluent aphasia (anomic) received training with similar tasks used in the earlier study described above (word recall, category sorting, feature selection, and yes/no feature questions). Four categories were introduced: church, hospital, museum, and courthouse. Each category consisted of concrete and abstract semantic features, e.g. a concrete word for *'church'* was *'candle'*, whereas *'holy'* was considered an abstract word. Similar results to previous studies were found. Training of abstract words led to generalization of untrained concrete words (within the same category), and training of concrete words did not show any generalization to untrained abstract items.

Further evidence of the semantic complexity effect was provided in a recent study by Kiran, Sandberg, & Sebastian, (2011) in which six individuals with fluent aphasia received word retrieval training of typical and atypical ad hoc category items. Items were trained for *'things at a garage sale'* and *'things to take camping'*. Treatment focused on training specific semantic features of specific items, which would then generalize to its adjacent lexical neighbors. According to the authors, ad-hoc categories are less common, implying they represent looser semantic memory, but do possess graded structures that determine member

typicality (Kiran et al., 2011). It was then hypothesized that trained atypical items would result in generalization to untrained typical items within the same category, and with less facilitation of categorical generation, trained typical items would show no generalization to atypical items. All participants revealed generalization to untrained typical items when trained on atypical items. No generalization was found after training of typical items.

Results have also shown that the training of atypical items leads to a spreading activation of items outside of a category. For example, typical items for a 'things needed at a camp' ad hoc category were noted as core attributes of this specific category, such as mosquito spray and sleeping bags. Atypical items, contrastingly, activated a wider range of semantic features. Items listed for this category were things to do at a camp, but also included items for personal hygiene. Accordingly, it was claimed that atypical items strengthen the connection across semantic attributes, resulting in a greater variety of retrieved words.

The current study is based on application of the CATE (Thompson et al., 2003) to individuals with fluent aphasia. This study investigates the training of atypical category members to improve naming performance.

Complexity Account of Treatment Efficacy (CATE)

The CATE has been described by Thompson et al. (2003) as an 'overarching principle' of a rather new, but effective treatment approach (Thompson et al., 2003). The authors define CATE as follows: "Training complex structures results in generalization to less complex structures when untreated

structures encompass processes relevant to (i.e., are in a subset relation to) treated ones" (Thompson et al., 2003, p. 602). Although this approach is relatively novel, several studies have led to the conclusion that decreased language function can be rehabilitated with greater success if complex rather than simple sounds (phonological complexity), words (semantic complexity), and syntactical structures (syntactic complexity) are treated. The complexity approach represents a new treatment paradigm, because its conceptual approach is in strong contrast to the traditional treatment approach, which promotes training simple structures (single sounds, easy/ typical words, less complex language structures) before more complex structures.

Research evidence has been imprecise on whether the improved naming abilities from CATE training will generalize over to an enhanced overall communicative performance used on a daily basis, such as improved conversational performance through the use of trained, atypical stimulus items. Stimuli used in previous studies, such as birds, furniture, clothes, and vegetables, have used concrete categorical boundaries and were found to show a generalization effect from trained to untrained words within a category. The use of the semantic application of the CATE with ad-hoc stimulus categories has also lead to a generalization effect within the semantic network even when categorical boundaries are less predefined compared to concrete stimulus items (birds, vegetables, etc.). Nevertheless, a functional application of increased semantic performance in regards to conversational patterns has not been investigated.

In summary, clear evidence has been provided for the positive impact of the complexity approach on an individual's word retrieval within and across categories, specifically in individuals with fluent aphasia. Overall, 17 out of 24 individuals with aphasia who participated in previous studies investigating the semantic application of the CATE revealed generalization from trained atypical to untrained typical items. Nonetheless, a major goal of treatment interventions for individuals with aphasia is to extend communication to a level where the individual can re-integrate him-/herself into the social network of a community. Word retrieval performance within discourse performance has not been investigated by previous studies, i.e. the use of complex stimuli has not yet been proven to increase the informational content and efficiency of a message as it can be measured by discourse performance. A legitimate reason behind not incorporating a conversational analysis might be the nature of previously used stimuli, e.g. concrete categories such as *clothing, birds*, and *furniture*. Very recent research has also led to the conclusion that not only complex stimulus items lead to a greater generalization within a category, but that training of complex items of categories with rather loose boundaries might create an even greater generalization effect (Kiran et al., 2011). Although categories with loose boundaries including '*Things to take camping*' and '*Things to sell at a garage sale*' were used by Kiran et al.(2011), outcome measures were not tailored to look at single word retrieval and word retrieval within connected speech (discourse performance).

This study will use stimulus categories with loose boundaries to investigate the effect of semantic complexity training on discourse performance. The following section therefore depicts a contrast between concrete and loose boundaries to provide the underlying argument for the use of stimulus categories with loose boundaries.

Concrete versus Loose Category Boundaries

The semantic representation of words is based on semantic features (Moss & Gaskell, 1999). Semantic features define a word that we hear or read and through these features a meaning is assigned to a word. Our semantic network consists of many of these features and corresponding words which are interrelated and may be part of overlapping semantic categories. A categorical arrangement of words in our semantic network enables us to process language and its meaning. For the purpose of this study we will focus on spoken language only.

Semantic processing is a crucial underlying neurological feature that enables us to not only retrieve words during a conversation, but also to processes and recognize auditory input. Once sensory information is transmitted to the cortex, an interpretation of this incoming information takes place. Language units are processed and forwarded to specific branches of the semantic network. A greater semantic activation is achieved if a word is recognized as a meaningful unit to the individual. Less activation is created when the meaning of a specific word is unknown (Moss & Gaskell, 1999). As described in the next section, a greater familiarity (personal relevance) of stimulus items is

related to a concept based on "relate the to-be-remembered information to one's self" (Viskontas, Quian Quiroga, & Fried, 2009, p. 21329). This concept reflects a self-reference effect in which the individual is more likely to associate items with higher familiarity, i.e. personal relevance, because they have undergone a more semantically elaborative encoding process (Viskontas et al., 2009).

Semantic processing also takes place during word retrieval when forming the idea of an utterance. During word retrieval processing, specific neuronal pathways are activated, leading to a specific word. In individuals with anomia (word retrieval difficulties), the pathway to a concept and its semantic properties is disabled, blocking the retrieval of the intended word (Nickels, 2002).

Many semantic categories have concrete boundaries, and there must be specific features to have membership in the category. A word that fits in the category, for example *furniture*, could not be stored in the *food* category, because it is not edible, but is used for sitting or storing.

An advantage of using words with concrete boundaries in a research study is that training focuses on strengthening activation within the same semantic dimension. A considerable disadvantage of this approach is that retrieval of activated sections within a category is restricted to only a limited number of items since these categories consist of pre-defined, concrete boundaries. For example, 'birds' and 'fish' are both items from the category 'animals'. Both items have concrete boundaries in which their semantic features are 'pre-defined', i.e. members of this category have to have specific features,

such as '*can swim/fly*', '*has a beak/fin*', etc. Activation therefore only spreads within this category, but not beyond it (across categories).

Although items from a category with concrete boundaries cannot become a member of another category with concrete boundaries, these items can become a member of a category with loose boundaries. For example, the category "*Things to take camping*' can consist of items taken from a variety of concrete categories, such as *chair, food*, or *kitchen utensils*. However, these items cannot change membership from one concrete category to the other. That is, a *chair* cannot be a member of the category '*food',* because it would not fall under the required semantic feature distinction of food items. Their membership is therefore only interchangeable when used within the context of categories with loose boundaries, because there is no definite feature boundary that establishes membership. Categories with loose boundaries do not have rigid semantic features which constitute the category membership. Ad-hoc categories have a fairly loose combined thread of common features instead (Kiran et al., 2011), such as the category '*Things to take camping'*. As mentioned above, the category '*Things to take camping'* consists of a variety of features taken together from other categories, such as '*personal hygiene'*, '*food for easy quick meals'*, etc. The membership of words within these categories is determined by an individual depending on how they define each category (in this case dependent on what they would take camping). Words from categories with rather loose boundaries result in wider semantic activation, i.e. a greater variety of words is being retrieved to form a conglomeration of items.

Concrete categories illustrate a fairly simple and structured organization of the semantic network. Barsalou (1983) has approached the categorization theory and took the taxonomical structure of representational concepts to a different level by specifying 'ad hoc' categories. Ad hoc categories are defined as "highly specialized and unusual sets of items constructed spontaneously for use in particular contexts" (Hough, 1989, p.554). These categories are not comparable to common categories, such as furniture and fruits. Instead, these categories are created and used for a specific purpose, such as to decide which things to pack for a camping trip or what to sell at a garage sale. Barsalou (1983, 1991) states that ad-hoc categories, such as 'things to take camping/sell at a garage sale', are categories that are novel to the individual's semantic memory, i.e. ad-hoc categories represent new categories for specific situations. Once an ad-hoc category becomes well-established in the semantic memory, it forms a category on its own and turns into a thematic, permanent category, also termed as a 'goal-derived' category. Goal-derived categories subsequently underlie the same graded semantic structure as concrete categories and hence also reveal a typicality effect.

To expand the idea of complexity to other, less restricted domains of stimulus items, Kiran et al. (2011) have approached ad-hoc categories, such as '*Things to sell at a garage sale*' and '*Things to take camping*', as mentioned earlier. Kiran and colleagues (2011) hypothesized in their study that goal-derived categories are also represented in terms of typicality and potentially have the same effect on complexity training as previously trained items with concrete

boundaries. Due to the fairly loose boundaries of semantic categories, it was expected that training of atypical items would generalize to items with similar semantic features. An important postulation here was that the authors assumed that items stored in one category potentially generalize over to the retrieval of semantic items from another category (e.g. things to take camping) if relevant atypical items were used for training (e.g. 'Items used for personal hygiene', compared to 'Items used for setting up a camp'). Results from this study were in concordance with previous results, namely that training of atypical items generalized to untrained, typical items. Specifically, Kiran and colleagues (2011) found that lexical access was increased within a generation naming task of categories with loose boundaries only, i.e. goal derived ad-hoc categories. Training of semantic word retrieval with atypical items generalized to untrained typical items. Training of typical items on the other hand resulted in retrieval of only core features of the category (Kiran et al, 2011). According to the authors, a theoretical application of these findings is that training categories with loose boundaries increases the communicative ability of individuals with impaired semantic representations.

Additionally, responses generated for each category were also subject to a qualitative and quantitative analysis to examine the "nature of retrieval" (Kiran et al., 2011, p.1110) and if complexity treatment would create a greater number of overall responses. The authors state that trained atypical examples elicited more responses of a targeted category compared to responses from trained typical examples. Although this study provides additional evidence that

complexity within loose categories increases the outcome effect of semantic retrieval processing, it does not clearly state whether participants received any benefit beyond generative naming. The efficiency of complexity training in reference to the individual's discourse performance ability in daily living situations is therefore yet to be determined.

Personal Relevance

A major decisive factor of treatment studies in aphasia is the challenge to make the outcome most efficient for the individual's language use. The use of complex stimulus materials has been shown to generalize from trained to untrained items. Categories with loose boundaries also result in greater retrieval across multiple concrete categories, and hence shown an increase of an individual's word retrieval performance.

Another facilitating factor to promote a more effective outcome is personal relevance to pictures, people, items, and/or places. Personal relevance reflects a great familiarity to an individual. A concept based on "relate the to-be-remembered information to one's self" (Viskontas et al., 2009, p. 21329) reflects a self-reference effect in which the individual is more likely to associate items with higher familiarity, i.e. personal relevance, because they have undergone a more semantically elaborative encoding process (Viskontas et al., 2009).

The neuronal and behavioral impact on processing speed of personally relevant stimuli has been studied in healthy adults, as well as in brain damaged individuals. The use of personally relevant stimulus items leads to greater activation of brain regions. Bear, Connors, & Paradiso (2001) explain this

phenomenon of greater activated brain regions through expertise or greater interest in a specific object or topic as 'highly developed specialized processing of visual features needed to classify particular examples' (p.751), i.e. activation is based on sensory input. The authors also state that these highly activated brain regions encode memory networks for a specific item category. A study by Gauthier, Skudlarski, Gore, & Anderson (2000) has confirmed that an acquired expertise for specific objects elicits different brain responses if individuals are confronted with picture stimuli in the area of their expertise. Consequently, greater activation of specific brain regions correlated with greater expertise of participants with the stimuli. Moreover, Gauthier et al. (2000) states that subordinate levels of categorization were accessed automatically by experts, but not by novices to a specific category. For example, access to 'barn owl', two levels below the basic level category 'bird' was found for experts, while both experts and novices accessed the intermediate category 'owl'. Scott, Tanaka, Sheinberg, & Curran (2008) supported this view of neural subordinate activation based on higher knowledge about the stimulus category. Previous studies by Tanaka, Curran, & Sheinberg (2005) and Scott, Tanaka, Sheinberg, & Curran (2006) have shown that training of subordinate semantic levels leads to greater generalization from trained to untrained stimuli, as well as a greater discrimination performance for trained items (Scott et al., 2008). Consequently, greater familiarity and/or personal relevance of the stimuli to the individual contribute to greater automatic neural activation (Gauthier et al., 2000).

The interaction between expertise and long-term memory has been further investigated by Groussard et al. (2010), who examined the effects of musical expertise on functional and structural plasticity in the hippocampus. Their results revealed that familiar musical tasks, e.g. semantic memory for music, activate neuronal areas that are of higher gray density, i.e. result in greater neural activation (Groussard et al., 2010). Although this experiment was conducted with healthy, normal adults, results support the fact that specific brain regions (hippocampus, parts of the medial temporal lobe (MTL)) contribute to the recollection and processing of familiar information (Milner, 1962; Skinner & Fernandez, 2007; Eichenbaum, Yonelinas, & Ranganath, 2007; Groussard et al., 2010). Viskontas, et al. (2009) provided additional evidence that neurons in the healthy brain are highly selective when personally relevant photograph stimuli (e.g. family members and personal acquaintances) are presented. Neuronal areas in the MTL and hippocampus specifically affected by the enhanced triggered speed response and memory retrieval are activated during semantic memory associations. Viskontas et al. (2009) states that the MTL reveals a higher neural activation when personally relevant photographic stimuli were presented compared to a lower activation when photographs consisted of entirely unfamiliar people. Gauthier et al. (2000) provided evidence that the recognition of familiar faces activates neural regions in the Fusiform Face Area (FFA) which is located in the temporal lobe. The authors also state that expertise and categorization level of homogenous categories (*birds* and *cars* in this case) also cause greater activation in the same FFA area. This brain region therefore

serves retrieval for familiar faces, people, and other categorical processes as long as they are of great familiarity to the individual (Kanwisher, McDermott, & Chun, 1997).

McKelvey (2007) investigated the impact of personal relevance on picture stimuli for language treatment with people with aphasia and states that relevance does not only reflect a 'relationship to the matter of hand' (p.31), but that personally relevant material also has a strong social applicability to the individual. Consequently, the author refers to personally relevant stimuli as especially useful items for treatment of aphasia, because they imply a personal connection to the client, are of high interest and potentially support a longer attention span for therapy activities. Personally relevant stimulus items trigger the contextual retrieval of information stored in a memory system and enable the individual to activate a holistic retrieval process, i.e. access contextual information within a category (e.g. Wilkinson & Jargoo, 2004; Dietz, McKelvey, Beukelman, Weissling, & Hux, 2006; McKelvey & Dietz, 2007).

A more recent study by McKelvey, Hux, Dietz, & Beukelman (2010) has additionally demonstrated that individuals with aphasia are more accurate on word-picture naming with the application of personally relevant stimulus items, as opposed to non-personally relevant items. The study investigated whether participants preferred personally relevant pictures as opposed to non-personally relevant pictures. Accuracy on word-picture matching using personally relevant versus personally contextual and decontextualized pictures was also investigated. Results supported a strong preference for personally relevant items

and a higher accuracy on personally relevant pictures during the matching task. The authors concluding statement is that personally relevant stimulus items trigger a higher performance in language processing in that the familiarity with the stimulus items offers a greater ease of recognizing and retrieving information from the semantic network. Moreover, solely the 'naturalness and intrinsic motivation' associated with personally relevant items triggers better language performance as compared to less personal stimulus items (McKelvey et al., 2010).

As a result, it is hypothesized that the use of personally relevant stimulus items in semantic complexity training will facilitate discourse performance in comparison to non-personally relevant items. Hence, a more functional use of the CATE with personally relevant categories is hypothesized to increase the individual's daily communicative ability and consecutively improve his/her life quality.

This study is a partial replication and extension of Kiran et al's (2011) work, investigating the typicality effect in ad hoc category training and production, and comparing the potential effect of personally relevant stimuli with those that are less personally relevant.

Research Question

Does the use of personally relevant stimuli in semantic complexity training improve discourse performance compared to training on non-personally relevant items among adults with chronic fluent aphasia?

Hypothesis

Due to the strong evidence from previous studies on the positive treatment effect of the CATE on word retrieval, it is hypothesized that semantic complexity training of personally relevant items from ad hoc categories will produce greater generalization to associated, untrained items than training of non-relevant items. The improved access to personally relevant semantic categories will consequently increase discourse performance.

Chapter Two:

Method

Experimental Design

This study used a Single Subject Research Design (SSRD) to investigate the effect of personally relevant stimuli in semantic training of atypical exemplars on discourse performance. This flexible type of design allows one to introduce a treatment to clients while isolating the behavior from other influences (e.g. extraneous influences such as social factors, spontaneous recovery, etc.) and provides strong experimental control to the study (McReynolds & Thompson, 1986; Kearns, 1986). It provides an opportunity to closely investigate performance across behaviors among closely matched individuals, and is the basis on which the effectiveness of semantic complexity training has been established (e.g. Kiran, 2008; Kiran et al., 2010).

A multiple baseline within and between series was applied to investigate training effects of two complex stimulus sets. An ABAB (baseline-treatment-baseline-treatment) design was administered to allow the application of two different treatment phases. Multiple baseline studies are required to have functional independent behaviors and homogenous subjects (Kearns, 1986). A between and within series was applied to implement the comparison of different treatment types, personally relevant and personally non-relevant stimulus sets. The between series served to compare two intervention phases, whereas the

within series compared the clients' behavioral change within the two treatment

phases. Treatment lasted for seven weeks total; it was divided into two treatment

phases consisting of three and a half consecutive intervention weeks each.

To examine the effect of personal relevance, the order of the stimulus sets

was counterbalanced. Hence, two participants received training of items with

greater personal relevance first. The third participant received treatment for

personally non-relevant items first. Table 1 illustrates the timeline for each project

task for this study.

Table 1

Timeline

	Pre-phase	W 1	W 2	W 3	W 4	W 5	W 6	W 7	W 8	W 9	W 10	W 11	W 12	W 13	Post-phase	W16-20
Project Tasks																
Recruitment																
1. Distribute recruitment materials to community 2. Set appointments for screening measurements	x															
Data Collection																
Administering Screening	x															
Baseline Measurement		x														
Intervention I																
1. Semantic treatment for 4 consecutive weeks 2. Treatment Probes		x	x	x	x											
Data Collection																
Baseline Measurement					x											
Intervention II																
1. Semantic treatment for 4 consecutive weeks 2. Treatment Probes						x	x	x	x							
Outcome Measurements									x	x						
Follow-up																x

Note. W = Week.

Participants

Three monolingual native speakers and readers of English with fluent aphasia participated in this study.

Inclusion criteria were based on measurements administered by the Principal Investigator (PI) of this study. Two standardized language tests, the *Western Aphasia Battery* (WAB) (Kertesz, 1982) and the *Boston Naming Test* (BNT) (Kaplan, Goodglass, & Weintraub, 2001) were administered prior to any other testing to determine the aphasia type and severity of each potential participant. The eligibility of each participant was based on the following inclusion criteria (inclusion criteria are the same as used by Kiran et al., 2003/2008/2011).

Participants were selected based on the following criteria: 1) classification of mild to moderate fluent aphasia (Wernicke's, transcortical, and/or anomic aphasia) as measured by the WAB and the BNT, 2) a single stroke in the left cerebral hemisphere with damage to the temporal lobe or temporal-parietal area determined by a CT/MRI scan as self-reported and/or stated in medical records if available, 3) onset of stroke at least 6 months prior to the baseline phase of the study, 4) pre-morbid right handedness, 5) pure-tone hearing screening at 40 dB HL bilaterally at 500, 1000, 2000 Hz, 6) at least a high school degree, 7) no language treatment at the time of the study onset, 8) monolingual, native English speaker, 9) indication of at least two activities on the presented topics as personally relevant, 10) age 21 to 90 years. The selection of two personally relevant activities merely served as a measure of precaution in the event of a potential ceiling in performance on one topic throughout the conversational

measurement during baseline testing. Participants with a pre-treatment performance level above 65% (40/60) on the BNT were excluded from this study to ensure naming impairments were present.

All participants were diagnosed with mild to moderate fluent aphasia. Diagnosis of the type and severity was established through the WAB and results were verified by licensed certified speech-language pathologists (SLPs). Participants with an Aphasia Quotient of <50 on the WAB were excluded from this study due to their severity of aphasia. The WAB is a norm-referenced test designed to evaluate a patient's language functioning following stroke or other acquired neurological disorders. It is also used to provide a comprehensive assessment of the patient's language in order to guide treatment. The type of aphasia is determined by an individual's performance on spontaneous speech, auditory verbal comprehension, naming, writing, reading, apraxia, and constructional visuospatial and calculation tasks. Since the WAB merely measures the presence, degree, and type of aphasia, data from this battery was only collected pre-treatment in order to describe the participants.

A summary of the participant's characteristics and scores on the independent measures for each participant can be found in Table 2. Three male individuals participated in this study. Their age ranged from 57 to 74 years (M=67 years, SD=8.71 years). Two participants had a high school diploma; one participant had a Master's degree. Time post onset ranged from 7 to 74 months (M=49.33 months, SD= 36.69 months). The Aphasia Quotient (AQ) ranged from

52.8 to 76.9 (M=68.23, SD=13.55). Performance on the BNT ranged from percentile of 18.3 to 65 (M=36.63, SD=24.91).

Table 2

Participant Characteristics

Participant	Gender	Aphasia Type & Severity	Age	MPO	Education	Aphasia Quotient
P1	M	Transcortical Sensory, moderate	57	7	HSD	75.6
P2	M	Wernicke's, moderate	71	69	HSD	52.8
P3	M	Anomic, mild	74	72	M.S.	76.9

Note. M = Male; HSD = High School Diploma; M.A. = Master of Science

Stimuli Development

Two lists consisting of atypical items served as stimuli for the treatment phase. One stimulus list consisted of items with greater personal relevance, whereas items of the second stimulus list contained items from the 'golfing' category, which served as the personally non-relevant topic for all participants.

Development of personally relevant topics. To determine a personally relevant topic, a list with a variety of different activities, such as golfing, boating and pet care, was presented to each participant. A selection of activities taken from the Life Interest and Values (LIV) Card Sorting System (Haley, Helm-Estabrooks, Womack, Caignon, & McCracken, 2007) was used for the development of stimulus items. The LIV activity is a pictorial, binary-sorting system, which provides an individual with aphasia to self-determine their personal interest by choosing from a selection of specific life interests and values. These interests are divided into four categories, 1) everyday activities, 2)

social activities, 3) recreational activities with high physical demands, and 4) recreational activities with low physical demands. Although not all choices were provided for the participants in this study, at least two items from each category were given. Prior to the finalization of the activity list, five individuals with different types of aphasia and within the same age range as the three participants were consulted to rate the list based on the overall variety of activities provided. Two activities were replaced with activities that were added by three out of five individuals.

Each of the three study participants were instructed to determine two activities of greatest interest to them. To help verify the participant's choice-based preference, a family member or close friend completed the LIV card questionnaire which is composed of identical activities from the LIV card picture choice selection. The family member or friend was instructed to choose activities that the participant with aphasia "wants to do more" or "does not want to do more" (Haley et al., 2007).

Development of stimuli, semantic features, and typicality rankings. Activities identified by the three participants as personally relevant became the ad-hoc categories for stimulus development through the following procedures. The development of stimuli, semantic features, and baseline procedures were partially replicated from Kiran et al. (2011, p. 9-11). A total of 45 normal participants were recruited for the development of the study's stimuli. Students and employees of the University of South Florida assisted in the stimulus development.

The first group consisted of 15 normal individuals (21-63 years of age) who generated as many items as possible for all categories. All collected words were summarized for each category. Verbs and synonyms within each category were eliminated.

The second group, consisting of 15 normal individuals, was randomly assigned to rate the previously generated items of each category on a seven-point scale. The rating scale technique was previously used by Rosch (1975) and used by Kiran & Thompson (2003), Kiran (2008), and Kiran et al. (2011) for stimulus development of atypical and typical items. On this scale, number 1 represents a good example, 4 a moderate fit, and 7 indicates a poor example of this category. The weighted average and standard deviation (SD) for each item was calculated across all normal participants. Items were eliminated (a) whose average typicality rating had a SD of two or more, (b) whose weighted average was below 4.0, c) that consisted of more than three words. Items meeting these criteria were deleted from each category. Fifteen items with the highest ratings were selected for treatment. Following Kiran et al.'s (2011) stimulus development, each category consisted of a final set of 15 atypical items.

The third group of 15 additional normal participants listed as many semantic features as possible for each item provided. Selected categories were *golfing* (control category), *boating, library*, and *yard work*, as selected by the participants with aphasia, representing their most personally relevant activities and a control topic (*golfing*). All participants went golfing once, but stated that they do not like this activity; hence it was agreed on to be non-relevant to them.

Instructions to participants were the same as used by Kiran et al. (2011).

Each atypical stimulus item was grouped with five semantic features pertinent to

the item, five semantic features not pertinent to the item, but belonging to a

different item of the same category, and five semantic features from a distractor

category. The distractor category was library, yard work, or boating. Hence, all

participants were confronted with all categories. Table 3 provides a summary of

the category combinations for each participant.

Table3

Combination of Training and Distractor Categories

Participant	Personally Relevant Topic/ Distractor Category	Personally Non-relevant Topic/ Distractor Category
P1	Boating/ Yard Work	Golfing/ Library
P2	Yard Work/ Boating	Golfing/ Library
P3	Library/ Yard Work	Golfing/ Boating

Treatment Procedures

A summary of the procedural components of this study is provided in a

chart below (Table 4). Data from the BNT was collected pre- and post-treatment

to assess confrontational naming performance of each participant to compare the

generalization of treatment effects on naming, i.e. semantic retrieval

performance. Additionally, two psycholinguistic assessments were administered,

i.e. subtests of the *Psycholinguistic Assessment of Language Processing in*

Aphasia (PALPA) (Kay, Lesser, & Coltheart, 1992) (for semantic processing

testing in all modalities), and the *Pyramids and Palm Trees* (PAPT) (Howard &

Patterson, 1992). Both tests provided vital information about single word

processing performance within a variety of written and spoken modalities

(LaPointe, 2005). These performance patterns of semantic processing were also assessed during pre- and post-testing and are used for data analyses only. They did not contribute to the participant's inclusion criteria.

Baseline measures. Baseline measures included generative naming of all categories and a conversational analysis. Three baseline sessions were administered to establish a stable pattern of naming performance. Guidelines on the data collection during this phase are described in detail below.

For the generative naming baseline, participants were instructed to name as many words associated with each category as they can. There was a two minute time limit. The number of (a) target typical words and (b) target atypical words was tabulated.

Named items were considered correct if they were clear and intelligible productions of the target words, semantically similar variations of the target word, or a very close synonym of the target word. The examiner kept track of (1) untrained atypical words and (2) untrained typical words which represent category exemplars of the typicality norms that were previously established and that are spontaneously generated by each participant. These responses were considered correct untrained words if they were intelligible productions of words that are appropriate for the category and that were identical or consist of semantic variations of items rated previously from the stimuli data set. These procedures were partially replicated from Kiran et al. (2011).

To analyze the discourse performance before and after treatment, CIUs were analyzed during a conversational task. The assessment of CIUs in brain-

damaged individuals has been shown to be a sensitive measurement to determine and detect changes in connected speech samples (Nicholas & Brookshire, 1993). During this baseline measure, the clinician confronted the participant with prompts about a personally relevant topic (e.g. "Tell me about the last time you went *boating*"). A discourse sample elicited by open-ended questions provides an accurate and valid sample of the individual's communicative performance level. Rules for this scoring system are taken from Nicholas & Brookshire (1993b).

Treatment probes. Generative naming probes for all items in training were administered before every second session. These naming probes were used to assess retrieval of trained and untrained items. Generalized retrieval of untrained items was considered to occur when levels of performance change to at least 40% over baseline levels.

Treatment protocol for target stimulus items. Depending on the participant's preference and availability, treatment was provided twice a week for three hours per session or three times a week for two hours per session (6 hours total). Both stimulus lists were counterbalanced, i.e. two participants received treatment of atypical personally relevant items first, while the remaining participant was confronted with the distractor list 'golfing' first.

Treatment was divided in two 3 ½ week training periods. Following the initial baseline measurement, the first training period was introduced. Subsequently, a second baseline measure and the second training phase of the second stimulus list were initiated.

The treatment protocol for this study included four different tasks, replicated from Kiran et al. (2011, p. 25). These tasks were category generation, category sorting, semantic feature selection, and answering yes/no feature questions about a target item. Each task was administered once and the four tasks/activities were presented in the same order during each treatment session.

1. *Category Generation/Naming*. The patient was asked to generate as many examples as possible for the category in training. This task was conducted as a probe measure every second session.

2. *Category Sorting*. The clinician placed two superordinate category cards (e.g. boating and library) and all written category cards on the table (stimulus category and distractor category). The patient was instructed to sort the word cards according to their superordinate category by placing them on the matching category cards (e.g. *mast - boating*). If the patient incorrectly categorized a word, immediate feedback was provided: "*Are you sure that a gooseneck is found in a library? It's actually something you use for boating*".

3. *Feature Generation/ Selection*. The clinician placed the target word at the center of the table and asked the patient to generate as many attributes as he could come up with regarding the target (e.g. *gooseneck*) that make it a good item to fit into the category (e.g. *boating)*. The clinician then presented the patient with the primed features of the item and asked the patient to select the semantic

features that are pertinent to the item. Selected features were read

aloud by or to the participant.

4. *Yes/No Feature Questions.* The clinician will remove the written

   phrases and instruct the participant "*I am going to ask you some*

   questions about (gooseneck) now. Please answer yes or no for each of

   these questions" The clinician asked a total of 15 questions, up to five

   questions that were relevant to the target example (e.g. is made out of

   metal); five that belonged to the category but not to the example (e.g.

   used to sail upwind); five that did not belong to the category (e.g. play

   area for children).

Follow-up

Two out of three participants were available for follow up probes. Follow

up probes measured naming accuracy and conversational speech for both topics

between six to ten weeks after the completion of the study. Ratings of response

accuracy were identical to those procedures used during the treatment study.

Data Analysis

Following Kiran et al.'s (2011) procedure of data analysis, the average

baseline probe scores were subtracted from the post-treatment score to calculate

the effect sizes (ES). Results were then divided by the standard deviation of the

baseline scores. Based on Beeson & Robey (2006), an ES of 4.0 is considered a

small effect, whereas 7.0 is considered as medium, and 10.1 as a large effect.

Table 4

Summary of Procedures

Determination of Personally Relevant Stimuli	Topic Selection by each participant from the LIV card activity task

Baseline Assessment: Week 1	3. Inclusion/ Exclusion Criteria
	4. Screening Measures
	5. Outcome Measures
	    a. Generative naming of items from personally relevant and personally non-relevant categories
	    b. Conversational Task to collect Correct Information Units (CIUs) and Social Validity Ratings
	    c. Boston Naming Test (BNT)
	    d. Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)
	    e. Pyramids and Palm Trees (PAPT)

Intervention I: Weeks 1-4	1. Semantic treatment of stimuli set A (counterbalanced) twice a week for three hours or three times a week for two hours (20 hours total) 2. Naming probes before every second session

Baseline Assessment Week 4	1. Screening measures
	    a. Generative naming of items from personally relevant and personally non-relevant categories
	    b. Conversational Task to collect CIUs

Intervention II: weeks 4-7	1. Semantic treatment of stimuli set B (counterbalanced) twice a week for three hours or three times a week for two hours (20 hours total) 2. Generative naming probes before every second session

Post- Intervention Assessment: Week 8/9	Outcome Measures
	    a. Confrontational naming of items from personally relevant and non-relevant categories
	    b. Conversational Task to collect Correct Information Units (CIUs) and Social Validity Ratings
	    c. Boston Naming Test (BNT)
	    d. Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)
	    e. Pyramids and Palm Trees (PAPT)

Follow-Up Probe: 6-10 weeks after completion of study	a. Generative naming of items from personally relevant and personally non-relevant categories b. Conversational Task to collect CIUs

The ES benchmark for an effective generalization effect is 2.0 with an

improvement of 40% accuracy over baseline levels for the untrained items

(replicated procedures taken from Kiran et al., 2011). To consider treatment of

complex items to be successful and effective, the benchmark of the ES for the

trained items is 6.5 with an improvement to 80% accuracy in confrontation naming for two consecutive treatment sessions.

For inter-judge reliability, two SLPs other than the SLP who originally scored the discourse samples calculated both words and CIUs of each speech sample. All scores were compared. Any rater disagreement was resolved by accepting ratings as accurately scored when two out of three ratings were in agreement.

To investigate the functionality of any potential change in discourse performance, perceptual ratings of the conversational speech samples were collected from naïve raters, as a measure of social validity (SV). SV has been used to measure treatment outcomes, goals and procedures to emphasize a clinically significant change of social communicative contents (Hickey & Rondeau, 2005). Including such a qualitative measure as SV, i.e. a subjective evaluation of the listener, compared to a quantitative analysis like CIU measures, provides a valid discourse analysis to investigate the relation between the speaker and communication context (Jacobs, 2001). SV determines communication abilities measured as a socially relevant change and provides information on generalization of treatment effects to discourse conditions (Jacobs, 2001).

All data were examined for clinical significance. Clinical significance determines if a treatment intervention can be justified by its applied value to a meaningful improvement (Kazdin, 1999). Accordingly, it provides the researcher with valid information about the impact of the treatment on clients and

accompanying future application of the treatment. For the purpose of this study, percentage change indicating clinical significance will be determined based on data range initially provided by Nicholas & Brookshire (1993b) for CIUs (3.2% for %CIUs). A change of at least ten percentage points determined clinically significant change on standardized tests (Katz & Wertz, 1997; Elman & Bernstein-Ellis, 1999)

Reliability

All baseline and treatment sessions were recorded on videotape. Two certified Speech-Language Therapists and one graduate student performed interrater reliability measurements. Reliability on the dependent variable (generative naming) was calculated for 75% of the probe sessions and there was a 100% agreement on probe scoring. An independent observer rated 50% of the treatment sessions for procedural reliability. The observer marked the presence of each of the four treatment tasks in each rated session. The four treatment tasks were rated as present in 100% of the sessions.

Chapter Three:

Results

Direct treatment effects were measured based on analysis of the

generative naming probes, standardized pre- and posttest results, and discourse

performance measured by CIU analysis. Results of atypical training are reported

first, then potential effects between personally relevant and non-relevant

categories are reported. Pre- and post measures including BNT, PALPA, PPT,

and CIU discourse analysis are then reported. Finally, the results of the social

validity measure are described as an indicator of whether discourse changes

were perceptible by naïve raters.

Generative Naming Performance

Total number of items produced in generative naming for the personally

relevant and non-relevant categories during baseline, treatment probes, and final

testing were analyzed for each participant. Words produced at each time point for

each category were sorted based on whether they were atypical items that were

trained during the treatment phase (atypical trained), atypical items that were not

practiced during treatment (atypical untrained), or typical items. Atypical trained

items were separated during the analysis from typical and atypical untrained

items. Typical items were defined as typical when they appeared on the stimulus

list created before the start of the study. Stimuli were judged to be atypical in one

of two ways: 1) they appeared on the stimulus list created before the start of the

42

study and fell in the range of atypicality in the item ratings (see stimulus development) or 2) they did not appear on the stimulus list and were automatically rated as atypical due to their low frequency of occurrence.

P1 participated in 20/20 hours for both treatment phases. He received treatment three times per week for approximately two hours. Baseline, probe, and final testing performances during naming in the personally relevant topic are shown in Figure 1. Visual inspection revealed no stable increase of production of atypical trained items in generative naming probes between the baseline and final testing. For the personally relevant topic, ES for atypical untrained items was 2.31, and for typical untrained items was -0.17. For the personally non-relevant topic, ES for atypical untrained items was 2, and for typical untrained items 0.6. The generative naming data for the personally non-relevant topic are shown in Figure 2. Although there were no notable changes in generative naming as a result of treatment, there was a suggestion of potentially more generalization to untrained items in the personally relevant topic compared to the non-relevant category.

P2 participated in 18/20 hours for the first treatment phase (personally non-relevant topic; golfing) due to time constraints of the participant. He participated in 19.5/20 hours for the second treatment phase (personally relevant topic; yard work) due to time constraints of the participant. No stable performance throughout treatment was observed. Based on the visual observation of Figures 3 and 4, treatment elicited more atypical untrained items
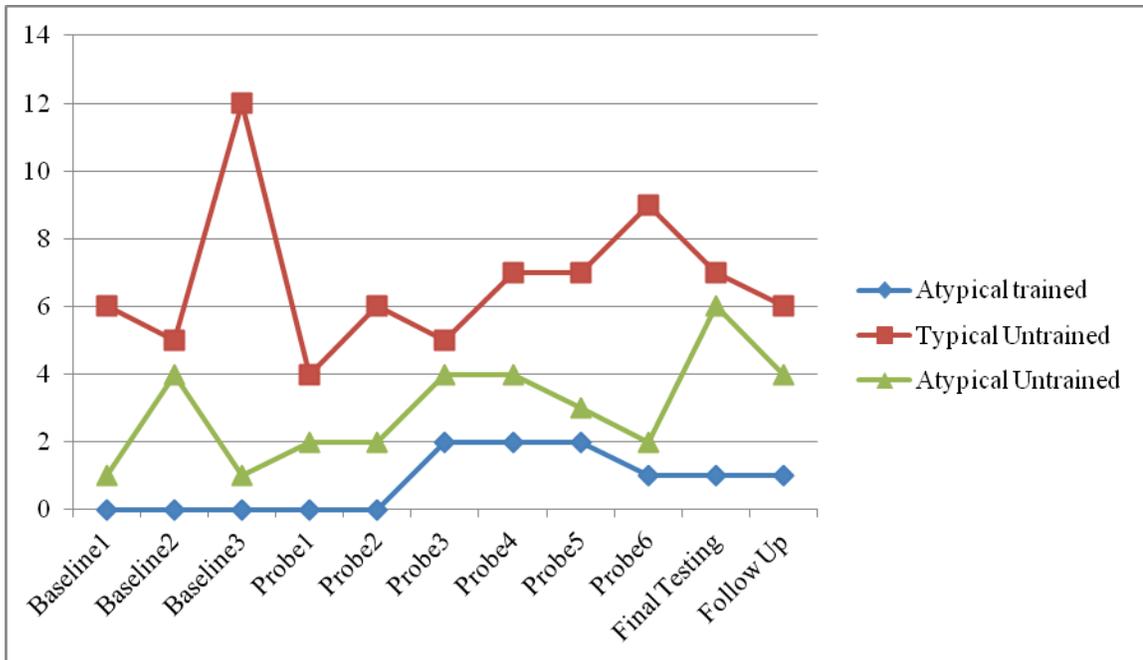
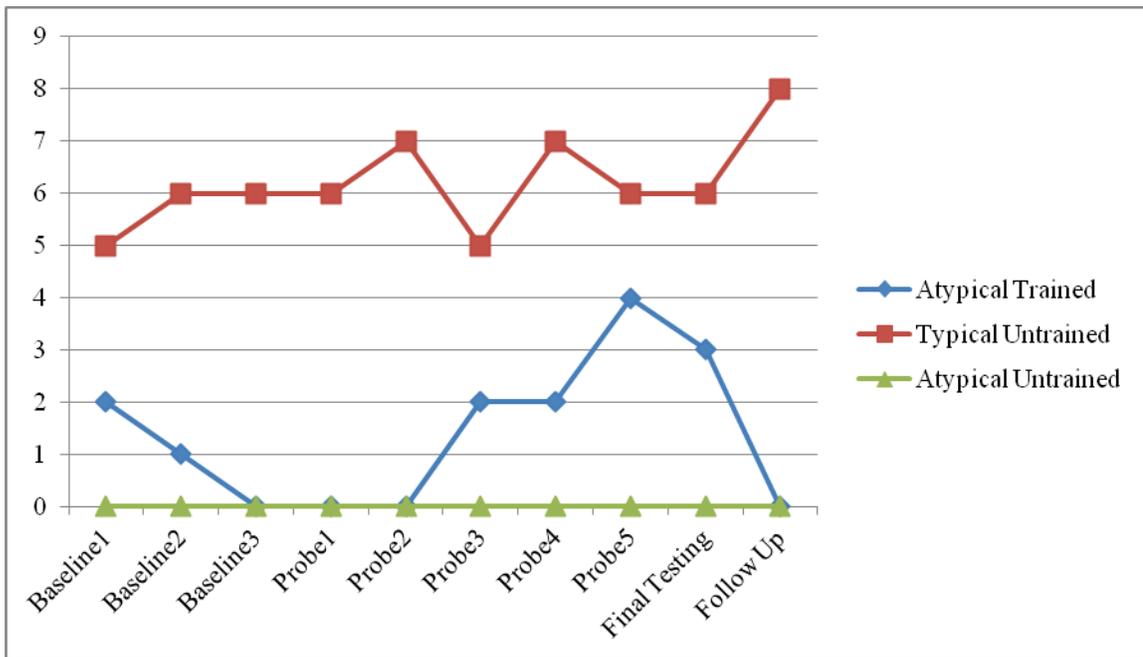*Figure 1.* P1 Generative naming; Personally Relevant Topic 'Boating'

Figure 2. P1 Generative Naming; Personally Non-Relevant Topic 'Golfing'

for yard work (personally relevant topic) compared to atypical untrained items of golfing (personally non-relevant topic).

Performance on atypical trained items revealed no difference meeting the study requirements between baseline and post-training performance on generative naming.

For the personally relevant topic, ES for atypical untrained items was 1.16, and for typical untrained items was 0.45. Probe data are shown in Figure 3. For the personally non-relevant topic, ES for atypical untrained items was -1.16, and for typical untrained items -0.58. None of these effect sizes met the criteria.

The generative naming data for the personally non-relevant topic are shown in Figure 4. Although there were no notable changes as a result of treatment, there was a suggestion of potentially more generalization to untrained items in the personally relevant topic compared to the non-relevant category.

P3 participated in 17/20 hours for the first treatment phase (personally relevant topic; library). He participated in 17.5/20 hours for the second treatment phase (personally non-relevant topic; golfing). P3 was not available for all treatment sessions due to transportation issues. Baseline, probe, and final testing performances during naming in the personally relevant topic are shown in Figure 5. Visual inspection revealed no stable increase of production of atypical trained items in generative naming probes between the baseline and final testing. For the personally relevant topic, ES for atypical untrained items was 3.0, and for typical untrained items ES was 2.89. For the personally non-relevant topic, ES for atypical untrained items was -0.29 and for typical untrained items 1.78.
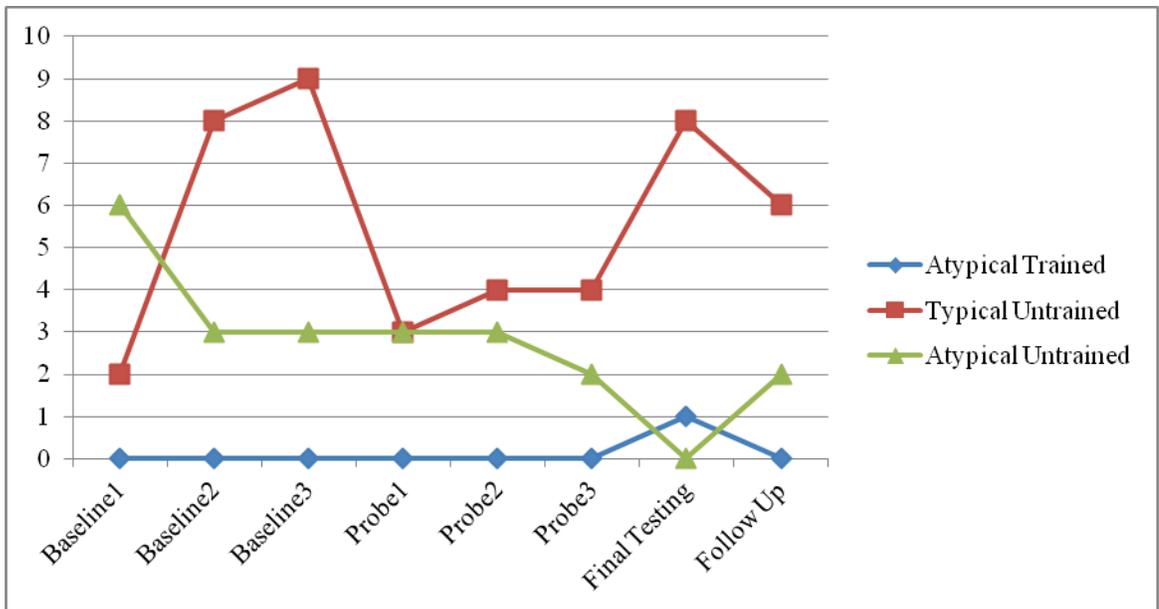
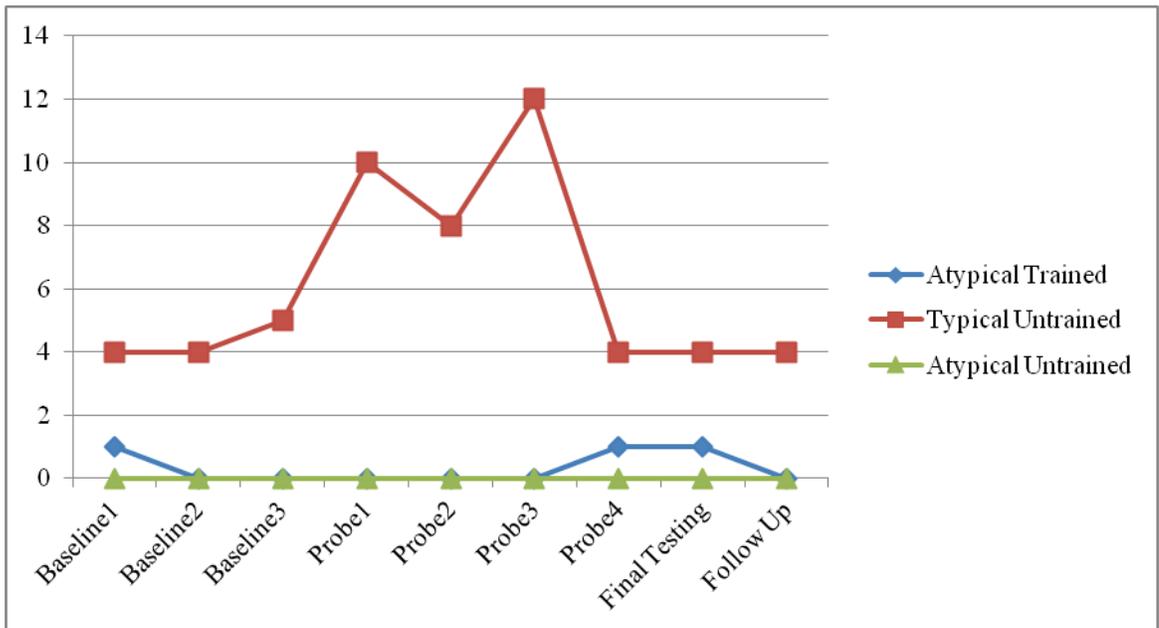*Figure 3.* P2 Generative naming; Personally Relevant Topic 'Yard Work'

Figure 4. P2 Generative Naming; Personally Non-relevant Topic 'Golfing'

The generative naming data for the personally non-relevant topic are shown in Figure 6.
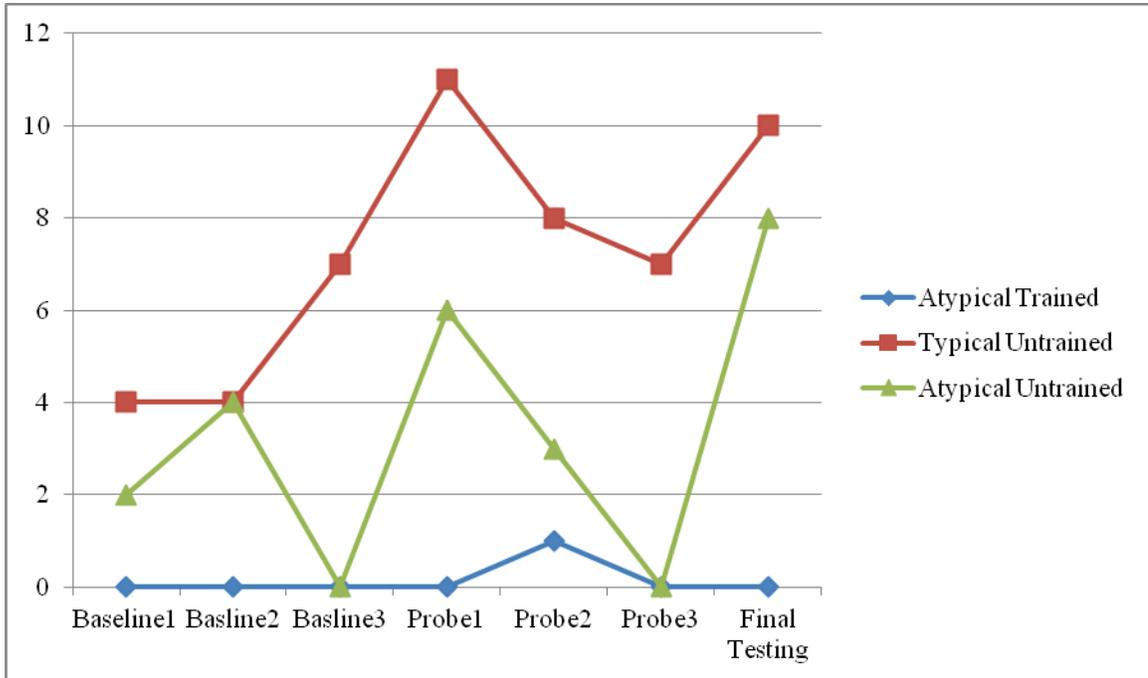
Figure 5. P3 Generative naming; Personally Relevant Topic 'Library'
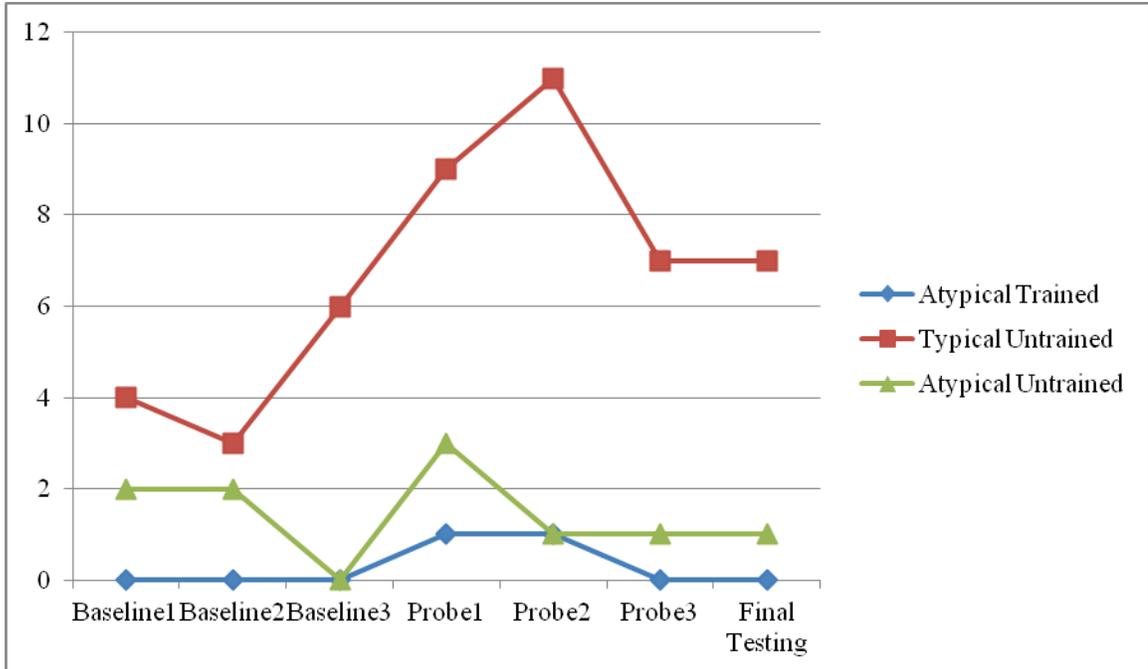
Figure 6. P3 Generative naming; Personally Non-relevant Topic 'Golfing'

47

Although there were no notable changes as a result of treatment, there was a suggestion of potentially more generalization to untrained items in the personally relevant topic compared to the non-relevant category. P3 was not available for follow-up measures due to personal reasons.

No effect was observed for either trained or untrained items produced during generative naming for any participant, based on visual inspection and effect sizes. Regarding a difference in direct naming effect for personally relevant and personally non-relevant items, results indicate that there was no difference in performance for all three clients for atypical trained items between personally relevant and personally non-relevant categories, although there was the suggestion of more generalization to untrained items in the personally relevant topic for all three participants.

Standardized Test Performance

Difference scores were calculated for each standardized measure by subtracting the pretest percentage score from the post-test percentage score. These scores are shown in Table 5.

Clinical significance determines if a treatment intervention can be justified by its applied value to a meaningful improvement (Kazdin, 1999). A criterion for clinically significant change was set at 10 percent of the total possible points, e.g. a significant clinical change for the BNT was set at 6 points out of 60 possible total points (Katz & Wertz, 1997; Elman & Bernstein-Ellis, 1999).

Pre-test performances and difference scores for all standardized measures for all three participants are shown in Table 5. Pre-test performance on

the BNT was 39/60 (65%) for P1, 11/60 (18.3%) for P2, and 16/60 (26.6%) for P3. P1 achieved clinically significant change with an increase of 23% on the BNT.

Pre-test performance on the PALPA was 68/80 (85%) for Auditory Lexical Decision for P1, 60/80 (75%) for P2, and 77/80 (96.25%) for P3. P2 achieved a clinically significant change with an increase of +11.25% on the Auditory Lexica Decision subtest.

Pre-test performance was 103/120 (85.83%) for Visual Lexical Decision for P1, 110/120 (91.66%) for P2, and 118/120 (98.33%) for P3. P2 and P3 performed near ceiling on both lexical decision subtasks, and there was little room for improvement on this task for any of the three participants.

Pre-test performance for Spoken and Written Word-Picture Matching was 39/40 (97.5%) and 37/40 (92.5%) for P1, 31/40 (77.5%) and 35/40 (87.5%) for P2, and 38/40 (95%) and 40/40 (100%) for P3. P2 achieved a clinically significant change with an increase of +17.5% on the Spoken Word-Picture Matching. Both P1 and P3 performed at a very high pre-test performance level on these subtests, limiting their potential for change.

Pre-test performance for Auditory and Written Synonym Judgment subtests were 50/60 (83.3%) and 53/60 (88.3%) for P1, 36/60 (60%) and 44/60 (73.33%) for P2, and 55/60 (91.66%) and 60/60 (100%) for P3. There were no clinically significant changes observed on this task.

Pre-test performance for Picture Naming was 36/40 (90%) for Spoken Picture Naming, 0/20 (0%) for Writing Picture Names, and 39/40 (97.5%) for

Reading Picture Names for P1. P2 scored 39/40 (97.5%) on Spoken Picture Naming, 16/20 (80%) on Writing Picture Names, and 34/40 (85%) on Reading Picture Names. Pre-test performance was 35/40 (87.5%) for Spoken Picture Naming, 20/20 (100%) for Writing Picture Names, and 40/40 (100%) for Reading Picture Names for P3. P1 achieved a clinically significant change with an increase of +10% on the Spoken Picture Naming subtest.

Table 5

Standardized Test Results

Test	P1		P2		P3	
	Pre	**DS**	Pre	**DS**	Pre	**DS**
BNT (%)	65	**+23.33**	18.3	**+5**	26.6	**+8.4**
PALPA (%)						
Auditory Lexical Decision (%)	85	**+2.5**	75	**+11.25**	96.25	**-1.25**
Visual Lexical Decision Task (%)	85.83	**+3.34**	91.66	**+4.94**	98.33	**+0.83**
Spoken Word-Picture Matching (%)	97.5	**+ 2.5**	77.5	**+17.5**	95	**0**
Written Word-Picture Matching (%)	92.5	**-5**	87.5	**+2.5**	100	**0**
Auditory Synonym Judgments (%)	83.3	**+8.36**	60	**0**	91.66	**-3.33**
Written Synonym Judgments (%)	88.3	**-6.64**	73.33	**+5**	100	**-1.67**
Spoken Picture Naming (%)	90	**+10**	97.5	**+2.5**	87.5	**+7.5**
Writing Picture Names (%)	0	**+0.05**	80	**+5**	100	**0**
Reading Picture Names (%)	97.5	**+2.5**	85	**0**	100	**0**
PAPT-3 Pictures (%)	96.15	**+1.92**	82.69	**5.54**	94.23	**-4.07**

Note. DS = Difference Score

Pre-test performance on the PAPT-3 was 50/52 (96.15%) for P1, 43/52 (82.69%), and 49/52 (94.23%) for P3. Both P1 and P3 performed at a high pre-test performance that precluded improvement.

Results for two participants indicate that their semantic abilities improved based on their overall performance on standardized language tests. P1 and P2 demonstrated clinically significant change on the BNT and on subtests on the PALPA, including lexical decision making and picture naming. A variety of the PALPA subtasks revealed a ceiling effect, specifically for P3, with little room to improve.

CIU Analysis

Changes found after CIU measures to baseline and final testing speech samples were significant enough to fall within the established range of values determined by Nicholas & Brookshire (1993a). In order to reach a clinically significant change in discourse performance, the authors report that the difference between the post-test and pre-test %CIUs score exceed 3.2% (Cameron, Wambaugh, & Mauszycki, 2010). As revealed by the CIU analysis (see Table 6), participants increased their performance by at least +3.2% for one or both topics. No consistent pattern of clinical significance across personally relevant or non-relevant topics was identified.

P1 achieved the greatest improvement, showing clinically significant change for both trained topics (personally relevant and non-relevant). P1 improved from 57.29% to 61.84% (+4.55%) on 'boating', his personally relevant topic. Performance on 'golfing', his personally non-relevant topic, showed an increase of 63.48% to 67.6% (+4.12%).

P2 demonstrated a clinically significant change on his personally non-relevant topic only. His CIU performance increased from 72.77% to 74.06%

(+1.29%) on 'yard work' (personally relevant topic). Performance on 'golfing', his personally non-relevant topic, showed an increase from 80.51% to 85.99% (+5.48%).

P3 showed a clinically significant change for the personally non-relevant topic, 'golfing'. His performance on this topic increased from 56.21% to 70.37% (+14.16%). P3 increased from 52.07% to 53.42% (+1.35%) on his personally relevant topic, 'library', a change that was not clinically significant.

Results of the CIU analysis revealed that all three of the participants achieved clinically significant improvement on content produced during discourse. All three improved their content production in the non-relevant topic, but only one participant (P1) improved in the personally relevant topic.

Table 6

CIU Analysis

Participant	Category trained	# of baselines	% pre($\bar{x}$) /post	% difference	Clinical significance
P1	PR	3	($\bar{x}$) 57.29% 61.84%	4.55%	Yes
	PNR	3	($\bar{x}$) 63.48% 67.6%	4.12%	Yes
P2	PR	3	($\bar{x}$) 72.77% 74.06	1.29%	No
	PNR	3	($\bar{x}$) 80.51% 85.99%	5.48%	Yes
P3	PR	3	($\bar{x}$) 52.07% 53.42%	1.35%	No
	PNR	3	($\bar{x}$) 56.21% 70.37	14.16%	Yes

Note. PR = Personally Relevant; PNR = Personally Non-Relevant

Social Validity

A summary of the difference scores of pre- and post-treatment ratings for each discourse sample is provided in Table 7. All participants were rated by the

same group of nine naive raters. Raters were students and employees recruited from the University of South Florida. Raters were asked to rate each discourse sample based on four parameters: a) Amount of information provided in the narrative, b) Person's ability to transmit the message, c) Person's ability to find the adequate words, d) Degree of ease in retelling the narrative (Cupit, Rochon, Leonard, & Laird, 2010). All discourse samples were randomized for both treatments phases. All participants received the randomized samples in the same order. A 7-point Likert rating scale was provided under each discourse sample (1-extremely poor; 7-extremely well).

Table 7 shows the median ratings across all nine raters for each participant's discourse at pre-test and post-test for the personally relevant and non-relevant topics. SV ratings for personally relevant speech samples showed that P1 final discourse performance was rated lower than his baseline samples for amount of information provided, his ability to transmit the message, and finding adequate words. No change of ratings occurred for ease of retelling the story.

SV ratings for personally relevant speech samples of P2 revealed that his discourse performance increased (+1) for all four parameters.

SV ratings for personally relevant speech samples for P3 showed a decrease of amount of information provided (-1) and no change of transmitting the message, finding adequate words, and the ease of retelling the story. All ratings for personally relevant and personally non-relevant discourse ratings are shown in Table 7.

Also shown in Table 7 are non-personally relevant discourse ratings. SV ratings for P1's discourse performance on 'golfing', the personally non-relevant discourse topic, was rated higher (+2) for his ability to transmit the message, find adequate words, and the ease of retelling the story. The amount of information provided was rated as no change from baseline to final testing.

P2's ratings revealed an increase (+1) for the amount of information provided. The ability to transmit the message and ease of retelling the story received a lower rating from baseline to final testing (-1). No change of discourse performance was noted for finding adequate words.

P3 received lower SV ratings for the amount of information provided and ease of retelling the story. No change was noted for his ability to transmit the message and finding adequate words.

Table 7

Median Social Validity Ratings Across Nine Raters for Pre- and Post-Test Discourse Samples

Relevance of Category	Participants	Parameters							
		Informativeness		Transmission of Message		Word Finding		Ease of Retelling Story	
		PRE	DS	PRE	DS	PRE	DS	PRE	DS
Personally Relevant	P1	4	-1	3	-1	3	-1	3	0
	P2	3	+1	3	+1	3	+1	3	+1
	P3	4	-1	3	0	3	0	3	0
Personally Non-Relevant	P1	3	0	2	+2	2	+2	2	+2
	P2	3	+1	3	1	3	0	3	-1
	P3	4	-1	3	0	3	0	3	-1

Note. Pre = Average Baseline Ratings; DS = Difference Score; PR = Personally Relevant; PNR = Personally Non-Relevant.

Overall, no consistent pattern was observed across the ratings. No difference was noted between ratings on personally relevant and personally non-relevant discourse samples.

A Wilcoxon paired-samples test was conducted to compare discourse performance in pre- and post-test conditions. There was not a significant difference in the scores for pre- and post-testing.

Chapter Four:

Discussion

The purpose of this study was to replicate semantic treatment of atypical exemplars among adults with fluent aphasia, and to extend that work to explore potential differences in outcomes between personally relevant and non-relevant categories. Stimulus development and treatment procedures matched previous work (e.g., Kiran et al, 2011). Ad-hoc categories were selected and stimuli were developed to determine if there was a difference between personally relevant and personally non-relevant treatment material.

Results of this study showed a failure to replicate the magnitude and type of improvements previously reported for the typicality effect, although some improvements were observed. In contrast to the work of Kiran et al. (2011), none of the participants in this study demonstrated an improvement in generative naming probes during the treatment. Two of the three participants (P1 and P2) demonstrated improvements in at least one of the standardized measures of naming and semantic processing. All three participants achieved a clinically significant change in at least one of the discourse tasks.

The hypothesis that training personally relevant categories would produce better treatment outcomes was not supported. Indeed, neither training (personally relevant or non-relevant) produced generative naming improvements.

Observed improvements on standardized tests and discourse measures were distributed across both personally relevant and non-relevant tasks. There are a number of potential reasons for this, but the first factor to consider is the overall failure to replicate the improvements on generative naming that have previously been associated with this treatment (e.g., Kiran & Thompson, 2003; Stanczak et al., 2006; Kiran, 2008).

Failure to Replicate Generative Naming Results

A systematic replication explores "the effects of different settings, therapists or clients on a procedure previously demonstrated as successful in a direct replication series" and predicts the generality of the effectiveness of treatment approaches to other populations, disorders and therapists (McReynolds & Kearns, 1983, p. 112). Systematic (and direct) replications in aphasia research are fairly uncommon, but required to establish the most effective therapy based on a specific type of etiology and the symptoms of individuals with aphasia (Sigurdardottir & Sighvatsson, 2011; Barlow & Hersen, 1984). The importance of replicating semantic treatment for atypical exemplars is crucial for further scientific studies investigating the typicality effect since it generalizes the results beyond the experimental conditions (Jackson, 2009). The failure to systematically replicate previous results of this semantic treatment can help identify the limitations of these treatment procedures as tested and communicated thus far (Gast & Ledford, 2010).

There were two changes in the treatment and probe procedures in this study that could conceivably contribute to the failure to find generative naming

improvements as a result of the treatment. First, this is the only study thus far to investigate the outcomes of atypical training alone, without a comparison to typical training. Although previous studies have shown the positive effect of atypical versus typical training with a greater generalization outcome for the atypical training (Kiran & Thompson, 2003; Stanczak et al., 2006; Kiran & Johnson, 2008; Kiran, 2008; Kiran et al., 2009; Kiran et al., 2011), no study up to this date has investigated the outcome of atypical treatment alone. The current study is the first study to attempt training of atypical items only to elicit a typicality effect to generalized word retrieval over to trained and untrained typical and atypical items.

Preceding studies have used a diverse number of treatment hours and varied intensity of weekly treatment provided. Kiran et al.'s (2011) participants attended between six and ten weeks of treatment for one category (typical or atypical) for two treatment sessions (2 hours each) twice a week. Performance changes were found after three weeks of treatment, or after 12 hours of treatment total. In this study, treatment was scheduled for 20 treatment hours for each of the two treatment categories (personally relevant and personally non-relevant) at a rate of six hours for 3 ½ weeks. This compares to Kiran et al. (2011) who trained one set of atypical items for 40 hours. Future studies should address the amount of treatment hours necessary to establish the greatest treatment for semantic training of atypical exemplars.

Another influencing factor could be that the generative naming task was not included in every treatment, in contrast to previous work (Kiran et al., 2011).

Instead, generative naming tasks were only included for probe purposes, i.e. generative naming was administered every second session to avoid a testing effect by repeated exposure to the naming task (Nickels, 2002). This does constitute a change to the previously published treatment procedure. Kiran et al. (2011) seems to have administered the generative naming task during every session as part of the treatment, and also administered it every second session as the treatment probe. It is unclear whether this means that the generative naming task was administered two times on every second session – once for the probe, and once as part of the treatment. Unfortunately, personal correspondence did not result in any additional clarification of this issue. In any case, the participants in Kiran et al (2011) had a substantial amount of repeated exposure to the generative naming task, compared to the participants in this study, and that could make an important difference in the generative naming results.

Identifying core components of a treatment is critical to the ultimate implementation of any evidence-supported treatment (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005). Whether administration of the generative naming in half the sessions, compared to all of the sessions, produces an important change in the treatment outcome can be a subject for empirical investigation.

Finally, there could be characteristics of the participants between the previously published research and the current study that contributed to differences in outcomes. Although every attempt was made to select participants for this study that met the same inclusion/exclusion criteria as previous work,

there could be characteristics not included in these criteria that might be driving treatment outcome (e.g. cognitive abilities).

Since all previous research investigating semantic treatment of atypical exemplars has been single subject research, the total number of participants across all published studies is relatively small. A total of 24 participants are reported across this work (Kiran & Thompson, 2003; Stanczak et al., 2006; Kiran & Johnson, 2008; Kiran, 2008; Kiran et al., 2009; Kiran et al., 2011). A treatment that has been developed and tested in only single subject designs may be more susceptible to participant differences, and this could prevent outcome replication.

Relationship to Semantic Models of Naming

The theoretical conceptualization for semantic treatment of atypical exemplars assumes that semantic feature representation would trigger the phonological retrieval of words within its semantic neighborhood throughout all ad-hoc subcategories. The lack of widespread improvement observed in the present findings suggest a lower effect on peripheral activation of atypical items to spread over to centered, typical items. This result could be associated with stimulus selection that was too centered and did not reflect the variety of category words of the selected ad-hoc categories. Another underlying cause of these weak results could also be associated with the semantic feature selection itself. Germani & Pierce (1995) state that there is a strong correlation between the importance of semantic features, i.e. the significance that semantic features contribute to the definition of a word's representation, and an individual's comprehension and naming performance. Semantic features used in this study

were not analyzed based on their high and low frequency use. Semantic features used here were selected from a list of features provided by healthy participants. Features given by at least two participants were used for each stimulus word. This feature selection ensured the use of higher frequency features. However, overall feature relevance was not determined in this study.

According to Sartori & Lombardi (2004), relevance values for semantic features differ within category specific words. That is, individuals with aphasia exhibit diverse deficits for category types, such as inanimate/animate, concrete/abstract and well-defined categories. Category-specific semantic deficits were not individually addressed in this study and might have contributed to low effect on semantic feature activation throughout the treatment. Mason-Baughman & Wallace (2013) found that semantic impairments in individuals with aphasia are tied to distinctive feature knowledge, not common feature knowledge. Based on findings by the same authors (Mason-Baughman & Wallace, 2013; Mason-Baughman, *under review*), distinctive feature knowledge within an impaired semantic system is a decisive factor for the individual's language performance. Semantic feature knowledge of participants in this study was not investigated prior to treatment, hence, the possibility exists that there was a correlation between the semantic features trained and the participants' semantic knowledge deficits. It can be concluded that the distinctiveness of features might have influenced the outcome of this study.

Another interpretation of these findings might imply that topics used here underlie a domain specific retrieval of information storage. That is, conversational

speech samples require the interaction of episodic and generic semantic

retrieval. Episodic and semantic memories depend on each other to the extent

that they are both subsystems of long-term memory. The interaction of these two

systems on a long-term memory level allows an individual to retrieve semantic

information in item recognition, as well as to retrieve episodic information in

lexical decision making (McKoon & Ratcliff, 1979). Episodic memory and

semantic memory are therefore especially dependent on each other in the area

of communication, i.e. expressing our thoughts and ideas. Although both systems

have different conceptual functions, the semantic memory stores facts, relations,

and language-related concepts, while the episodic memory stores events in the

form of personal experiences, both systems intertwine in communication

processes.

     As depicted above, a loop between semantic and episodic memory is

established under semantic long-term memory storage. Sensory input stored in

episodic memory is related to experiences in specific temporal and spatial

locations (Crowder, 1976), whereas knowledge of the world around us is stored

in the semantic memory system. Consequently, the semantic memory system is

independent from time and place. This information can yet be retrieved in any

given situation, for example to share thoughts and experiences with other

individuals. Thus, we are able to share autobiographical events, i.e. episodic

memories, by using our semantic resources to communicate these memories.

Conclusively, memories triggered by a golfing event might be more concrete and

easier to retrieve, especially since all participants have only been golfing once in

their lifetime. Golfing is also a more structured activity, i.e. playing a round of golf involves a structured sequence of events whereas all other activities used in this study can vary in their course of action. A broader semantic and episodic area for these activities might potentially be more difficult to retrieve due to an overload of conceptual input.

Limitations and Future Directions

The failure to replicate generative naming improvements as a result of this treatment points to possible limitations of this study with corresponding future directions for further research.

First, participant selection should be further specified. Specifically, cognitive abilities such as attention, executive function, and memory should be assessed in order to better identify the individuals for whom this treatment is appropriate. There is reason to suspect that individuals with lower executive function are less likely to generalize treatment results without specific transfer training (Hinckley, 2011). Since previous research of this treatment has not assessed executive function, it is impossible to say whether that played a role in the disparate findings here. However, it should be investigated in the future.

Assessment of the individual's processing of semantic features should be incorporated into future research to investigate the degree and extent of feature processing. Mason-Baughman (2009) states that an incomplete semantic representation of a word is "partially reflected in the differential impairment of common versus distinctive feature knowledge" (p. 37) and can cause errors in

word comprehension and naming tasks due to a decreased distinctive feature identification.

The frequency of the administration of the generative naming task should also be addressed in future research. Kiran et al. (2011) performed a category generation task at the beginning of every session. This study limited the generation task to every second session to collect probe data. It is therefore unclear if the frequency of category generation influenced the generalization outcome in this study.

Other procedural changes to the treatment, such as incorporating lexical-semantic approaches (e.g., auditory-word to picture matching and/or written-word to picture matching) to facilitate a visual memory alleviating generative naming performance, should also be investigated. An enhanced visual representation of the stimulus items (through orthographic and phonological input) during the categorization task might increase processing of corresponding meaning of the presented item(s).

Appropriate outcome measures also should be re-considered. First, in the case of this study, ceiling effects were present for subtasks measuring lexical phonological input and output performance and visual object recognition which represent major components of the semantic system. The phonological output lexicon is an outcome of those two components of the semantic system and reveals strengths and weaknesses of word retrieval skills. It can be hypothesized that a high performance on these subtasks should be used to tailor the treatment

to the greatest strengths and weakness, such as implementing aforementioned visual support during categorization tasks.

Analyzing discourse samples in behavioral studies is important, because individuals with aphasia are often prone to display difficulties communicating at a discourse level. Discourse performance also provides information about on how well these individuals can communicate with others within their environment (Harris Wright, 2011). Moreover, changes of language performance measured by discourse samples also reflect a more meaningful change associated with the treatment. This study has not shown generalization to generative naming. However, clinical significance was found when a CIU analysis was conducted on discourse samples. Hence, future research should address discourse analysis to gain insights on an individual's ability to coherently provide information for daily communication (Harris Wright, 2011). Moreover, Ulatowska & Olness (2004) state that personal stories are a useful tool to evaluate coherence in discourse performance, as has been examined in this study by personal relevance of stimulus topics.

The use of functional outcome measures, such as measuring discourse samples should also be continued in future studies. Standardized tests often do not provide enough information about communication impairments as found in discourse samples. Including social validity ratings on discourse samples adds an additional important measurement of how a language change carries over to everyday living through objective raters (McReynolds & Kearns, 1983).

Single Subject Research Designs should be continued in future research as a way to define treatment procedures for semantic treatment of atypical exemplars outcomes can then be tested in group designs (McReynolds & Kearns, 1983). Since clinical studies with individuals with aphasia face the challenge of a diversity of symptoms (based on severity and type of aphasia), outcomes can be measured at various time points, specifically at baseline testing, when using SSRDs. This provides the researcher with the control to pinpoint behavioral changes within subjects, which can then be applied to a broader research group. It is also suitable for a small amount of participants which is ideal for aphasia research since it is challenging to find research participants with homogenous symptoms (McReynolds & Kearns, 1983).

Future research should also address the number of treatment sessions necessary to acquire a clinically significant change for generative naming generalization. Treatment effects can potentially be influenced by the number of treatment hours provided, combined with the type of stimuli used (written vs. visual).

Conclusion

This study provides weak support for the use of semantic treatment of atypical exemplars in ad-hoc categories. The study hypothesis that atypical training of personally relevant categories would produce better outcomes than training of non-relevant categories was not supported.

The study does demonstrate the critical role of replication across labs and across researchers in order to identify key issues in the candidacy, procedures,

and outcome measurement of any developing treatment. The results of this study provide focus to the areas needed for further investigation in order to verify previously reported results and specify the conditions under which the treatment produces those outcomes.

The findings of this study also suggest that future research should approach group comparisons of atypical/typical versus atypical treatment only to investigate if there is indeed a difference of performance based on atypical and typical stimuli differentiation. Such a comparison will provide more information about the nature of the semantic treatment of atypical exemplars, i.e. if its effectiveness is potentially triggered by a combination of stimulus frequency with atypical and typical items.

These current findings did not provide a strong support for the generalization of generative naming following a semantic application of the typicality effect; however, on a functional basis, the findings did support the notion that training of atypical items contributes significantly to discourse performance. Although results did not reveal an influence of the personal relevance of the topic, an overall increase in discourse performance was present, indicating that this is an efficient treatment method for individuals with anomia.

References

American Speech-Language-Hearing Association (2010). *Code of Ethics.*
Retrieved from http://www.asha.org/docs/html/ET2010-00309.html#sec1.2

Barlow, D. H. & Hersen, M. (1984). *Single case experimental designs. Strategies for studying behavior change.* New York: Pergamon Press.

Barsalou, L.W. (1983). Ad hoc categories. *Memory and Cognition*, 8, 211-227.

Barsalou, L.W. (1991). Deriving categories to achieve goals. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 27, pp. 1-64). San Diego, CA: Academic Press. [Reprinted in A. Ram & D. Leake (Eds.), *Goal-driven learning* (1995, pp. 121-176). Cambridge, MA: MIT.

Bear, M.F., Connors, B.W., & Paradiso, M.A. (2001). *Neuroscience: Exploring the Brain.* Baltimore, MD: Lippincott Williams & Wilkins.

Beeson, P. M. & Robey, R. R. (2006). Evaluating Single-Subject Treatment Research: Lessons Learned from the Aphasia Literature. *Neuropsychology Review*, 16 (4), 161-169.

Cameron, R. M., Wambaugh, J. L., & Mauszycki, S. C. (2010). Individual variability on discourse measures over repeated sampling times in persons with aphasia. *Aphasiology*, 24(6-8), 671-684.

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.

Crowder, R. G. (1976). *Principles of Learning and Memory*. Hillsdale, NJ: Lawrence Erlbaum.

Cruice, M., Worrall, L., Hickson, L., & Murison, R. (2003). Finding a focus for quality of life with aphasia: Social and emotional health and psychological well-being, *Aphasiology,* 17(4), 333-353.

Cupit, J., Rochon, E., Leonard, C., & Laird, L. (2010). Social Validation as a measure of improvement after aphasia treatment: Its usefulness and influencing factors. *Aphasiology*, 24(11), 1486-1500.

Davis, G. (2000). Aphasiology: Disorders and Clinical Practice. Boston: Allyn & Bacon.

Dell, G.S., Schwartz, M.F., Martin, N.Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in aphasic and non-aphasic speakers. *Psychological Review*, 104(4), 801-838.

Dietz, A. McKelvey, M., Beukelman, D., Weissling, K., & Hux, K. (2006). *Integrating Contextually Relevant Visual Scenes into Aphasia Interventions.* Seminar presented at the American Speech-Language and Hearing Association Annual National Convention, Miami, FL.

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

Edmonds, L.A., Nadeau, S.E., & Kiran, A. (2009). Effect of Verb Network Strengthening Treatment (VneST) on Lexical Retrieval of Content Words in Sentences in Persons with Aphasia. *Aphasiology, 23(3), 402-424.*

Eichenbaum, H., Yonelinas, A.R., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience,* 20, 123-152.

Elman, R.J., & Bernstein-Ellis, E. (1999). The Efficacy of Group Communication Treatment in Adults With Chronic Aphasia. *Journal of Speech, Language, and Hearing Research*, 42, 411-419.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). Implementation Research: A Synthesis of the Literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Fridriksson, J., Holland, A. L., Beeson, P., & Morrow, K.L. (2005). Spaced retrieval treatment of anomia. *Aphasiology*, 19(2), 99-109.

Gauthier, I., Skudlarski, P., Gore, J.C., & Anderson, A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3 (2), 191-197.

Gast, D., & Ledford, J. (2010). *Single-Subject Research in Behavioral Sciences*. New York, NY: Taylor & Francis.

Germani, M., & Pierce, R. (1995). Semantic attribute knowledge in adults with right and left hemisphere damage. *Aphasiology*, 9, 1-22.

Goodglass, H. (1993). *Understanding Aphasia*. San Diego, CA: Academic Press.

Greenwald, M.L., Raymer, A.M., Richardson, M.E., & Rothi, L.J.G. (1995).

Contrasting treatments for severe impairments of picture naming.

Neuropsychological Rehabilitation, 5, 17-49.

Grober, E., Perecman, E., Kellar, L., & Brown, J. (1980). Lexical Knowledge in

Anterior and Posterior Aphasia. *Brain and Language*, 10, 318-330.

Grossman M. (1980).The aphasics identification of a superordinate's referents

with basic level and subordinate level terms. *Cortex*, 16, 459–469.

Grossman, M. (1981). A Bird is a Bird is a Bird: Making Reference within and

without Superordinate Categories. *Brain and Language,* 12, 313-331.

Groussard, M., La Joie, R., Rauchs, G., Landeau, B., Chetelat, G., Viader, F. …

& Platel (2010). When music and long-term memory interact: effect of

musical expertise on functional and structural plasticity in the

hippocampus. *Public Library of Science One*, 5 (10), e13225.

Haley, K. L., Helm-Estabrooks, N., Womack, J., Caignon, D., & McCracken, E.

(2007). A pictorial, binary-sorting system allowing "self-determination"

despite aphasia. American Speech Language Hearing Association,

Boston, MA.

Harley, T. (2001). *The psychology of language*. Hove: Psychology Press.

Harris Wright, H. (2011). Discourse in aphasia: An introduction to current

research and future directions*, Aphasiology*, 25(11), 1283-1285.

Hickey, E. M., & Rondeau, G. (2005). Social validation in aphasiology: Does judges' knowledge of aphasiology matter? *Aphasiology, 19*(3/4/5), 389-398.

Hinckley, J. (2011). A Preliminary Model for Selecting Aphasia Treatment Type Based on Cognitive Profile. Paper presented at the Clinical Aphasiology Conference, Ft. Lauderdale, FL, May.

Hinckley, J., & Packard, M. (2001). Family education seminars and social functioning of adults with chronic aphasia. *Journal of Communication Disorders, 34*(3), 241-254.

Holland, A.L., Fromm, D.S., DeRuyter, F., & Stein, M. (1996). Treatment efficacy: Aphasia. *Journal of Speech and Hearing Research*, 39(5), 27-36.

Hough, M.A. (1989). Category concept generation in aphasia: The influence of context. *Aphasiology*, 3(6), 553-568.

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test*. Thames Valley Test Company.

Jackson, S.L. (2009). Research Methods and Statistics: A Critical Thinking Approach, 3rd Edition. Belmont, CA: Wadsworth.

Jacobs, B.J. (2001). Social Validity of Changes in Informativeness and Efficiency of Aphasic Discourse Following Linguistic Specific Treatment (LST). *Brain and Language*, 78, 115-127.

Kanwisher, N., McDermott, J., & Chun, M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for the Perception of Faces. *Journal of Neuroscience*. 17, 4302-4311.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). Boston Naming Test (2nd edition). Philadelphia: Lippincott Williams & Wilkins.

Katz, R.C., & Wertz, R.T. (1997). The Efficacy of Computer-Provided Reading Treatment for Chronic Aphasic Adults. *Journal of Speech, Language, and Hearing Research*, 40, 493-507.

Kay, J., Lesser, R., & Coltheart, M. (1992). *The Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)*. Hove, U.K.: Erlbaum.

Kazdin, A.E. (1999). The Meaning and Measurement of Clinical Significance. *Journal of Consulting and Clinical Psychology*, 67(3), 332-339.

Kearns, K. P. (1986). Flexibility of Single-Subject Experimental Designs. Part II: Design Selection and Arrangement of Experimental Phases. *Journal of Speech and Hearing Disorders*, 51, 204-214.

Kiran, S. (2007). Complexity in Treatment of Naming Deficits. *American Journal of Speech- Language Pathology*, 16(1), 18-29.

Kiran, S. (2008). Typicality of Inanimate Category Exemplars in Aphasia Treatment: Further Evidence for Semantic Complexity. *Journal of Speech, Language, and Hearing Research*, 51(6), 1550-1568.

Kiran, S., & Thompson, C.K. (2003a). Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain and Language*, 85, 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, 773-787.

Kiran, S., & Bassetto, G. (2008). Evaluating the effectiveness of semantic based

      treatment for naming deficits in aphasia: what works? *Seminars in Speech*

      and Language, 29(1), 71-82.

Kiran, S., & Johnson, L. (2008). Semantic complexity in treatment of naming

      deficits in aphasia: Evidence from well-defined categories. *American*

      Journal of Speech Language Pathology, 17, 389-400.

Kiran, S., Sandberg, C., & Abbott, K. (2010). Treatment for lexical retrieval using

      abstract and concrete words in persons with aphasia: Effect of complexity.

      Aphasiology, 23(7), 835-853.

Kiran, S., Sandberg, C., & Sebastian, R. (2011). Treatment of category

      generation and retrieval in aphasia: Effect of typicality of category items.

      Journal of Speech, Language, and Hearing Research, Published online

      December 20, 2010.doi: 10.1044/1092-4388(2010/10-0117)

Kertesz, A. (2006). *Western Aphasia Battery Revised*. San Antonio: Harcourt

      Assessment, Inc.

LaPointe, L. L. (2005). *Aphasia and Related Neurogenic Language Disorders*.

      New York: Thieme.

Levelt W. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA:

      MIT Press.

Mason-Baughman, M. B. (2009). *Choosing Among Related Foils in Aphasia: The*

  Role of Common and Distinctive Semantic Feature (Doctoral dissertation).

  Retrieved from

  http://etd.ohiolink.edu/view.cgi/MasonBaughman%20Mary%20Beth.pdf?k

  ent1239718973.

Mason-Baughman, M. B., & Wallace, S. E. (2013). The role of commonality,

  distinctiveness and importance of semantic features in persons with

  aphasia. *Brain Injury*, 27(4), 399-407.

McKelvey, M. (2007). *The impact of contextualization and personal relevance on*

  communicative performance by people with aphasia. (Doctoral

  dissertation, University of Nebraska-Lincoln). Retrieved from

  http://digitalcommons.unl.edu/dissertations/AAI3284031/

McKelvey, M., & Dietz, A. (2007). Visual Scene Displays for People with Aphasia:

  An Overview of the Research. *The Nebraska Speech-Language and*

  Hearing Association: Networker, 1(6), 5-12.

McKelvey, M., Hux, K., Dietz, A., & Beukelman, D. (2010). Impact of personal

  relevance and contextualization on comprehension by people with chronic

  aphasia. *American Journal of Speech Language Pathology*, 19, 22-33.

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic

  memory. *Journal of Verbal Learning and Verbal Behavior*, *18*, 463-480.

McReynolds, L. V., & Kearns, K. (1983). Single-Subject Experimental Designs in

  Communicative Disorders Baltimore, Maryland: University Park Press.

McNeil, M. R., Doyle, P. J., Spencer, K. A., Jackson Goda, A., Flores, D., & Small, S. L. (1997). A double-blind, placebo-controlled study of pharmacological and behavioral treatment of lexical-semantic deficits in aphasia. *Aphasiology*, 11(4-5), 385-400.

McReynolds, L. V., & Thompson, C. K. (1986). Flexibility of single-subject experimental designs, Part I: Review of the basics of single-subject designs. *Journal of Speech and Hearing Disorders*, 51, 194–203.

Milner, B. (1962). Laterality effects in audition. In: V.B., Mountcastle (Ed.) *Interhemispheric relations and cerebral dominance* (pp. 177-95). Baltimore, MD: John Hopkins University Press.

Moss, H. E., & Gaskell, M. G. (1999). Lexical semantic processing during speech. In S. Garrod, & M. Pickering (Eds). *Language Processing* (pp. 59-99). Hove: Psychology Press.

National Aphasia Association. (2009). *What is Aphasia [Facts]*. Retrieved from http://www.aphasia.org/Aphasia%20Facts/aphasia_facts.html

National Aphasia Association (2005). *Aphasia Bill of Rights*. Retrieved from http://www.aphasia.org/docs/Bill%20of%20Rights.pdf

Nettleton, J., & Lesser, R. (1991). Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *Journal of Neurolinguistics*, 6, 139-157.

Nicholas, L. E., & Brookshire, R. H. (1993a). Quantifying connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36*, 338-350.

Nicholas, L. E., & Brookshire, R. H. (1993b). A system for quantifying the
informativeness and efficiency of aphasic adults' connected speech.
Journal of Speech and Hearing Research, 36, 338-350.

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

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

Raymer A. M. (2005). Naming and word-retrieval problems. In: L.L, LaPointe
(Ed.). *Aphasia and Related Neurogenic Language Disorders (*p.68-
82). New York: Thieme.

Rogers, T. T. (in press). Computational models of semantic memory. To appear
in *The Cambridge Handbook of Computational Psychology*, R. Sun (Ed.),
Cambridge, UK: Cambridge University Press.

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

Rosch, E. (1975b). The nature of mental codes for color categories. *Journal of
Experimental Psychology: Human Perception & Performance*, 1(4), 303–
322.

Rosch, E., & Mervis, C. B. (1975). Family Resemblance: Studies in the Internal
Structure of Categories. *Cognitive Psychology*, 7, 573-605.

Sartori, G., & Lombardi, L. (2004). Semantic relevance and semantic disorders.
Journal of Cognitive Neuroscience, 16, 439–442.

Scott, L.S., Tanaka, J.W., Sheinberg, D.L., Curran, T., (2006). A reevaluation of
the electrophysiological correlates of expert object processing. *Journal of
Cognitive Neuroscience*, 18, 1453–1465.

Scott, L.S., Tanaka, J.W., Sheinberg, D.L., & Curran, T. (2008). The role of
category learning in the acquisition and retention of perceptual expertise;
A behavioral and neurophysiological study. *Brain Research*, 19, 204-215*.*

Sigurðardóttir, Z. G., & Sighvatsson, M. B. (2011), Treatment of Chronic Aphasia
with Errorless Learning Procedures: A Direct Replication. *Journal of
Speech-Language Pathology and Applied Behavior Analysis*, 5 (1-3), 47-
58.

Skinner, E.I., & Fernandez, M.A. (2007). Neural correlates of recollection and
familiarity: a review of neuroimaging and patient data. *Neuropsychologia*,
45, 2163–2179.

Stanczak, L., Waters, G., & Caplan, D. (2006). Typicality-based learning and
generalization in aphasia: Two case studies of anomia treatment,
Aphasiology, 20(02-04), 374-383.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of
real-world, perceptual expertise. *Psychological Science*, 16, 145–151.

Thompson, C.K. (2007). Complexity in Language Learning and Treatment.
American Journal of Speech-Language Pathology, 16, 3-5.

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

Ulatowska, H. K., & Olness, G. S. (2004). Discourse. In R. D. Kent (Ed.), *The
MIT encyclopedia of communication disorders* (pp. 300–302), Cambridge,
MA: The MIT Press.

Viskontas, I.V., Quian Quiroga, R., & Fried, I. (2009). Human medial temporal
lobe neurons respond preferentially to personality relevant images.
Proceeding of the National Academy of Science of the USA, 106(50),
21329-34.

Wallace, S., & Mason-Baughman, M. (*under review*). Relationship between
distinctive feature knowledge and word retrieval abilities in people with
aphasia. *Aphasiology.*

IRB Approval

USF
UNIVERSITY OF
SOUTH FLORIDA

DIVISION OF RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-5618

July 18, 2012

Stephanie Karidas Communication Sciences and Disorders PCD 1017 4202 East Fowler Avenue Tampa, FL 33610

RE: Expedited Approval for Initial Review IRB#: Pro00007250

Title: Effects of Personal Relevance and Complexity after semantic treatment among adults with fluent aphasia

Dear Ms. Karidas: On 7/18/2012 the Institutional Review Board (IRB) reviewed and APPROVED the above referenced protocol. Please note that your approval for this study will expire on 7/18/2013.

Approved Items: Protocol Document(s): IRB_protocol_report_May_2012.docx

Consent/Assent Documents: Name IC_minimal_revised_July 2012_IRBstudy_7250.docx.pdf Pictographic_informed_consent_minimal_risk_complexity_study_ dissertation.doc.pdf Please note, the informed consent/assent documents are valid during the period indicated by the official, IRB-Approval stamp located on the form - which can be found under the Attachment Tab. Valid consent must be documented on a

copy of the most recently IRB-approved consent form. It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review category:

(4) Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed, they must be cleared/approved for marketing.

(6) Collection of data from voice, video, digital, or image recordings made for research purposes.

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

John Schinka, PhD, Chairperson USF Institutional Review Board

John A. Schinka, Ph.D.

Appendix B:

Atypical Stimuli

Boating	Yard Work	Library	Golfing
Airbags	Aphids	Art show	Bets
Bleach	Basil	Brick	Dimples
Boom	Chicken poop	Cash register	Dinner
Canvas	Children	Envelops	Fence
Cleats	Coffee grounds	Exhibition	Green fee
Competition	Drive way	Music	Handicap
Grommet	Dust	Overhead projector	Lounge
Jib	Eggshells	Puzzle	Notepad
Keel	Firewood	Sofa	Rake
Launch	Lantern	Train table	Status
Porthole	Loppers	Ladder	Sunrise
Protest	Lubricant	Sleepiness	Sunscreen
Spirit	Mole	Voting place	Water
Trouble	Tree trunk	Water Cooler	Woof
Tack	Winter	Lion statue	Space