ABSTRACT

Title of thesis: "GO APHASIA!": EXAMINING THE EFFICACY OF CONSTRAINT-INDUCED LANGUAGE THERAPY FOR INDIVIDUALS WITH AGRAMMATIC APHASIA

Christine Rachel Virion, Master of Arts, 2008

Thesis directed by: Professor Yasmeen Faroqi-Shah
Department of Hearing and Speech Sciences

Recently, high intensity short-term therapy with a heavy emphasis on verbal language (called constraint induced language therapy, CILT) has gained momentum in aphasiology. However, the entire extent of its applicability and limitations has not been fully studied, especially with regard to specific aphasic deficits. This thesis sought to: 1) determine the efficacy of the originally published CILT protocol (o-CILT) with a deficit specific population (four individuals with agrammatic aphasia) and 2) examine the potential effect of a modified CILT protocol, which additionally focused on grammatical accuracy (g-CILT). Results revealed differences between the performance of individuals with agrammatism in this study and previously published CILT data. Findings also demonstrated that participants receiving g-CILT produced more significant gains on tests of aphasia severity and grammaticality, while individuals receiving o-CILT showed more highly significant changes on discourse measures of grammaticality. This paper suggests that, for individuals with agrammatism, CILT in its original form may not evince significant changes on tests of aphasia severity and grammatical production and a grammatical modification appears to increase the efficacy of CILT.

"GO APHASIA!": EXAMINING THE EFFICACY OF CONSTRAINT-INDUCED LANGUAGE THERAPY WITH AGRAMMATIC APHASIA

By

Christine Rachel Virion

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Arts
2008

Advisory Committee:

Professor Yasmeen Shah, Chair
Professor Rochelle Newman
Professor Froma Roth

© Copyright by
Christine Rachel Virion
2008

TABLE OF CONTENTS

List of Tables…….....iv

List of Figures…….....v

Chapter 1: Introduction……..……..1
Constraint Induced Language Therapy……4
Origins of CILT……...5
Techniques of CILT…….....6
Massed intervention…….....6
Constraint of compensatory strategies…….....6
Shaping……8
Theoretical underpinnings…….....8
Learned non use……...8
Cortical reorganization……...…..9
Investigations of CILT efficacy……12
Agrammatic aphasia……...16
Adaptation theory……...18
Treatment approaches for agrammatic speech……...20
Purpose of this study…..……23
Research questions and hypotheses……...24
Chapter 2: Methods……..……..27
Design……27
Participants…….28
Inclusionary criteria……...28
Exclusionary criteria……..29
Language Measures……...31
Aphasia severity measures…….31
Grammatical accuracy measures……33
Assessment administration…….37
Procedure……...38
Stimuli……40
Intervention task…….41
Constraints…….44
Verbal output constraint…….44
Material difficulty constraint…….45
Complexity rule constraint…….45
Grammaticality constraint and grammaticality judgments……48
Cueing……48
Data analysis……..51
Scoring and reliability……51
Statistical analyses…….52
Comparisons of this thesis with Maher, et al. (2006)……53

Chapter 3: Results……55
Research question #1: Determining the efficacy of CILT for agrammatism…….55
Aphasia severity measures…….55
Western aphasia battery-R aphasia quotient……..55
Boston naming test…….56
Object and action naming test……63
Summary of findings……..67
Research question #2: Determining the efficiency of g-CILT……...67
Aphasia severity measures……68
Western Aphasia Battery-R AQ……...…….68
Boston Naming Test……..69
Object and Action Naming Test……69
Grammaticality test measure outcomes……70
Verb inflection test ……....70
WAB Spontaneous speech ……71
Discourse measures of grammaticality……..76
Summary of findings……..77
Chapter 4: Discussion and future implications……81
Response to o-CILT by individuals with agrammatic aphasia……..81
Post treatment performance of individuals with agrammatism…….82
Maintenance of gains in o-CILT from individuals with agrammatism….86
Response to g-CILT by individuals with agrammatic aphasia……..88
Post treatment outcomes of aphasia severity measures…….88
Maintenance of aphasia severity measures……90
Post treatment outcomes of grammaticality measures……...92
Maintenance of grammaticality measures…….....98
Study Limitations……...99
Limitations of participant cohort and testing measures……...100
Limitations of protocol and analysis measures……101
Implications and future directions……...102
Implications……..102
Future directions……..103

LIST OF TABLES

Table 1: Participant characteristics and baseline test scores…….29

Table 2: Grammaticality measure formulae for discourse production…….37

Table 3: Descriptions of stimuli used in the o-CILT and g-CILT dual card game…...…40

Table 4: Description of complexity of utterance constraint……..48

Table 5: Actual scores, change scores, and effect sizes for participants on WAB AQ scores…….56

Table 6: Actual scores, change scores, and effect sizes for participants on BNT scores…….58

Table 7: Actual scores, change scores, and effect sizes for participants on O&A and ANT scores……..61

Table 8: Actual scores, change scores, and effect sizes for participants on VIT scores……....68

Table 9: Actual scores, change scores, and effect sizes for participants on WAB spontaneous speech grammatical content scores……...……70

Table 10. Discourse sample grammatical analyses reported in proportions (proportion of sentences (S), proportion of well-formed sentences (WFS), tense accuracy (TA), and tense variety (TV) for individuals who received o-CILT…….72

Table 11. Discourse sample grammatical analyses reported in proportions (proportion of sentences (S), proportion of well-formed sentences (WFS), tense accuracy (TA), and tense variety (TV) for individuals who received o-CILT……..73

LIST OF FIGURES

Figure 1: The proposed role of CI therapy in overcoming learned non use in aphasia.....9

Figure 2: Examples of stimuli……....42

Figure 3: o-CILT participants utilizing a visual barrier ……44

Figure 4: g-CILT participants play the dual card game with a card holder……..45

Figure 5: Western Aphasia Battery-R Aphasia Severity Quotient (Kertesz, 2007) raw score changes from baseline to post treatment (max score =100)…….55

Figure 6: Boston Naming Test (Kaplan, et al., 2000) raw score changes from baseline to post treatment (maximum score= 60)……57

Figure 7: Object and Action Naming Battery (Druks & Masterson, 2000) raw score changes from baseline to post treatment (maximum score= 60) for participants in this thesis and Action Naming Test (Nicholas, et al., 1985) raw score changes from baseline to post treatment (maximum score=unknown) for Maher, et al. (2006) participants……60

Figure 8: Verb Inflection Test (Faroqi-Shah, unpublished) raw score changes from baseline to post treatment (maximum score= 20)……..……....67

Figure 9: Western Aphasia Battery Spontaneous Speech Fluency, Grammatical Competence, and Paraphasia Score (Kertesz, 2007) raw score changes from baseline to post treatment (maximum score= 10)……68

CHAPTER 1: INTRODUCTION

Each year, approximately 780,000 people have a new or recurrent stroke in the United States (AHA, 2008). It is estimated that two-fifths of these stroke victims, or 312,000 individuals, will be diagnosed with aphasia (Pedersen, Jorgensen, Nakayama, Raaschou, & Olsen, 1995). Aphasia is a condition which results from brain damage and can cause deficits in areas such as word retrieval, sentence formulation, grammatical speech production, and/or comprehension (ASHA, 2007). The prevalence of this condition motivates researchers to establish efficacious approaches for the rehabilitation of individuals with aphasia. While speech and language therapy for patients with aphasia is generally considered efficacious (Holland, Fromm, DeRuyter, & Stein, 1996), efficacy data for specific treatments with this population is limited (ANCDS, 2008).

Beeson and colleagues are in the process of compiling all of the treatment studies used for individuals with aphasia, and then subdividing them by target of intervention (e.g., lexical retrieval, writing, aphasia group therapy) and describing each study in terms of study design (e.g., between group, within group, single subject, or case study design), class of study as defined by the American Academy of Neurology guidelines (e.g., strong, intermediate, weak), and phase of treatment research according to Robey & Schultz, 1998 (e.g. pre-efficacy or efficacy). The strongest class of study is defined as a well designed randomized controlled study, intermediate studies are well designed observational studies with controls or single subject multiple baseline studies, and the weakest studies are expert opinion, case studies, studies utilizing historical controls, or single subject multiple baseline studies across behaviors. There are five phases of treatment research (Robey & Schultz, 1998), and the first three phases are considered pre-efficacy data while phases

four and five are considered efficacy data. Phase one studies determine if there is evidence to suggest the therapeutic value of a treatment; phase two attempts to develop, standardize, validate, and optimize procedures or to explain why a treatment works and identify the ideal candidate population; phase three tests the treatment under ideal conditions; phase four determines treatment effectiveness under ordinary conditions of use, and phase five explores the efficiency, cost-benefit, patient-family satisfaction, and the influence of the treatment on patients' quality of life. The Beeson group is also developing effect size scores on the basis of significant effect sizes which have been found in aphasia treatment literature.

This massive undertaking of identifying the strength and phase of every research study has illuminated the evolution of aphasia treatment research as well as the areas which are not as well explored. The preponderance of treatment studies have focused on weak levels of evidence and in pre-efficacy studies. When examining the specific treatments approaches, it appears that few are able to report generalization of treatment effects to conversational settings or long term maintenance of skills gained in therapy. Limited effectiveness of aphasia treatments may be the result of an overemphasis on deficits associated with aphasia, such as anomia or agrammatism in isolated trials in the clinical setting, and less of a focus on incorporating conversational activities within the therapy setting. It has been suggested that generalization has to be planned and does not occur automatically within administration of impairment-based treatments (Kearns & Thompson, 2006). In addition, learned non use may limit the effectiveness of therapy (Meinzer, Elbert, Djundja, Taub, & Rockstroh, 2007; Sterr, Szameitat, Shben, & Freivogel, 2006; Taub, 2006).

Learned non use describes the difficulty of a patient to use spoken output, not directly as a result of neurological deficits, but as a byproduct of a learned suppression of speech. The proponents of learned non use argue that, at the onset of aphasia, individuals may attempt to use speech, but these attempts may be negatively reinforced when the individual is not able to effectively convey his or her message. After several unsuccessful attempts to use speech, the individual may begin to rely on compensatory strategies, such as gesturing, pointing, or writing, to get their message across, since these are less effortful and more likely to succeed. This develops into a learned non use of the verbal modality (Morris & Taub, 2001; Sunderland & Tuke, 2005; Taub, 2006). A relatively recent approach which has been used to help individuals with aphasia overcome learned non use via intensive practice with verbal communication and constraint of alternates to speech is called constraint-induced language therapy, or CILT (Barthel, Meinzer, Djundja, & Rockstroh, 2008; Breier, Maher, Schmadeke, Hasan, & Papanicolaou, 2007; Breier, Maher, Novak, & Papanicolaou, 2006; Maher, Kendall, Swearengin, Rodriguez, Leon, Pingel, et al., 2006; Meinzer, Streiftau, & Rockstroh, 2007; Meinzer, Djundja, Barthel, Elbert, & Rockstroh, 2005; Meinzer, Elbert, Wienbruch, Djundja, Barthel, & Rockstroh, 2004; Pulvermuller, Hauk, Zohsel, Neininger, & Mohr, 2005; Pulvermuller, Neininger, Elbert, Mohr, Rockstroh, Koebbel, & Taub, 2001; Richter, Miltner, & Straube, 2008; Szaflarski, Ball, Grether, Al-fwaress, Griffith, Neils-Strunjas, et al., 2008). CILT is a somewhat radical departure from the majority of aphasia treatments for several reasons: 1) most traditional aphasia treatments focus on either ameliorating the impairment (e.g., phonological access) or on teaching compensatory strategies (e.g., definition of a word or use of gestures in order to circumvent word

retrieval impairment), while CILT stresses on excessively practicing the impaired modality, 2) traditional therapies typically focus on patient selection criteria with the assumption that therapy techniques need to be matched with patient deficits. In contrast, CILT assumes a "one size fits all" approach, with all prior studies of CILT having used the same treatment protocol with a heterogeneous group of individuals with aphasia.

This thesis sought to 1) examine the efficacy of CILT for individuals with agrammatic aphasia by comparing the results of this thesis with previously published CILT efficacy data, and 2) to assess the effect of adding a grammaticality constraint to the CILT protocol for individuals with agrammatic aphasia when compared to the original CILT protocol. A total of four agrammatic participants were recruited, two participants each received original CILT and the grammatically-modified CILT. The next section reviews the technique and theory behind this treatment, as well as relevant CILT research. This is followed by a description of agrammatic aphasia and a comparison of the theory of learned non use with the adaptation theory of agrammatism. Finally, typical treatment approaches for agrammatism will be reviewed and the research questions and hypotheses for this thesis will be presented.

1.1 Constraint-induced language therapy (CILT)

Constraint-induced language therapy (CILT), also referred to as constraint-induced aphasia therapy (CIAT), initially piqued the interest of aphasia researchers because of its significant results, long-lasting benefits, and successful versatility when used with heterogeneous classifications of patients with chronic aphasia (e.g., fluent, nonfluent, anomic). CILT utilizes the techniques of massed intervention (intensive intervention over a short period) in an aphasia group setting, constraint of compensatory

strategies, and shaping of target behaviors to generate significant functional and clinical gains. In the CILT protocol, participants play a dual card game, much like *Go Fish*, in which they attempt to find matches for the cards in their hand. Participants are required to use speech only and not compensatory strategies, such as gesturing or writing, to communicate. In addition, the complexity of utterances required for communication exchanges progressively increases throughout the game.

1.1.1 Origins of CILT

The three core techniques of CILT (e.g., massed intervention, constraint of compensatory strategies, and shaping) were originally used in the early nineties for rehabilitation of motor skills for stroke-induced hemi-paresis in the physical therapy domain. In constraint-induced motor therapy (CIMT), participants with hemi-paresis constrain their unaffected arm with the use of a sling, and then complete everyday activities and therapy tasks utilizing their affected (paretic) arm. Participants wear the sling for approximately six hours a day, for ten consecutive days, thereby satisfying the massed intervention technique requirements. Therapists utilize the shaping technique by designing tasks which challenge participants, but ensure ultimate success at the task. Modeling and motivation are provided as needed throughout all activities (Morris & Taub, 2001; Sunderland & Tuke, 2005).

1.1.2 Techniques of CILT

Three major techniques are used during CILT: massed intervention, constraint of compensatory strategies, and shaping. These techniques and their applications to therapy are discussed in the following paragraphs.

1.1.2.1 Massed intervention. Massed intervention refers to intensive practice over a short period of time. A meta-analysis of all published efficacy studies of aphasia treatments retrieved by a MEDLINE literature search revealed that, overall, aphasia interventions provided in a time-intensive format of a minimum of 8.8 hours a week for at least 11.2 weeks are more efficacious than aphasia treatments provided for 2 hours a week over 22.9 weeks (Bhogal, Teasell, & Speechley, 2003; Bhogal, Teasell, Foley, & Speechley 2003). CILT capitalizes on the massed intervention technique by administering 24 to 30 hours of the dual card game over the course of 2 weeks (Maher, et al., 2006; Pulvermuller, et al., 2001). The shortest duration of a CILT protocol reported to obtain significant effects utilized fifteen hours of therapy for one week (Szaflarski, et al., 2008), underscoring the importance of massed practice in this treatment.

1.1.2.2 Constraint of compensatory strategies. Compensatory strategies typically used by individuals with aphasia include pointing, gesturing, pantomiming, drawing, and writing (Szaflarski, et al., 2008). The primary constraint in the CILT protocol requires all participants to use only speech to communicate. Participants are discouraged from using compensatory strategies with verbal reminders and the use of a physical barrier, which prevents the each participant from seeing the other participants' lower faces and bodies (see Figure 3). This barrier prevents the successful use of writing and gesturing as means of intentional communication (Maher, et al., 2006) and can be used as a card holder for participants with hemi-paresis (Meinzer et al., 2007). Constraint of compensatory strategies is the only constraint which has been directly examined for contribution to CILT, and it appears as though this constraint may assist in maintenance of therapy gains (Maher, et al., 2006).

Material difficulty constraints and complexity rule constraints provide additional challenges within the CILT protocol. Material difficulty is incorporated as a function of stimuli design. Cards used to play the therapy card game are designed to differ from each other by one feature such as color or number, so that participants' utterances need to be specific and accurate to ensure a successful card request (e.g., *Give me the blue/red table, Give me one/two apple(s)*). If they describe something other than the card they want, participants may not gain the desired match.

> The complexity rule constraint places hierarchically increasing requirements on the production of utterances. This hierarchy is individualized, and depends on the participant's initial performance level. At the most basic level, a participant may be asked to label an object on a playing card (e.g., "Apple."). The second level might require the participant to label the object and make a request (e.g., "Apple. Can I have that card?"), and the difficulty would increase accordingly upon successful completion of each level. Most complexity constraints described in previous CILT studies involve a gradual increase in utterance length and number of utterances. Constraints used in previous studies have not explicitly focused on grammaticality as a criterion for success, and this thesis explores this concept. For example, an utterance such as "Apple give please" would be considered a successful attempt for a three-word utterance complexity constraint despite the utterance's ungrammaticality. As described in a later section, there is one group of individuals with aphasia who primarily produce ungrammatical sentences (section 1.2. Agrammatic aphasia). The specificity of a deficit in agrammatism raises the question of whether existing CILT complexity constraints are suitable and

sufficient to improve utterance quality in individuals with agrammatic aphasia, and this issue is discussed further in a later section.

1.1.2.3 Shaping. Shaping is defined as utilizing cues and assistance from the clinician to produce a successful response according to the complexity constraint level. This approach allows participants to experience success at incremental levels with cueing as necessary to gradually achieving the target goal while minimizing frustration throughout the process. Shaping is accomplished through a procedure which involves providing explicit, positive performance-based feedback and rewards, selecting tasks at which the participant should be successful, assisting the participant in being successful at that level with cueing as necessary, and systematically increasing the difficulty of the task in accordance with participant performance (Morris & Taub, 2001). Shaping in CILT is in some ways analogous to operant conditioning, in which the participant is rewarded for each improvement, but not criticized for any difficulty with the task (Richter et al., 2008). In relation to the dual card task, shaping is achieved by providing cueing as necessary to ensure a successful response which meets the complexity constraint (Meinzer et al., 2007).

1.1.3 Theoretical underpinnings

As mentioned earlier, the theoretical basis of CILT originates from the work of constraint-induced motor therapy investigators. Current findings attribute the significant behavioral changes following constraint-induced therapies to overcoming learned non-use and associated cortical reorganization.

1.1.3.1 Learned non use. Learned non use, as mentioned earlier in this thesis, proposes that individuals with chronic aphasia are inefficient at using the verbal modality

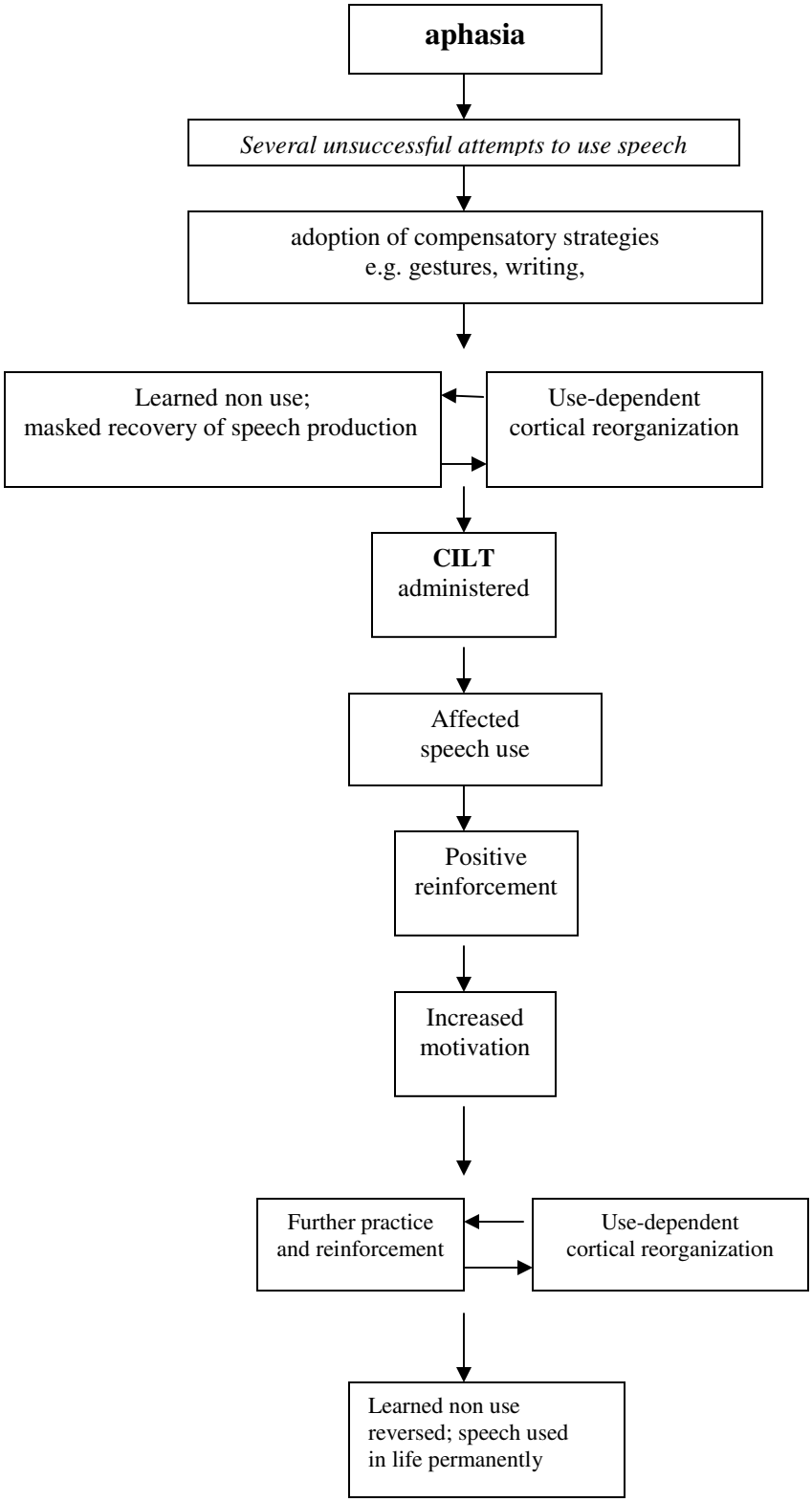
as a byproduct of a learned suppression of speech, over and above the direct result of a neurological insult. Immediately following a stroke, damaged areas of the brain demonstrate depressed neural activity; however, in the weeks following the stroke, initially-damaged neural connections begin to reactivate. This process is known as spontaneous recovery, and it occurs up to a year following the stroke (Saur, Lange, Baumgaertner, Schraknepper, Willmes, Rijntjes, & Weiller, 2006; Thompson, 2000). While the individual with aphasia regains some speech functions as a result of spontaneous recovery, some individuals are less likely to use speech as an effective method of communication as a result of prior failures. Thus, the individuals may not take advantage of their new gains. Constraint-induced therapies theoretically address non use by requiring the patient to utilize the affected process deemed unsuccessful at the onset of the stroke (e.g., CILT- verbal output, CIMT- affected arm use). Proponents of CILT suggest that learned non-use is overcome with massed intervention, constraints, and shaping techniques, participants learn the extent of the speaking abilities, and they become more effective in utilizing verbal output for communication. Positive reinforcement and emphasis on success without frustration are essential for a participant to successfully overcome learned non use. This chain of events is illustrated in Figure 1, adapted from Taub (2006).

1.1.3.2 Cortical reorganization. As discussed above, minimal neural activity occurs in the lesional and perilesional brain regions immediately following a stroke (acute stage) (Morris & Taub, 2001). Attempts to utilize speech during the acute phase may be unsuccessful, but continued trials during the subacute (e.g., 2 weeks post) and chronic (e.g., 1 year) phases of stroke recovery are likely to promote ideal cortical

reorganization, leading to improved efficiency with speech over time and practice (Saur, et al., 2006). It has been postulated that when individuals with aphasia decrease attempts to use verbal output following a stroke, the cortical networks devoted to speech are less likely to be activated, and as a result, hinder further recovery of speech (Lillie & Mateer, 2006; Sunderland & Tuke, 2005). Several studies document cortical reorganization following the administration of CILT and neural changes described in results from these studies correlate to significant behavioral gains on aphasia severity measures (Breier, et al., 2007; Breier, et al., 2006; Meinzer, et al., 2004; Pulvermuller et al., 2005; Richter, et al., 2008). These studies have demonstrated significant reorganization utilizing event-related potential amplitude (ERP), magnetencephalography (MEG), and functional magnetic resonance imaging (fMRI) techniques. This is particularly noteworthy because it has been shown that neural networks enter a stable phase approximately one year following a stroke (e.g., the chronic phase), so it is highly likely that reorganization can be attributed to therapeutic intervention (Saur, et al., 2006).

While the proponents of constraint-induced therapies attribute treatment effects to overcoming learned non use and neural reorganization, the effective components of this therapy technique are hotly debated. One of the reasons is that constraint induced therapies differ from other rehabilitation methods in multiple ways, and it is not clear which aspects actually contribute to its efficacy. In a critical evaluation of constraint induced motor therapies, Sunderland and Tuke (2005) propose that it is not intensity nor constraints that are responsible for the post-treatment changes, but a more efficient use of compensatory strategies.

Figure 1. The proposed role of CI therapy in overcoming learned non use in aphasia, Taub (2006).



1.1.4 Investigations of CILT efficacy

The first study to apply constraint-induced theory to aphasia therapy compared the efficacy of CILT to standard aphasia therapy (Pulvermuller, et al., 2001). Two groups were compared: one condition received thirty-five hours of CILT over the course of ten days (17.5 hours per week) and one condition received thirty-five hours of standard aphasia group therapy over four weeks (8.75 hours per week). Each group was comprised of heterogeneous classifications and severities of aphasia patients (e.g., types- Wernicke, Broca, Transcortical, Conduction; severities- mild, moderate, severe). In addition to differences in intensity between the two groups, the participants engaged in different treatment protocols. CILT participants played a dual card game, in which they obtained matches for the cards in their hand while only using speech to communicate. Physical barriers prevented participants from seeing each other's faces, and the responses were shaped to increase in difficulty as the participants progressed throughout the game. The standard aphasia treatment group received deficit specific "conventional group aphasia therapy", which involved practice with naming, repetition, sentence completion, and conversational tasks. The method of communication involved in the conventional therapy approach was not restricted to speech and involved all modalities of communication.

Pulvermuller, et al. (2001) found substantial post-treatment gains for the CILT group (N=10), but no significant improvements for the standard aphasia group (N = 7). The CILT condition demonstrated gains in multiple areas of deficits, including comprehension, naming, and production. Participants also completed a survey detailing the amount of communication used before and after treatment, quality of communicative

exchanges, and transfer to everyday life and marked improvements were found with the CILT group, but not with the standard treatment group. Additionally, blind raters judged the participants' communicative effectiveness before and after treatment, and scores for individuals in the CILT group increased in these ratings relative to standard therapy participants. These results were remarkable, and the research raised an important question. Since CILT departs from conventional therapies in several respects, which elements of CILT contributed to these significant improvements? Although these investigators demonstrated the importance of the massed intervention technique for aphasia therapy, it was not clear if there were additional factors contributing to the success of CILT. Several subsequent studies have attempted to identify other variables which may contribute to the success of CILT and have achieved better control of variables by matching intensity (Barthel, et al., 2008; Maher et al., 2006).

Maher, et al. (2006) investigated the role of modality constraints within the context of CILT. CILT (N=4) was compared with "Promoting Aphasic Communicative Effectiveness" (PACE) (N=5), which is an approach that utilizes all modalities of communication for treatment, but is similar to CILT in other aspects such as the use of a group format and in activities such as card games. A heterogeneous group of participants in terms of aphasia severities and classifications comprised both treatment groups. Each approach was administered in the same, time-intensive format (three hours a day, four days a week, for two weeks = twenty-four hours), and participants engaged in an identical group task (e.g., the dual card game), with shaping provided throughout both treatments. The only difference between conditions was that participants in CILT were only allowed to use speech to communicate while multimodality communication was allowed for

participants in PACE. Maher, et al. (2006) found that both groups improved significantly in the post-treatment test scores; however, only the CILT participants maintained these gains when retested at a one month maintenance session. Additionally, subjective assessments of narrative discourse quality conducted by trained speech-language pathologists blind to condition and time of sample indicated that CILT members improved the most in their ability to tell a cohesive narrative. These results demonstrated the significant effect of the constraint within the context of CILT, primarily as a means of increasing maintenance of skills. However, a major confound in this study was that three out of five participants who received PACE had severe apraxia, while only one CILT participant was found to have severe apraxia. Hence it is not clear if the co-morbidity of apraxia disadvantaged the PACE group.

Barthel, et al. (2008) attempted to isolate the effects of a deficit-specific focus within a time-intensive format. CILT, as described by Pulvermuller, et al. (2001), was compared with "Model-oriented aphasia therapy", or MOAT. MOAT, which is a typical treatment approach used in aphasia efficacy studies, integrates several approaches to create a treatment which focuses on a specific deficit and is administered in an individual therapy session format. For example, a normal model of language production, such as Levelt's (1989) model or Ellis and Young's (1988) model of lexical representation may be used to identify impaired processes and a treatment protocol is designed to target the impaired process (Chapey, 2001; Mitchum & Berdnt, 1993). Barthel, et al. (2008) ensured that the participants received MOAT therapy in the same time intensive format as participants in the CILT condition (three hours a day for ten consecutive days= thirty hours of therapy). Twelve participants in the MOAT study were given individual therapy

and tested immediately before therapy, immediately after therapy, and six months post therapy and compared to CILT data from twenty-seven previously published participants (Meinzer, et al., 2005), who were tested at the same time points. The results indicated that there was no significant difference in treatment gains between the MOAT and CILT participants on aphasia severity measures, but MOAT participants showed more pronounced improvements on writing and naming subtests, as those skills had been featured explicitly in the MOAT protocol, but not emphasized in CILT. Significant post treatment changes were maintained after six months for both the MOAT and CILT conditions. Barthel, et al. (2008) suggested that, although intensity is crucial for significant treatment gains, deficit-specific training may contribute additional benefits and may be essential for effective, high-intensity treatments for individuals with chronic aphasia. While the results of MOAT and CILT are comparable, the increased gains in naming and writing scores evinced by MOAT relative to CILT are noteworthy. These results highlight the efficacy of deficit-specific treatment relative to general language stimulation treatments. It is worth examining if CILT would provide the same gains as a deficit specific individual treatment (e.g., MOAT) if a deficit-specific modification was incorporated into the group CILT protocol.

To summarize, a review of literature reveals that constraint-induced therapy is efficacious when applied to individuals with chronic aphasia and the massed intervention technique is essential to the success of participants in CILT (Pulvermuller, et al., 2001). The use of constraints may ensure maintenance of skills rehabilitated with CILT (Maher, et al., 2006), and a deficit-specific focus may be necessary to increase the efficacy of CILT relative to individualized treatment (Barthel, et al., 2008). CILT is unique in that it

produces significant benefits for heterogeneous classifications of aphasia through the use of one protocol; however, CILT to date has not been used with a homogenous group of aphasia classifications, as is typical for efficacy studies in this field (see Holland, et al., 1996). Given the shortage of third party reimbursement available for aphasia therapy, CILT is attractive for financial reasons, since it uses a group therapy format (hence cost effective) and it is also practiced over a relatively short period of time. Additionally, the game-like format and opportunity to interact with other patients with aphasia may increase the interest of chronic aphasia patients in speech therapy, thus providing an opportunity for "buy-in" to the therapy. The appeal and success of CILT warrant further efficacy research; however, the Barthel, et al. (2008) finding that CILT is less effective than MOAT is noteworthy. As researchers have pointed out, it is essential to determine which components of CILT are fundamentally necessary to produce significant and long lasting benefits, as this could lead to a more efficacious CILT protocol and increased efficacy of other treatments for individuals with chronic aphasia. One classification of aphasia patients which may benefit from CILT is agrammatic aphasia. In the following paragraphs, agrammatic aphasia is described, and efficacy of treatments for agrammatism is presented. An argument is then made for the use of CILT with this population.

1.2. Agrammatic aphasia

Agrammatic aphasia is characterized by syntactic and morphological (tense marking) impairments (Faroqi-Shah & Thompson, 2007; Thompson & Shapiro, 2005). Agrammatic speech is associated with non-fluent classifications of aphasia (e.g., Broca's aphasia) and it is often described as telegraphic with reduced phrase length, simplified syntactic complexity, impaired production of main verbs and passive sentences, and

omission or substitution of free and bound grammatical morphemes (Chapey, 2001). The following excerpt from a pre-treatment Cinderella narrative typifies many of the characteristics of agrammatic speech. "*Run down stairs in change and her drop in the shoes.*" While the content can be somewhat inferred from the context of the story, this utterance demonstrates difficulty with the use of necessary features for constructing a grammatically-correct sentence. One of the most salient features of this utterance is the lack of tense-encoding morphology. Verbs, such as "run" and "drop" are used in nonfinite form. While one article (e.g., the) is used in this phrase, there are other locations in which the use of an article is grammatically required, but absent. The utterance includes inaccurate uses of pronouns (e.g., "her" for "she") and prepositions (e.g., "stairs IN change" as opposed to "stairs AND change"), as the production of function words tends to be particularly challenging for individuals with agrammatism. Although it is not demonstrated in this utterance, individuals with agrammatism often overuse open class words (e.g., nouns, verbs, adjectives) and have difficulty with the production and comprehension of sentences in which the noun phrase has been moved out of the subject-verb-object order (e.g., canonical sentences such as "Who did the boy kiss?").

Agrammatism has generated considerable interest among aphasiologists and neurolinguists due to its interesting linguistic manifestations. There are several theoretical proposals regarding the underlying impairment in agrammatic aphasia, and a majority of these linguistic proposals focus on syntactic impairments (Friedmann & Grodzinsky, 1997; Grodzinsky, 1986; Kegl, 1995). In contrast, the adaptation theory describes the use of agrammatic speech as a compensatory strategy used to overcome lengthy processing

and production times associated with the formulation of grammatically correct utterances (Kolk & van Grunsven, 1985; Kolk & Heeschen, 1992).

1.2.1 Adaptation theory

Initially presented as the economy hypothesis (Isserlin, 1985), the adaptation theory was formalized by Kolk and colleagues (Kolk & van Grunsven, 1985; Kolk & Heeschen, 1992). The adaptation theory describes agrammatic speech as a strategy used to adapt to the deficits associated with nonfluent aphasia. According to this theory, characteristics such as telegraphic speech, decreased syntactic complexity, and omission of grammatical elements are adopted to overcome a slow rate of speech, avoid computational overload, and maintain the interest of the conversational partner (Heeschen & Schegloff, 1999; Salis & Edwards, 2004). Kolk and Heeschen (1992) discuss the overuse of ellipses in agrammatic speech and present data to suggest that individuals with agrammatic speech overuse ellipses, and therefore produce primarily content words which are absolutely necessary to convey meaning. Ellipses are utterances that lack constructional elements for grammaticality; however, the meaning of these utterances can be inferred and understood from the context of the conversation. Kolk and Heeschen (1992) compared ellipses use by individuals with paragrammatic speech to individuals utilizing agrammatic speech. Paragrammatic speech, which is common in individuals with fluent aphasia (e.g., Wernicke's aphasia), involves morphosyntactic substitution errors. They found that individuals with agrammatic speech produced a significantly higher proportion of ellipses compared to individuals with paragrammatic speech characteristics. Additionally, whereas the individuals with paragrammatism displayed a stable proportion of errors in a variety of elicitation contexts, individuals with

agrammatic speech show showed significant intra-individual variation depending on the elicitation task (e.g., picture description, sentence completion, narration, conversation). It has been noted in several studies (Bastiaanse, 1995; Heeschen & Schegloff, 1999; Salis & Edwards, 2004) that many individuals with agrammatism switch between the use of highly complex grammatically-correct utterances and agrammatic speech.

Based on the individual variability in grammatical competence demonstrated by an individual with agrammatic aphasia, the adaptation theory suggests that the ability to produce grammatically-correct speech is not lost following a stroke; rather, the agrammatic speaker consciously or unconsciously uses an *agrammatic register* to compensate for an underlying sentence processing and/or lexical processing deficit. It may take longer for an individual with agrammatism to process and produce utterances with grammatical components, and use of agrammatic speech lessens this cognitive load. This theory stresses that individuals with agrammatism are not devoid of the ability to use grammaticality; but they will use grammatical speech at the expense of slow rate of speech, frequent hesitations and reformulations (Kolk & Heeschen, 1992). Bastiaanse (1995) compares the strategy adopted by individuals with nonfluent aphasia as a "register", which these patients use to conserve energy and get a message across with increased efficiency. The register of agrammatic speech decreases the requirements for grammatical production, therefore reducing the time and effort afforded to verbal communication.

The adaptation theory of agrammatic speech bears striking resemblance to the theory of learned non use which is proposed to underlie the success of CILT. Both theories emphasize that communication deficits are a result of compensation, or

adaptation, to underlying brain dysfunction and recovery mechanisms. It is stressed in both theories that while these processes are not consciously chosen, they are learned in response to the deficit, and therefore it can be postulated that a successful treatment would address the underlying learned non use or adaptation of communication difficulties instead of the particular deficits associated with the adopted manner of production. According to this rationale, if agrammatic speech is a strategy adopted to increase communicative efficiency at the expense of grammaticality, a treatment which targets overcoming learned non use would be successful at the remediation of agrammatic speech characteristics. This thesis chose individuals with agrammatic speech for the target participant cohort because, based on the adaptation theory of agrammatism, the focus of learned non use which is addressed by CILT should prove to be beneficial for individuals with agrammatic aphasia. A review of literature revealed a paucity of interventions that utilize the framework of the adaptation theory for agrammatic aphasia. Typical treatments for individuals with agrammatic aphasia target the remediation of characteristics of agrammatic speech, and several common treatment approaches are discussed in the following paragraphs.

1.2.2 Treatment approaches for agrammatic speech

The *Sentence Production Program for Aphasia* assumes that agrammatism stems from a difficulty in accessing syntactic information (Helm-Estabrooks, Fitzpatrick, & Barresi, 1981). This intervention trains the production of types of sentences (e.g., imperative intransitive, wh- interrogative, yes/no questions), and the patient learns and uses these sentence formulations in a story completion format and proceeds from less to more complex syntactic structures. There are limited efficacy studies for this procedure.

Individuals showed good acquisition of trained structures, and generalization was found to occur in within-class untrained sentences, sentences with patterns similar to the trained sentences, and some novel sentence structures (Fink, Schwartz, Rochon, Myers, & Socolof, 1995). Generalization to spontaneous speech and narrative productions; however, was limited, and there was no improvement in morphosyntactic production abilities (Doyle, Goldstein, & Bourgeois, 1987).

Response elaboration training (RET) capitalizes on the spontaneous productions of the individual with agrammatism (Kearns & Scher, 1989). The focus of this treatment is on incidental learning. The individual produces an utterance to describe a picture, and the clinician elaborates on what has been produced. Then the clinician attempts to elicit a causal relationship understanding by asking why an agent is performing the action. The clinician elaborates on the provided response again, and then has the individual repeat the entire sentence. This approach requires intensive verbal reinforcement and prompting. Again, very limited efficacy data are available, and while individuals were successful in treatment, overall generalization to spontaneous speech was limited and highly variable between participants in studies (Wambaugh, Martinez, & Alegre, 2001).

In the *Treatment of Underlying Forms* (Thompson & Shapiro, 2005), therapy is based on the concept of treating more complex sentence structures to assist in transfer of explicitly targeted grammatical knowledge to less complex sentence structures, thereby promoting untrained generalization. Sentences trained during treatment address both the argument structure of the verb and syntactic movement of the noun phrase. Argument structure describes the information provided by the verb as to the number of participants in a sentence and their connection to the action. For example, a verb such as “snore” has

one argument, as it only needs one agent to perform this action, and the verb "tickle" requires both an agent and a theme to complete the action. The noun phrase of a sentence can either be merged or moved to address the syntactic movement within a sentence. If the noun phrase is merged, it is combined with the verb to create a sentence (e.g., The boy is sleeping.), and if the noun phrase is moved, it is placed in a different part of the sentence, and can thus affect meaning (e.g., Is the boy sleeping?). In therapy, linguistically complex syntactic targets are used for practice. The approach then trains participants to formulate questions and manipulate the sentence to facilitate grammatical sentence production. By completing this complex sentence manipulation work, patients receiving this treatment have demonstrated significant generalization to syntactically similar and less complex sentence structures.

Mapping therapy assumes that the difficulty in agrammatism is in mapping semantics (thematic roles such as agent, undergoer, etc.) onto syntactic positions (subject, object) in reversible sentences. This approach primarily targets asyntactic comprehension, or difficulty understanding sentences that have multiple verb arguments or show semantic reversibility (e.g., "the boy kissed the girl."). Therapy is conducted by arranging printed materials, and it focuses on proper assignment of agency in active and passive sentence contexts. Research utilizing this approach demonstrates individual success in therapy, but generalization to untrained sentences is limited (Rochon, Laird, Bose, & Scofield, 2005).

To summarize, most existing treatment approaches for agrammatic production are impairment-based, attempting to address the impaired syntactic processes. Some approaches, such as Treatment of Underlying Forms and Mapping Therapy, have

multiple efficacy studies; whereas other approaches have limited empirical validation. While some studies have found impressive generalization to linguistically similar untrained sentences in elicited tasks, most studies have found limited generalization to spontaneous speech. This difficulty with generalization has limited the clinical and functional utility of these existing approaches for agrammatic aphasia. It is possible that, although individuals with agrammatism develop syntactic competence as an outcome of traditional impairment-based therapies, limited generalization to conversational and spontaneous speech is a result of learned non use. That is, grammatical competence may be amenable to traditional impairment based therapies, but the effects of these therapies only show up in elicited speech contexts. A different approach is warranted to overcome the learned non use that limits functional gains.

1.3 Purpose of this study

As CILT gains momentum as a mainstream approach for aphasia rehabilitation, there remain several unresolved issues that warrant further investigation. First, given the heterogeneous participant cohorts of previous CILT studies, it is necessary to examine the effects of CILT for a deficit-specific population. Recent research (Barthel, et al., 2008) proposed that a deficit-specific focus within the structure of a time-intensive aphasia treatment produces greater gains than CILT alone. CILT is cost effective and participants are motivated to engage in therapy because of its emphasis in social interactions, thus making this treatment approach an attractive vehicle for chronic aphasia rehabilitation. It is necessary to examine the effects of CILT for a deficit-specific population so that this information can be used to determine suitability of various treatment options and to predict the likelihood of success.

Secondly, most treatment approaches for individuals with agrammatic aphasia lack generalization to spontaneous speech, and it is necessary to develop treatments which demonstrate functional carryover of skills learned in the clinical setting. Studies suggest that CILT has the potential for this type of generalization. It is possible that individuals with agrammatic speech may have been included and/or benefited from previous CILT studies but aphasia severity measures, such as the commonly used Western Aphasia Battery Aphasia Quotient (AQ) (Kertesz, 2007), can improve even if grammaticality does not improve. The AQ is a composite score of aphasia severity, and does not reflect grammatical well formedness. This score is based on comprehension, repetition and naming tests. Hence, it is possible that participants with agrammatism in prior CILT studies may have improved in severity scores, but not in grammaticality measures. Because these changes have not been thoroughly documented, more distinctive measures directed at grammaticality are needed to properly assess the effect of CILT for agrammatic aphasia.

1.4 Research Questions and Hypotheses

The primary purpose of this thesis was to investigate the efficacy of CILT for individuals with agrammatic aphasia. The following research questions were posed:

Question #1: Does CILT produce improvements in aphasia severity measures for individuals with agrammatic aphasia and are these improvements comparable to what has been found in previous CILT studies (specifically, Maher, et al., 2006)? In order to address this question, CILT was administered to two individuals with agrammatism, and the outcomes were compared with those published for a heterogeneous group of individuals with aphasia. Severity measures obtained from the following tests were

compared with values obtained by Maher, et al. (2006). The present results were compared with Maher et al. because this group produced one of the few studies with English-speaking participants to report individual participant scores (to enable comparison) and use commercially available English tests: Western Aphasia Battery-Revised, Aphasia Severity Quotient (Kertesz, 2007), Boston Naming Test (Kaplan, 2000), and the Object and Action Naming Test (Druks & Masterson, 2000).

It was hypothesized that if generalized mass practice and constraint of compensatory strategies alone were beneficial to the rehabilitation of agrammatism, then improvements in aphasia severity scores from this thesis would be comparable to those obtained by previous CILT investigators. If massed practice and constraints alone were not sufficient to evince significant change, then it was hypothesized that individuals with agrammatism would not demonstrate changes in severity scores. The secondary purpose of this paper was to determine the benefit of a deficit-specific modification to CILT for individuals with agrammatism.

Question #2: Would a modification to the CILT protocol to include a grammaticality constraint produce greater changes in aphasia severity and grammatical measures than results obtained by CILT in its original form? To answer this question, grammatically-modified CILT was administered to two participants and results were compared with data from two participants who received original CILT. The following measures were used to compare aphasia severity between individuals receiving grammatically-constrained CILT and those receiving original CILT: Western Aphasia Battery- Revised, Aphasia Severity Quotient (Kertesz, 2007), Boston Naming Test (Kaplan, 2000), and the Object and Action Naming Test (Druks & Masterson, 2000).

In order to specifically examine the change in grammatical well formedness in agrammatic patients, the following measures of grammaticality were compared between participants receiving original CILT and participants receiving grammatically modified CILT: Verb Inflection Test (Faroqi-Shah, unpublished), Fluency, Grammatical Competence, and Paraphasias score of the Western Aphasia Battery-Revised, Spontaneous Speech subtest (Kertesz, 2007), and on discourse measures of grammaticality, including proportion of sentences, proportion of well-formed sentences, accuracy of tense, and variety of tense.

On the basis of the finding that a deficit-specific focus produced greater gains than massed practice and constraint of compensatory strategies alone (Barthel, et al., 2008), it was hypothesized that the inclusion of a grammaticality constraint to CILT would produce greater gains than original CILT for individuals with agrammatic aphasia on both measures of aphasia severity and grammatical accuracy. This change would be seen on tests and linguistic measures of grammaticality in spontaneous speech (e.g., proportion of sentences, proportion of well-formed sentences, tense accuracy, and tense variety).

To summarize, the aim of this thesis was to determine the efficacy of CILT for individuals with agrammatism and to compare original CILT (henceforth called o-CILT) to a variation of o-CILT which stressed grammatical accuracy (henceforth called grammatical CILT, or g-CILT) for individuals with agrammatism. Both o-CILT and g-CILT were administered using identical tasks and stimuli. The only difference between treatments was the type of constraints employed.

CHAPTER 2: METHODS

2.1 Design

This was a repeated measures, single-subject design, and the aim of the study was to determine the efficacy of CILT for individuals with agrammatism and to evaluate improvements in aphasia severity and grammatical production measures (dependent variable) for individuals with agrammatic aphasia as a factor of the presence of a deficit-specific grammaticality modification to the CILT protocol (independent variable). The dependent variables were measured at three time points: baseline (within 2 months prior to treatment), post-treatment (within two days after the completion of treatment) and maintenance testing (three to four months post treatment). Since the first question was to compare the findings of individuals with agrammatic aphasia receiving o-CILT with the results of Maher, et al. (2006), test measures used by Maher, et al. were replicated for the design of this study, with two exceptions. First, due to unavailability of testing materials, the Action Naming Test (Nicholas, Obler, Albert, & Goodglass, 1985) used by Maher, et al. (2006) was replaced by another commercially available verb naming test (An Object and Action Naming Battery, the action naming subtest, Druks & Masterson, 2000). Second, additional tasks were included to directly assess grammaticality of speech production, as the second research question sought to compare o-CILT and g-CILT on this dimension. Grammatical ability was evaluated through elicited speech tests and variety of narrative and conversational samples, which are described in section 2.3. Treatment was conducted for a two-week period, and tasks and materials were modeled after those used by Maher, et al. (2006).

2.2. Participants

2.2.1 Inclusionary criteria.

Four volunteers with chronic agrammatic aphasia participated in this study. Participants were recruited through the Aphasia Research Center at the University of Maryland, College Park, and were all enrolled in a previous intensive individual treatment program (Faroqi-Shah, in press), as well as other experimental studies at the Center. All participants in the study had a diagnosis of Broca's aphasia with agrammatic speech. This aphasia classification was operationally defined as an accuracy score below 70% on the Verb Inflection Test (VIT) (Faroqi-Shah, unpublished), as well as discourse production characterized by ungrammatical sentences, paucity of verbs, and syntactic formulation difficulties. The VIT assessed the participants' abilities to produce grammatically-correct verb tenses in a picture description task.

All participants sustained a single, left-hemisphere lesion, and were at least 24 months post stroke (chronic phase, scores: 24-138 months) with a non-progressive aphasia. The corrected vision and hearing of all participants were tested to be within normal limits. The primary language of all participants was American English, and there was no age-limit associated with participation. Age and duration of chronic aphasia have not been found to have an effect on improvement of language function in the CILT protocol (see Meinzer et al., 2005). Other inclusionary criteria were: naming accuracy above 60% as defined by the WAB-R naming subtest, auditory comprehension abilities as defined by above 60% on the WAB-R auditory comprehension subtest, and relatively preserved reading skills, with WAB-R scores above 80% on the reading subtest. Participants also demonstrated comprehension of article use as defined by 60% or greater

on an article contrastive judgment task (Virion, unpublished). Article grammaticality comprehension was assessed to ensure that grammatical use of articles in a sentence context could be elicited, as this was a required step of the CILT protocol. In this task, participants were asked to judge if a phrase utilizing an article and a noun was grammatical or ungrammatical (e.g., "an apple", "an bird"). Participant characteristics and baseline test scores can be found in Table 1.

2.2.2 Exclusionary criteria.

Exclusionary criteria were: the existence of prior neuropsychiatric conditions, noticeable confusion, seizures, drug or alcohol dependence, and transportation issues that precluded the individual from attending all treatment sessions. Participants were also excluded if they had error-free spontaneous speech, as indicated by a score of 100% on the WAB-R spontaneous speech subtest. The BDAE Oral Expression subtest (Goodglass, Kaplan, & Barresi, 2000) and Apraxia Battery for Adults-2 (Dabul, 2000) ruled out the existence of severe apraxia, and all but one participant showed no signs of apraxia. P4 was identified with mild apraxia, as indicated by the standard scoring for the apraxia batteries utilized. A depression screener (adapted from AllPsych online, 2008) and as defined by the *Diagnostic and Statistical Manual of Mental Diseases*, DSM-IV; APA, 1994) was used to detect the presence of depression in potential participants, as it was suggested by Pulvermuller, et al. (2001) that depression negatively impacts performance in this type of therapy. Eligible participants were then contacted by letter and phone call, and those who responded were invited to participate on the basis of availability. All participants were enrolled in community aphasia groups at the time of treatment, but the duration of individual participation in these groups did not exceed one

Table 1

Participant characteristics and baseline test scores

TXT	ID	AGE	SEX	H	MPO	MPIT	WAB Name /60	WAB A.C. /200	WAB Read /82	ACJ /10
o-CILT	P1	68	F	R	138	23	55	149	72	10
o-CILT	P2	56	M	R	24	7	56	197	79	9
g-CILT	P3	62	M	R	59	23	44	137	70	6
g-CILT	P4	45	M	R	24	5	50	200	71	9

Abbreviation Key:

TXT= treatment type; **ID=** thesis participant identification code; **H**= handedness, as reported by participants; **MPO**= months post stroke onset; **MPIT**= months post individual treatment; **WAB Name**= Western Aphasia Battery-Revised, Object Naming subtest baseline score/60 points (Kertesz, 2007); **WAB A.C.**= Western Aphasia Battery-Revised, Auditory comprehension subtest baseline score/200 points; **WAB Read**= Western Aphasia Battery-Revised Reading subtest baseline score(including the following subtests: *Reading comprehension of sentences, reading commands, written word stimulus-object choice matching, written word stimulus-picture choice matching, picture stimulus-written word choice matching,* and *spoken words-written word choice matching*)/82 points; **ACJ=** Article contrastive judgment test baseline score/10 points (Virion, unpublished).

hour per week during the two weeks in which this thesis study was conducted. Additionally, one participant (P1) received less than one hour of individual speech and language services per week from the departmental clinic during the course of the study. This research was conducted with the approval of the University of Maryland, College Park Institutional Review Board (IRB) and all participants provided written informed consent prior to commencing study participation.

Participants receiving each of the two treatments were matched primarily on the basis of time lapsed from individual treatment at the Aphasia Research Center. In this way, any effects which could be attributed to the recency of a previous intensive treatment were controlled to the extent possible. Time post individual treatment is reported in Table 1. After participants were matched on the basis of recency since individual treatment, matching was secondarily conducted on baseline VIT scores and approximate ages, to the extent possible. Because this study focused, in part, on change in grammatical abilities as a function of treatment, matching on the basis of recency of individual treatment for agrammatism and severity of agrammatism as defined by VIT score was considered more important than matching on factors such as overall aphasia severity or age. One individual from each pair was randomly assigned to a treatment condition, and the participant's match was automatically assigned to the alternate treatment condition. P1 receiving o-CILT was matched with P3 in g-CILT and P2 was matched with P4 receiving g-CILT. In this way, participants had a matched pair in the alternate treatment condition to allow for comparison between treatments.

2.3 Language Measures

2.3.1 Measures of Aphasia Severity

As previously mentioned, one of the aims of this study was to compare the outcomes of this thesis with the results from Maher, et al. (2006). Hence, the measures reported by Maher et al. were used to assess change in the severity of aphasia for this thesis. The measures used to assess aphasia severity included:

1. The Aphasia Quotient (AQ) from the Western Aphasia Battery-R, maximum score=100, (WAB-R) (Kertesz, 2007). The WAB Aphasia Quotient is a composite score comprised of subtest scores in areas of spontaneous speech information content and grammaticality, auditory verbal comprehension, repetition, naming, and word finding. According to Shewan and Kertesz (1980), the original WAB (1982) was over 88% reliable for test-retest administrations. As a result, any changes seen on this measure across test administrations should be attributable to therapy changes.
2. The Boston Naming Test, maximum score= 60, (BNT) (Kaplan, et al., 2000) which describes word retrieval abilities based on the patient's ability to provide single word labels for the black and white picture stimuli. Stimuli range from easiest to very difficult to name.
3. The Object and Action Naming Battery, maximum score=50, (O&A) (Druks & Masterson, 2000). Maher, et al. used the Action Naming Test (Nicholas, Obler, Albert, & Goodglass, 1985), but this test is not commercially available, so the Action Naming subtest of the Object and Action Naming Battery was used instead, as these tests provide similar outcome measures. Both of these tests assess the patient's ability to name actions in a black and white picture naming task.

All of the above tests were scored as per published scoring procedures in the test manual. Baseline scores on these standardized measures are provided alongside the patient demographics in Table 1.

2.3.2 Measures of Grammatical Accuracy

The following measures were included to assess the effects of o-CILT and g-CILT on grammatical aspects of speech production:

1. Verb Inflection Test, maximum score=20, (VIT) (Faroqi-Shah, unpublished) This measure required the participant to produce verb morphology in the context of a picture description task where a temporal adverb is provided. Scoring criteria are based on appropriate use of verb morphology in the context of the temporal adverb. Hence, *The girl will wipe the table* is scored as incorrect if produced in response to the cue word *Yesterday*. This scoring point is highlighted here because the grammaticality constraints provided to the g-CILT group were similarly structured. The test-retest reliability for this measure has yet to be established.
2. Western Aphasia Battery-Revised, Spontaneous Speech-Fluency, Grammatical Competence, and Paraphasia Score, maximum score=10 (WAB-R) (Kertesz, 2007). This score is determined on the basis of fluency and grammatical speech abilities, as well as the prevalence of paraphasias a participant utilizes when describing the Picnic picture on the Western Aphasia Battery-R. Possible scores range from zero to ten, with zero representing “no words or short, meaningless utterances” and ten representing sentences of normal length, complexity, and cadence. The subtest test-retest reliability

3. Discourse Measures of Grammaticality

Narratives were elicited using three types of stimuli: narration of the Cinderella story, the Narrative Discourse subtest of the BDAE (Goodglass, et al., 2000), and a description of a cartoon video. The primary purpose of using three types of narrative stimuli was to obtain a larger corpus of narrative speech than has been previously used in CILT literature. The Cinderella narrative was chosen primarily to maintain continuity with the Maher, et al. (2006) group protocol. Although the Cinderella story is technically not a standardized test of grammatical ability, it is administered frequently in aphasia research and has associated normative scores (see Rochon, et al., 2000; Webster, Franklin, & Howard, 2001). As per the typical procedure for eliciting the Cinderella story from individuals with aphasia, participants were allowed to review the picture scenes of a wordless version of a Cinderella story book before retelling the story to ensure that failure to remember story details did not impact the quality of the narrative. The Narrative Discourse subtest of the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass, et al., 2000) was administered. In this subtest, participants are shown a 4-6 picture cartoon sequence and verbally read a story which accompanies the cartoon pictures. Participants are instructed to use the pictures to retell the story. The BDAE includes four of these stories, and a different one was randomly selected at each testing time point (e.g., baseline, post treatment, and maintenance) in order to avoid repeated exposure effects. The third narrative sample was elicited using videos from the "Tom and Jerry's Greatest Chases" DVD (Barbera & Hanna, 2000), which featured a cat and mouse performing action-oriented scenes during a

wordless cartoon. "The Yankee Doodle Mouse" was used at baseline, "Mice Follies" was shown at post treatment testing, and "Salt Water Tabby" was played at maintenance testing. In order to properly assess the effects of CILT on grammaticality in discourse, it was necessary to prompt participants to attempt to use tense encoding morphology with action verbs and to produce sentence structures derived from actions. These clips were ideal for eliciting a grammatically-focused sample, as they did not typically provide any written or verbal language, so it was necessary for participants to describe the actions seen in order to accurately recall the details of the video.

In addition to the narratives described above, conversational samples were elicited a minimum of two times throughout each testing session, and the first 100 words of both samples combined were used for analysis. Participants were informed that they were going to talk informally with the researcher, and although these conversations were being recorded, participants were asked to behave as naturally as possible. It was emphasized to participants that the researchers would not be judging the content of speech, rather the type of sentences used. While this cannot be considered a conversation in the most natural of contexts, it was an attempt to elicit less formal discourse between the participant and another speaker.

All narrative and conversational samples were transcribed by the primary researcher and subsequently coded using the Systematic Analysis of Language Transcripts software program (SALT; Miller & Chapman, 2000). Reliability of transcription and coding is discussed in section 2.6.1. The following lexical and sentence codes were used to compute a variety of grammatical measures: formulaic utterances,

sentence fragments, non-sentences, grammatical sentences, ungrammatical sentences, verb tense accuracy, and verb tense (e.g., simple present, past progressive).

Formulaic utterances were operationally defined as utterances that occur with such a high frequency in normal discourse that they are most likely stored as a unit within the mental lexicon (e.g., you know, oh dear). Sentence fragments were defined as utterances which are so fragmented that it was impossible to consider them as either a sentence or a non-sentence. A non-sentence was defined as an utterance that does not contain a verb, but can stand alone (given conversational context, e.g., The end, The story of Cinderella). This category included single words or phrases. A well-formed sentence was defined by the grammatically-correct inclusion of a verb either in the matrix or embedded clause. An ungrammatical, or flawed, sentence was defined as a sentence which contained a matrix or clausal verb which included enough information to distinguish an argument structure, but used an incorrect tense. Verb tense was coded as correct, incorrect, or absent based on the morphological marking of the verb within the context of the sentence. After grammaticality of the verb was identified, all verbs were further coded for tense (e.g., nonfinite, simple present tense, progressive present tense, perfect present tense, conditional tense). SALT computed the following grammatical measures: proportion of sentences overall, proportion of grammatical sentences, accuracy of tense, and variety of tense based on the definitions of codes provided above. Formulas describing how each of these measurements was calculated can be found in Table 2.

In this study, the linguistic analysis of narrative and conversational samples was more comprehensive than analyses performed by Maher, et al. (2006) in order to document changes in grammaticality. This study chose to code sentences as grammatical

or ungrammatical, on the basis of the operational definitions listed in Table 2, while Maher et al. coded sentences irrespective of grammaticality. Additionally, several previous CILT studies have called for the inclusion of more functional speech assessments to better determine the carryover effect of this treatment into everyday discourse (see Barthel, et al., 2008; Maher, et al., 2006, Meinzer, et al., 2007), so a greater breadth of discourse samples were collected for this thesis. This protocol is limited in that it only collected discourse samples at one point in time, and thus did not account for the tremendous variability of productions from participants with aphasia. The implication of this limit is reviewed in the discussion chapter.

2.4 Assessment administration

The following language measures were administered during all three assessment sessions (baseline, post-treatment and maintenance): WAB-R AQ (Kertesz, 2007), BNT (Kaplan, et al., 2000), Object and Action Naming Battery (Druks & Masterson, 2000), BDAE Narrative Discourse subtest (Goodglass, et al., 2000), Cinderella story retelling, cartoon video description, and a conversational sample. Test order varied across participants. Any test begun on one day of testing was completed within that same session to ensure that the scores were representative of abilities at that point in time. Individuals administering the tests were sensitive to levels of participant fatigue, and scheduled breaks or ended testing at the request of the participant or as deemed necessary by the test administrator. Baseline assessments were administered by the primary researcher or Aphasia Research Center research assistants. Post-treatment testing was administered by research assistants for the Aphasia Research Center only, so that familiarity with the primary researcher would not impact performance. Maintenance

testing was administered to P2, P3, and P4 by the primary researcher, and P1 completed maintenance testing with her individual therapist from the university clinic, who is also associated with the Aphasia Research Center. In this way, a potential familiarity effect was equated for all participants at maintenance testing.

All test measures (e.g., WAB, BNT, O&A, and VIT) used the identical assessment forms at each administration (e.g., baseline, post-treatment, maintenance), which may have produced a repeated administration effect. This potential effect is an undesirable result of a lack of alternate forms for commonly used aphasia tests, and this is addressed further in the discussion chapter. Additionally, all assessments were administered only one time, and this score may not be a true representation of the abilities of these participants. Aphasia patients are often highly variable in their abilities (e.g., severity of speech production, comprehension, anomia), and aphasia research has begun to utilize multiple baseline assessments to identify a stable baseline ability. This thesis did not obtain multiple baselines because of scheduling conflicts; however, the benefit of this is discussed in the final chapter of this paper.

2.5 Procedure

Participants received one of two treatments: original constraint-induced language therapy (o-CILT) or grammatical constraint-induced language therapy (g-CILT). All therapy sessions for both conditions took place in the same quiet therapy room, and participants were seated across from each other. Sessions were video and audio recorded using the Linksys Wireless-G Internet Video Recorder, a Panasonic PV-L750 Camcorder and an Olympus Digital Voice Recorder VN-3100PC for transcription and coding

Table 2

Grammaticality measure formulae for discourse production

1. Proportion of Sentences	([sentences] + [flawed sentences])/[total utterances]
2. Proportion of well-formed sentences	[sentences]/ ([sentences] +[flawed sentences])
5. Accuracy of tense	[correct use of tense]/([correct use of tense] +[incorrect use of tense])
6. Variety of tense	total different/unique tense codes/([correct use of tense]+[incorrect use of tense])

purposes, as well as to ensure treatment fidelity. Both treatments were administered for 3 hours a day for 8 days over the span of 2 consecutive weeks. This was equivalent to 24 hours of therapy. There were two participants in each treatment condition.

2.5.1 Stimuli

The stimuli were identical for both treatment conditions. There were 18 cards with black and white objects, 18 cards with colored objects, and 18 cards with multiple objects in each deck, for a total of 54 cards in each deck. Ten decks were used throughout treatment, and Table 3 includes a description of the stimuli used in treatment. The necessity for a deck including objects with several characteristics (e.g., black and white, colored, multiple) is explained further in section 2.4.3. Since CILT's therapy task is a card game, images were printed and centered on 5x8" index cards, (see Figure 2). Images used on these cards were procured from a number of sources including the International Picture Naming Project (Szekely, et al., 2004) and a Google image search. Each card displayed either a boy or a girl with an object in his or her hand. This was a slight departure from the Maher, et al., (2006) protocol, as this group only included objects on their stimuli, not individuals with objects in their hands. The arrangement utilized in this thesis encouraged the verbal production of agent+action+object sentences by requiring all participants to request a card by describing the individual (agent) and what they were doing (action) to the object. The design of the stimuli and the semantic categories used were adapted from the Maher, et al., (2006) study description. Each deck of cards was organized according to objects that corresponded to a semantic theme (e.g., fruits, vegetables, animals, clothes, kitchen items). There were two sets of cards for each semantic category. As with Maher et al.'s procedure, one set of cards contained high

frequency exemplars of the category (e.g., animals- dog and cat) and the other contained low frequency exemplars (e.g., giraffe and camel) according to Francis & Kucera's (1982) word frequency counts. Exemplars were included primarily on the basis of availability of a clear picture representation of the object, and pictures obtained were ranked according to frequency identified by Francis and Kucera (1982). The exemplars were then split in half to produce a high frequency and low frequency deck of cards for each semantic category. All pictures were identified by two adults without aphasia as being clear pictures of the target exemplars. See Table 3 for further details.

2.5.2 Intervention Task

The primary therapy activity for both o-CILT and g-CILT was a dual card game and, all therapy for both groups was administered by the primary researcher. The objective of this game was for participants to collect matches for all of the cards in their hands, much like the game *Go Fish* (hence the title *Go Aphasia* for this thesis). Participants were initially dealt five cards each, and these were all high-frequency exemplars from the same semantic category. Participants took turns asking each other for matches to the cards in their hands. Each participant had a minimum of 4 turns in each game, and several games were played each session. A turn consisted of one participant (the requester) asking the other participant (the responder) for a specific card, and the responder provided a specific answer to the requester's question. If the responder had the card, he or she would give it to the requester, and the requester would be able to ask for another card. If the responder did not have that card, the requester would draw a card from the deck, and participants would switch roles as requester and responder. The game was over when a player received matches for all of the cards in his or her hand. When a

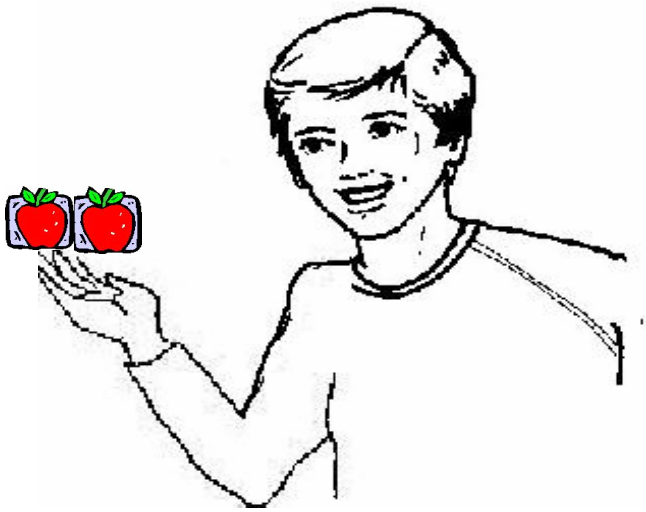
Table 3.

Descriptions and examples of stimuli used in the o-CILT and g-CILT dual card game

Deck number	Category name	High/low frequency	Example of B&W card	Example of colored card	Example of multiple card
1	Animals	High	B&W dog	red dog	2 red dogs
2	Animals	Low	B&W camel	blue camel	3 blue camels
3	Fruits	High	B&W apple	purple apple	4 purple apples
4	Fruits	Low	B&W fig	green fig	1 green fig
5	Vegetables	High	B&W carrot	yellow carrot	2 yellow carrots
6	Vegetables	Low	B&W pumpkin	red pumpkin	3 red pumpkins
7	Clothing	High	B&W shirt	orange shirt	4 orange shirts
8	Clothing	Low	B&W socks	blue sock	1 blue sock
9	Kitchen items	High	B&W dish	green dish	2 green dishes
10	Kitchen items	High	B&W knife	yellow knife	3 yellow knives

B&W = Black and white

Figure 2. Examples of stimuli



game finished, participants played again, this time utilizing a deck featuring the low frequency set of examples from that semantic category (See Table 3 for further explanation). The games utilized a new deck of cards for each round, cycling through high-frequency and then low-frequency semantic exemplars. The decks were not used twice within the same therapy session.

2.5.3 Constraints

Three types of constraints were used with both o-CILT and g-CILT conditions: a verbal output constraint, material difficulty constraint, and a complexity rule constraint. This thesis also included a grammaticality constraint, which was only used by individuals receiving g-CILT. The following paragraphs describe how each constraint was utilized in treatment.

2.5.3.1 Verbal output constraint. Participants receiving both the o-CILT and g-CILT treatments were required to use speech only, and not compensatory strategies, such as gesturing or writing, to communicate. Visual barriers (see Figure 3) and verbal reminders were provided for individuals receiving o-CILT to prevent participants from using alternative modes of communication. This setup ensured exact replication of the Maher, et al. (2006) study. Participants receiving g-CILT did not utilize the visual barrier. Reminders to constrain alterative modes of communication were given only verbally to individuals in this treatment. It has been found that the barrier provides no significant value to the treatment and it provides an unnatural context for constraint of verbal communication. This treatment includes positive social reinforcement of target behaviors, and while individuals receiving both treatments were able to see the primary researcher's face at all times throughout treatment, more naturalistic reinforcement can be given

between participants without the barrier (Szaflarski, et al., 2008). Participants in the g-CILT condition were given a small cardholder to assist in managing their materials effectively, due in part to difficulty with hemi-plegia (Figure 4).

2.5.3.2 Material difficulty constraint. The cards were designed to vary on only one feature to effectively employ the material difficulty constraint. For example, there would be a pair of cards featuring a boy holding two red apples and another pair of cards featuring a boy holding two blue apples in the same deck. In this way, the participant was required to accurately describe the desired card in order to obtain a match.

2.5.3.3 Complexity rule constraint. Participants were asked to increase the complexity of their utterances as they moved throughout the games. Individuals in both o-CILT and g-CILT conditions advanced through the same hierarchy of difficulty, but at their own pace (e.g., one participant in the dyad could be working at level 2 while the other participant operated at level 6). The hierarchy included six levels, and participants moved up the hierarchy of difficulty after successfully completing two consecutive turns as the requester using the appropriate verbal output for their level without utilizing cueing from the primary researcher. This hierarchy can be seen in Table 4. All participants began at Level 1 at the start of each day, and progressed through the levels throughout the session. The primary researcher recorded data on the number of attempts needed to produce a successful utterance. This enabled the primary researcher to accurately move the participants through the hierarchy when they had completed two successful and independent turns, and it also allowed for the establishment of effective cueing hierarchies for individual participants.

Figure 3. o-CILT participants utilizing a visual barrier.

Figure 4. g-CILT participants play the dual card game with a card holder.

2.5.3.4 Grammaticality constraint and grammaticality judgments. A grammaticality constraint was employed throughout the entire g-CILT condition only. This constraint required participants to only produce grammatically-correct sentences. Participants in the g-CILT condition were provided with a temporal adverb (e.g., yesterday, tomorrow, right now, everyday) at the beginning of their turn, and asked to formulate their sentence using this time word (e.g., *Yesterday, the boy held the apple*). During each turn, the temporal adverbs were randomly assigned ensuring that participants had an equal amount of practice with each adverb.

In addition to producing only grammatically correct sentences, participants receiving g-CILT were asked to make grammaticality judgments. This was a metalinguistic task, which is very different from the grammaticality production constraint. When the g-CILT requester produced a sentence, the g-CILT responder was encouraged to judge the grammaticality of the sentence produced. Grammatical judgment accuracy was not recorded throughout treatment, as it was primarily included to engage the responder throughout the grammatically-focused treatment.

2.5.4 Cueing

Cues were provided to participants in both treatments when deemed necessary, and participants were provided with the minimal amount of cueing necessary to end each turn successfully (defined as producing a sentence which met the constraints listed in Table 4). If the initial response of the participant did not meet the accuracy criteria, participants were asked to independently provide another response (e.g., “Close, but not quite. Try again!”). If the participant had difficulty producing an intended word, a phonemic or semantic cue was provided (e.g., “The word begins with a ‘duh’ sound” or

"This animal barks."). If this cueing level was not sufficient to produce the target response, the participant was given a choice of two possible responses and asked to repeat the response which was most like the target (e.g., "Which one sounds right? The boy is holding a dog. The boy is holding a brown dog."). If the participant required maximal cueing, they were given the response and asked to repeat it (e.g., "Say this: "The boy is holding a brown dog").

During the first three sessions for each treatment, cueing was provided from a constraint reminder board and sample verb lists. Each participant was given a constraint reminder board which described the constraint required for their level of play (e.g., level 1: person+action+thing). Participants in both treatments referred briefly to these cues during the first session; but, presumably because of the intensive and repetitive nature of the treatment, participants did not find it necessary to utilize these aids as treatment continued. Similar to the constraint reminder board, participants were given a sample list of ten verbs which could be used to assist in sentence formulation. While all participants referred to this board during the first two days of treatment, they were able to generate sentences using provided and novel verbs without the aid of the list in subsequent sessions.

To reiterate, the treatment hours, intervention task, stimuli, cueing hierarchy, verbal output constraint, material difficulty constraint, and complexity rule constraint were the same for individuals receiving o-CILT and g-CILT. The difference between the two treatments was the addition of a grammaticality constraint and grammaticality judgment for individuals receiving g-CILT.

Table 4.

Description of complexity of utterance constraint

Level	o-CILT minimum accuracy scoring criteria	g-CILT minimum accuracy scoring criteria
1	Agent+action+object "Boy hold apple."	Temporal adverb+agent+well formed action+object "Tomorrow the boy will hold the apple."
2	Level 1+ request "Boy hold apple. Card?"	Level 1 + request "Tomorrow the boy will hold the apple. Can I have that card?"
3	Level 2 + politeness request "Boy hold apple. Bob, Card, please?"	Level 2 + politeness request "Tomorrow the boy will hold the apple. Bob, can I have that card, please?"
4	Level 3+ color descriptor "Boy hold red apple. Bob, card, please?"	Level 3+ color descriptor "Tomorrow the boy will hold the red apple. Bob, can I have that card, please?"
5	Level 4+ article "Boy hold the red apple. Bob, card, please?"	Level 4+ article "Tomorrow the boy will hold a red apple. Bob, can I have that card, please?"
6	Level 5+ plural "Boy hold two red apples. Bob, card, please?"	Level 5+ plural "Tomorrow the boy will hold two red apples. Bob, can I have that card, please?"

2.6 Data Analysis

2.6.1 Scoring and Reliability

All tests were scored by the primary researcher and additionally by research assistants who were familiar with the scoring procedures. The reliability scoring occurred either during the testing session or while watching videotapes of the session, and research assistants were blind to the treatment condition to ensure unbiased reliability for testing scores. The Cinderella story retelling, BDAE narrative discourse subtest, cartoon video description, and conversational samples were transcribed from audio and video recordings and then coded using SALT by the primary researcher. For reliability purposes, twenty percent of all discourse samples were then transcribed and coded again by a research assistant proficient with the coding strategy, but blind to condition and time of sample (e.g., baseline, post treatment, maintenance).

Reliability differences were addressed during both the transcription and coding processes. Initially, transcriptions from the primary researcher and reliability scorer demonstrated a discrepancy in the process for parsing sentences. This is a difficulty frequently encountered by researchers attempting to perform in-depth linguistic analyses (see Rochon, Saffran, Berndt, & Schwartz, 2000). Guidelines for establishing sentence boundaries were adapted from Miller & Chapman, 2000, and indicated that prosodic and contextual parsing cues were essential to accurately identifying sentence boundaries. After discussing these guidelines, major transcription differences were resolved. All discourse samples were then re-transcribed following these guidelines. Revised transcriptions were then recoded and a reliability scorer coded a different 20% of the revised discourse samples. Inter-rater reliability for the recoding was 95%. All discourse

materials were then analyzed for the linguistic measures described in Table 2 and statistical analyses were performed on the primary researcher's revised coding.

2.5.2 Statistical Analysis

Due to the single subject design and the inter-individual variation on severity and grammaticality measures, non-parametric statistics were used for analysis of the data. Each participant's test scores between any two time points (baseline, post treatment, and maintenance) were compared with McNemar's change test. Effect sizes for all tests were also calculated by dividing the change score (post-treatment minus baseline, maintenance minus baseline) by the pre-treatment variability of the group for that test. Effect sizes are becoming standard in aphasiology literature, and the values of Cohen's *d* scores, which are common to psychology literature, were used to determine the magnitude of the effect size for this calculation (e.g., small effect= .2, medium effect=.5, large effect=.8; Cohen, 1988). Results found in this study were also compared and analyzed with the measures used by Maher, et al. (2006). Maher, et al. used the following criteria to evaluate their data: WAB scores were considered significant according to the Shewan and Donner (1988) criteria of a 5 point difference in WAB AQ (Kertesz, 2007) scores and a change of 2 standard deviations, or 8 points, on the BNT (Kaplan, et al., 2000) determined significant change on this measure. Since Maher et al. did not use the McNemar test and did not compute effect sizes, those Maher, et al. (2006) raw data were analyzed for purposes of this thesis with McNemar's change test and effect sizes were calculated as described for this thesis.

Changes in discourse linguistic measures as analyzed by SALT were deemed significant using the criteria of a two standard deviation change between two scoring

points (e.g., post-treatment minus baseline, maintenance minus baseline). This was calculated by determining the pretreatment variability, or standard deviation, of all baseline performance scores for each test with individuals in this thesis. Because individuals with aphasia can show significant variation on measures of linguistic ability, the standard deviations used to calculate these measures were typically very large. In order to be considered significant by these standards, the change scores had to be considerable, thus ensuring that these changes were truly significant.

2.7 Comparisons of this thesis with Maher, et al. (2006)

While an effort was made to keep the design of this thesis as close to the Maher, et al. (2006) guidelines as possible, this study differed slightly in a few respects, and it is necessary to review these variations. The assessment test for verb naming was different, as the Object and Action Naming Battery (Druks & Masterson, 2000) replaced the Action Naming Test (Nicholas, Obler, Albert, & Goodglass, 1985) since the latter test was not commercially available. Additional tests of grammaticality were included to assess the effect of CILT on grammatical production and this thesis included a larger corpus of discourse samples to examine CILT's effect on measures of grammaticality in discourse. Stimuli used in this thesis differed slightly with the inclusion of an agent (e.g., boy or girl) on the stimulus card holding an object, as opposed to a card featuring only the object. The grammaticality constraint and judgment were not present in the Maher, et al. (2006) work. Statistical analyses procedures also varied between studies. Test measures were not analyzed by the Maher, et al. (2006) criteria as described in section 2.5.2, and McNemar's change test and effect sizes were calculated with Maher, et al.'s data. Sentences in discourse grammaticality measures were coded as grammatical or

ungrammatical and verb tense was identified and coded, while these measures were not present in the Maher, et al. (2006) study.

CHAPTER 3: RESULTS

3.1. Research Question #1: Determining the efficacy of CILT for individuals with agrammatic aphasia

The results are described in the order of the research questions. The first set of results compares the performance of individuals with agrammatic aphasia in this thesis and Maher, et al.'s participants on measures of aphasia severity and the second half of the results examines the response of individuals with agrammatism to g-CILT and o-CILT.

3.1.1 Western Aphasia Battery-Revised Aphasia Quotient

The Aphasia Quotient (AQ) composite scores for the Western Aphasia Battery-R (Kertesz, 2007) are given along with the change scores and effect sizes for individuals receiving o-CILT (Figure 5) and Maher, et al. (2006) study participants (Figure 6) in Table 5. The mean baseline AQ for o-CILT participants was 81.7 (scores: 69.7-93.7) while the mean baseline was 58.5 for Maher et al.'s participants (scores 40.6-70). Hence, there was a 23.2 point difference in severity between the participants of the two studies. Post-treatment scores improved by an average of 2.1 points for o-CILT (83.8; scores: 74.8-92.8) while the improvement was nearly 6.6 points for Maher, et al. (65.08, scores: 51.2 to 74.1). Maintenance scores remained relatively similar to pre-treatment levels for o-CILT participants (81.5, scores: 70.7-92.3), while maintenance scores improved by nearly 9 points for Maher, et al.'s participants (67.48, scores: 54.2-77).

To summarize, while three out of four participants in Maher, et al. improved in post-treatment scores and maintained these improvements, one o-CILT participant (P1) improved in post-treatment testing, but did not maintain the gains. The second o-CILT participant was in the top scoring tier for the WAB AQ, and it is possible that this ceiling

effect impacted performance. In addition to deficit specificity, the issue of aphasia severity seems crucial in determining the potential for change with CILT, and this aspect will be addressed in the section 3.2.1 (results of g-CILT) and in the Discussion chapter.

3.1.2 Boston Naming Test

Boston Naming Test (Kaplan, et al. 2000) scores varied between this study's participants receiving o-CILT (Figure 7), as well as in comparison to CILT participants in the Maher, et al. study (Figure 8). These results can be seen in Table 6. The mean baseline BNT score for o-CILT participants was 51.5 (scores: 47-56) while the mean baseline was 18 for Maher et al.'s participants (scores 4-38). There was a 33 point difference in severity between the participants in the two studies. Post-treatment scores declined by an average of 4.5 points for o-CILT (47; scores: 36-58) while Maher, et al. participants gained an average of 3 points (21, scores: 4 to 42). Maintenance scores remained relatively similar to pre-treatment levels for o-CILT participants (52, scores: 46-58), while maintenance scores improved by nearly 6 points for Maher, et al.'s participants (26.75, scores: 9-48).

To summarize, two of four participants in Maher, et al.'s study improved in post-treatment scores and maintained these improvements, and no o-CILT participants demonstrated or maintained significant post treatment gains. P2 was close to a perfect score on the BNT, and it is possible that this ceiling effect impacted performance. It is notable that while participants in the Maher, et al. study produced overall lower scores on the BNT than participants in this thesis, both sets of data showed at least one

Figure 5. Western Aphasia Battery-R Aphasia Severity Quotient (Kertesz, 2007) raw score at each assessment time for participants in this study (max score= 100).

Note: The g-CILT results will be discussed in section 3.2.1.

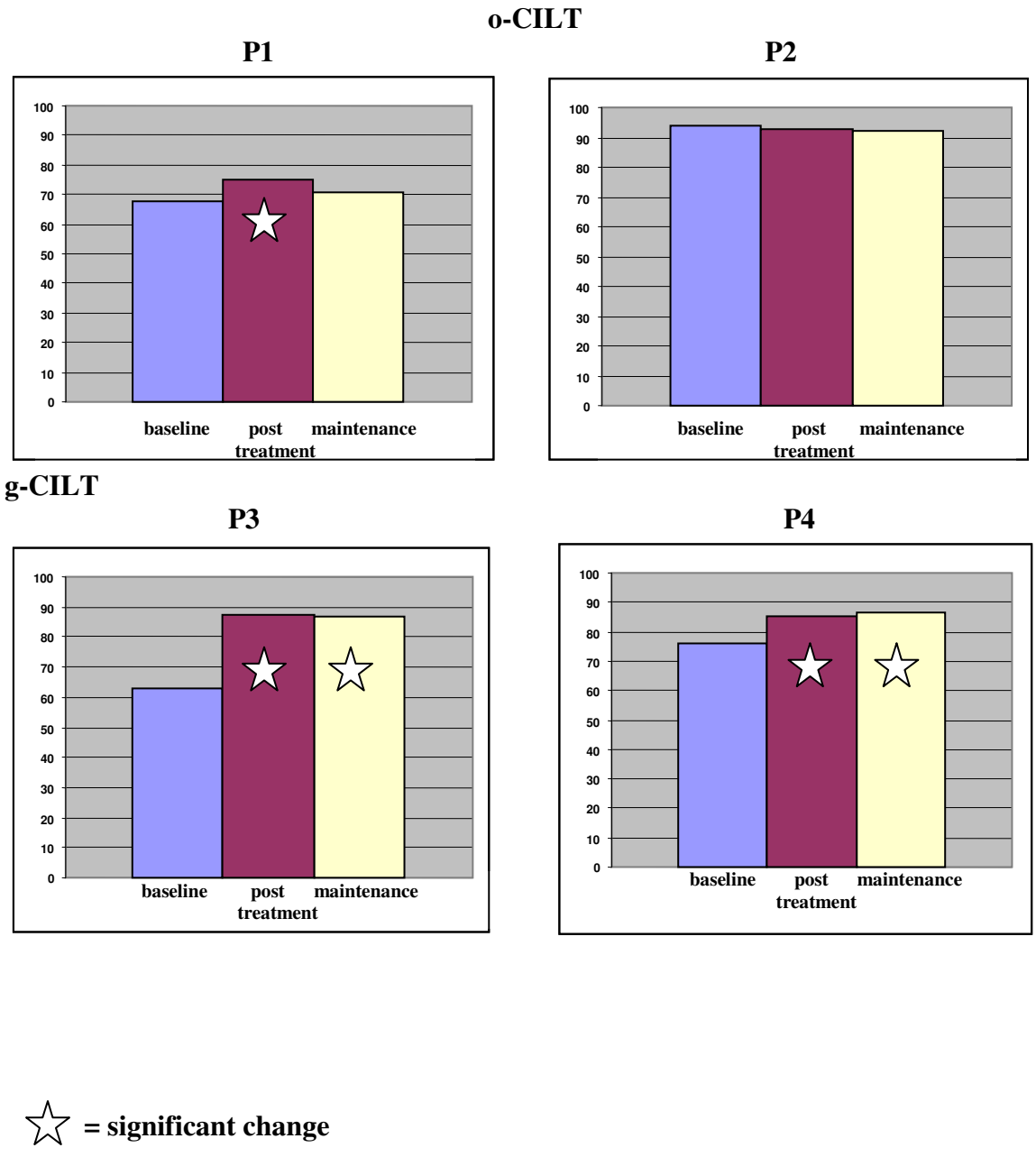


Figure 6. Western Aphasia Battery-R Aphasia Severity Quotient (Kertesz, 2007) raw score at each assessment time for participants in the Maher, et al. study (max score= 100).

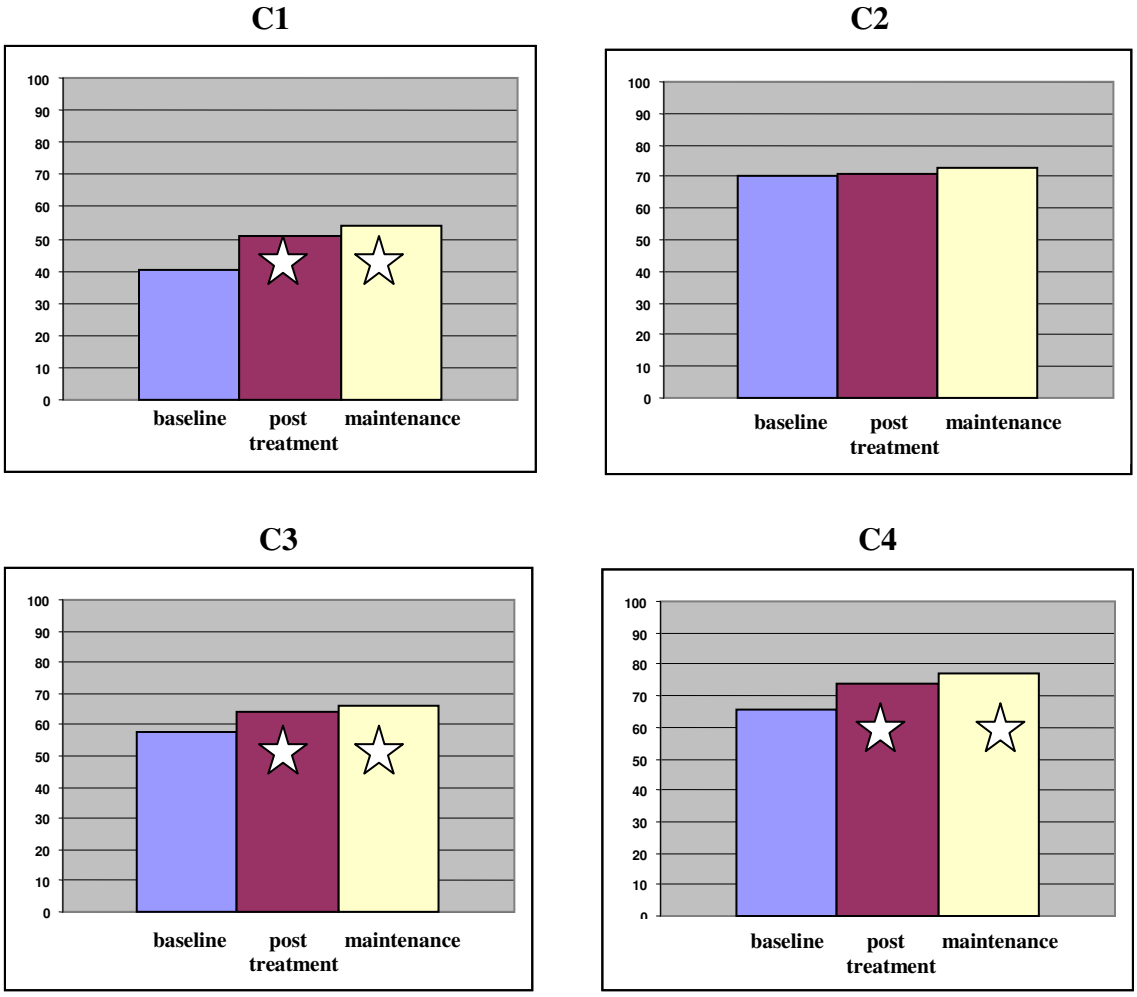


Table 5. *Actual scores, change scores, and effect sizes for participants on WAB-R AQ scores*

Treatment	Participant	Pre-Tx	Post-Tx	Post – Pre change (d)	Follow up	Follow up –pre change (d)
o-CILT	P1	69.7	74.8	***5.1**** (.39)^	70.7	1.0
	P2	93.7	92.8	-.9	92.3	-1.4
g-CILT	P3	62.9	87.3	24.4****(1.85)^^^	86.8+	23.9****(1.81)^^^
	P4	75.8	85	9.2***(.70)^^	86.5+	10.7***(.81)^^^
Maher, et al. (2006)	C1	40.6	51.2	10.6***(.47)^	54.2+	13.6****(.60)^^
	C2	70.0	71.2	1.2	72.8	2.8
	C3	57.8	63.8	6.0**(.27)^	65.9+	8.1***(.36)^
	C4	65.8	74.1	8.3***(.37)^	77.0+	11.2****(.50)^
	Mean	58.55	65.08		67.48	

Maximum score=100

Note: the text for g-CILT results is presented in section 3.2.1.
Effect sizes not listed were calculated to be below .20

Legend:
Bold- significant according to McNemar's change test
Italics- significant according to Maher, et al. (2006) criteria
* significant at $p \leq .05$ ** significant at $p \leq .02$
*** significant at $p \leq .01$ **** significant at $p \leq .001$
+ gains preserved between post treatment and maintenance testing
Magnitude of Effect Sizes (as per Cohen, 1988): ^ small effect size (d=.2);
^^medium effect size (d=.5); ^^^large effect size (d=.8)

Figure 7. Boston Naming test (Kaplan, et al., 2000) raw score at each assessment time for participants in this study (max score= 60).

Note: The g-CILT results will be discussed in section 3.2.2.

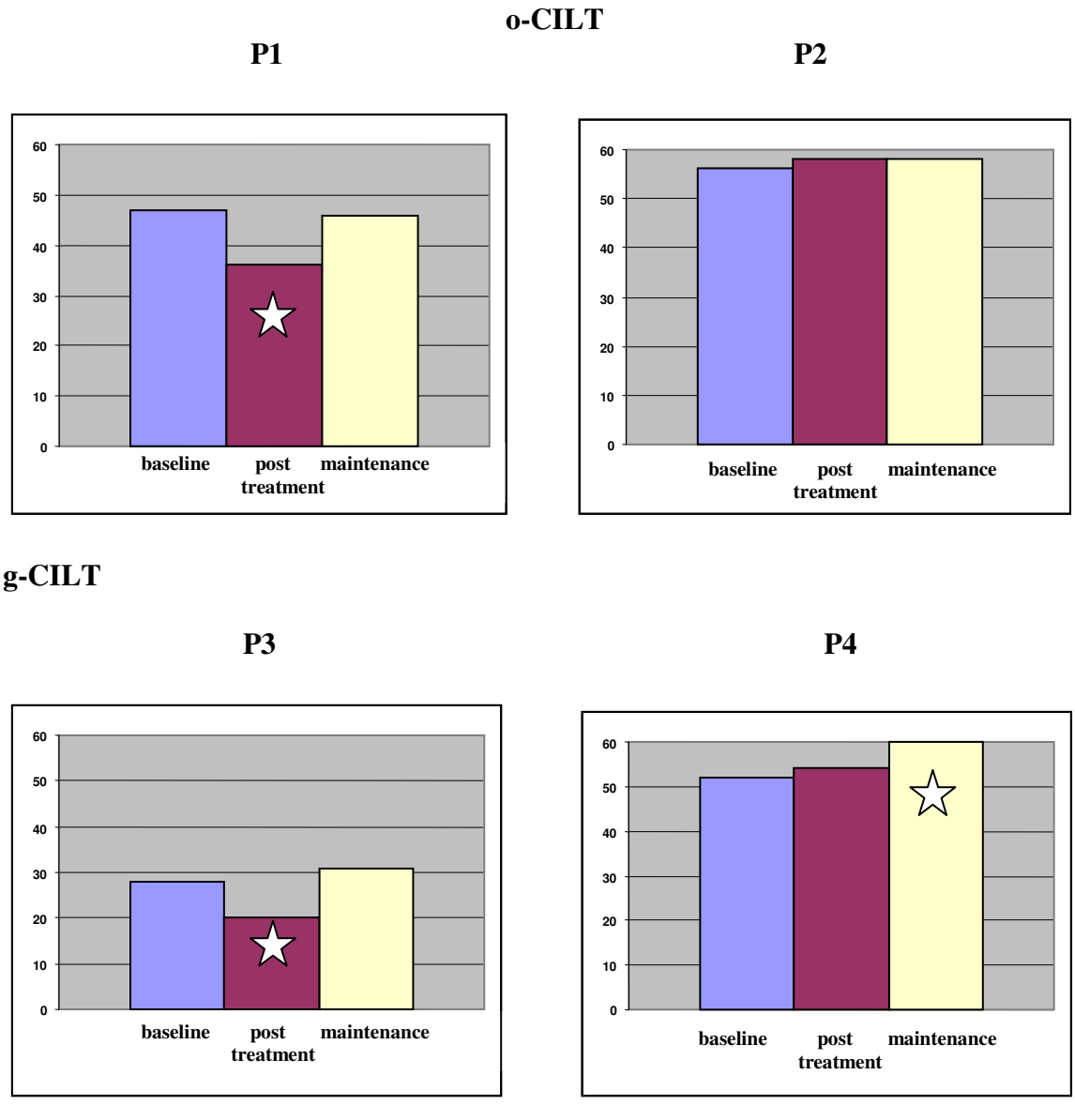


Figure 8. Boston Naming test (Kaplan, et al., 2000) raw score at each assessment time for participants in the Maher, et al. (max score= 60).

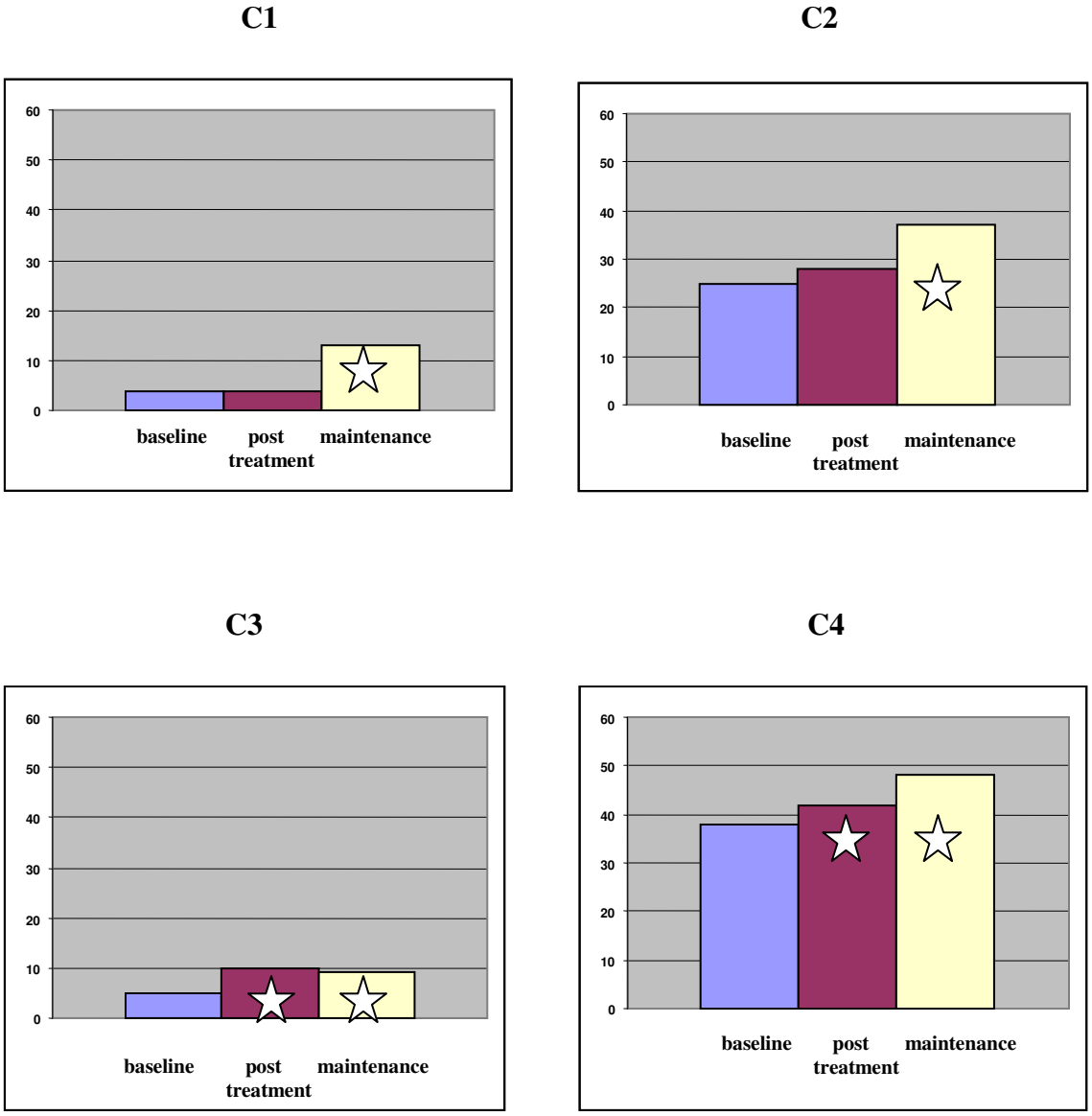


Table 6

Actual scores, change scores, and effect sizes for participants on BNT scores

Treatment	Participant	Pre-Tx	Post-Tx	Post – Pre change (d)	Follow up	Follow up –pre change (d)
o-CILT	P1	47	36	**-11****** (.89)^^^	46	-1
	P2	56	58	2	58	2
g-CILT	P3	28	20	**-8*****(-.65)^^	31	3(.24)^
	P4	52	54	2	60	**8*****(.65)^^
Maher, et al. (2006)	C1	4	4	0	13	**9*****(.56)^^
	C2	25	28	3	37	**12******(.73)^^
	C3	5	10	**5***(.30)^	9+	**4***(.24)^
	C4	38	42	**4***(.24)^	48+	**10*****(.61)^^
	Mean	18	21		26.75	

Maximum score=60

Note: the text for g-CILT results is presented in section 3.2.2.
Effect sizes not listed were calculated to be below .20

Legend:
Bold- significant according to McNemar's change test
Italics- significant according to Maher, et al. (2006) criteria
* significant at $p \leq .05$ ** significant at $p \leq .02$
*** significant at $p \leq .01$ **** significant at $p \leq .001$
+ gains preserved between post treatment and maintenance testing
Magnitude of Effect Sizes (as per Cohen, 1988): ^ small effect size (d=.2);
^^medium effect size (d=.5); ^^^large effect size (d=.8)

participant demonstrating significant increases in scoring at post-treatment and maintenance testing measures. Potential explanations for this are explored in the Discussion chapter.

3.1.3 Object and Action Naming Battery

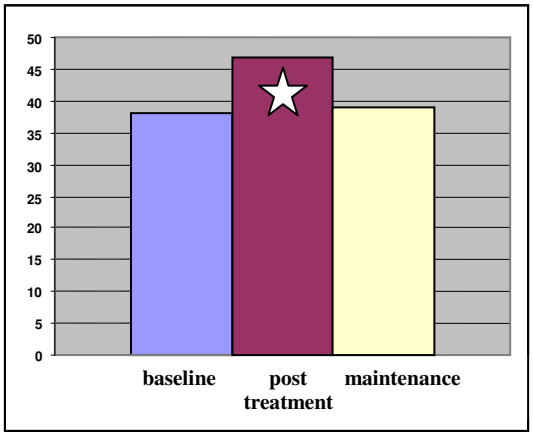
The Object and Action Naming Battery (Druks & Masterson, 2000) results for individuals receiving o-CILT (Figure 9) and scores for the Action Naming Test (Nicholas, et al., 1985) (Figure 10) can be seen in Table 7. The mean baseline O&A score for o-CILT participants was 38.75 (scores: 34-46) while the mean baseline was 18 for Maher et al.'s participants (scores 3-38). There was a 20.75 point difference in average severity between the participants in the two studies. Post-treatment scores improved by an average of 2 point for o-CILT (40.5; scores: 34-47) while Maher, et al. participants gained an average of 5 points (23.25, scores: 6 to 44). Maintenance scores increased 6.5 points for o-CILT participants (45.25, scores: 39-49), while maintenance scores improved by nearly 3 points for Maher, et al.'s participants (26 scores: 5-43).

To summarize, two of four participants in Maher, et al. improved in post-treatment scores and maintained these improvements, and one o-CILT participant demonstrated and maintained significant post treatment gains. It is notable that both P1 and P2's scores increased significantly at post treatment testing, but P1's scores decreased significantly to baseline while P2's scores significantly increased at post treatment and maintenance testing. This suggests that participant-specific factors may influence response to CILT, and this is further addressed in the Discussion. Additionally, the maximum score for the Action Naming Test used in the Maher, et al. study was

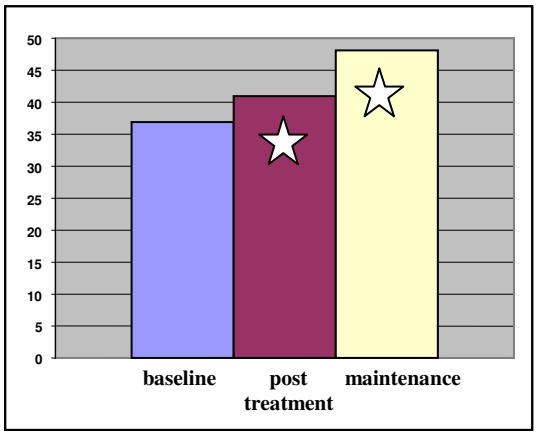
Figure 9. Object and Action Naming Battery (Druks & Masterson, 2000) raw score changes at each assessment time for participants in this study (maximum score=50).

o-CILT

P1

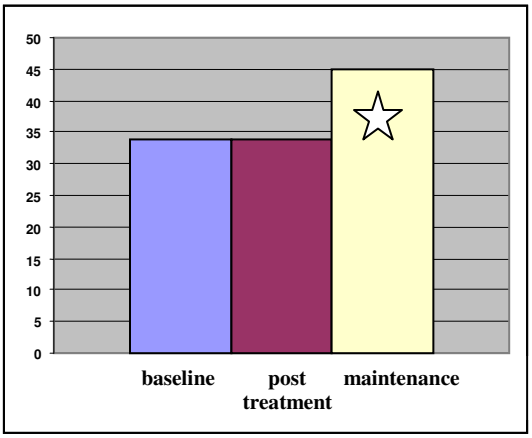


P2

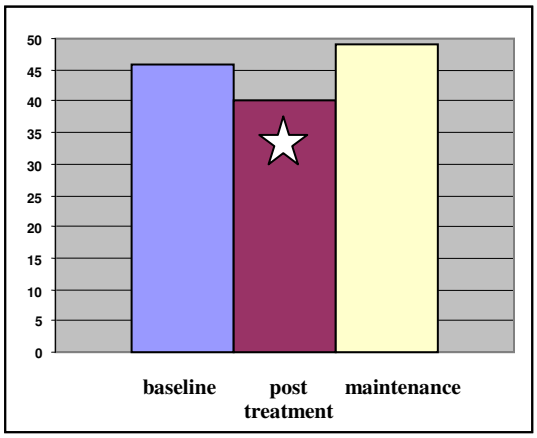


g-CILT

P3



P4



☆ **= significant change**

Figure 10. Action Naming Test (Nicholas, et al., 1985) raw score changes at each assessment time for participants in the Maher, et al. study (maximum score=unknown).

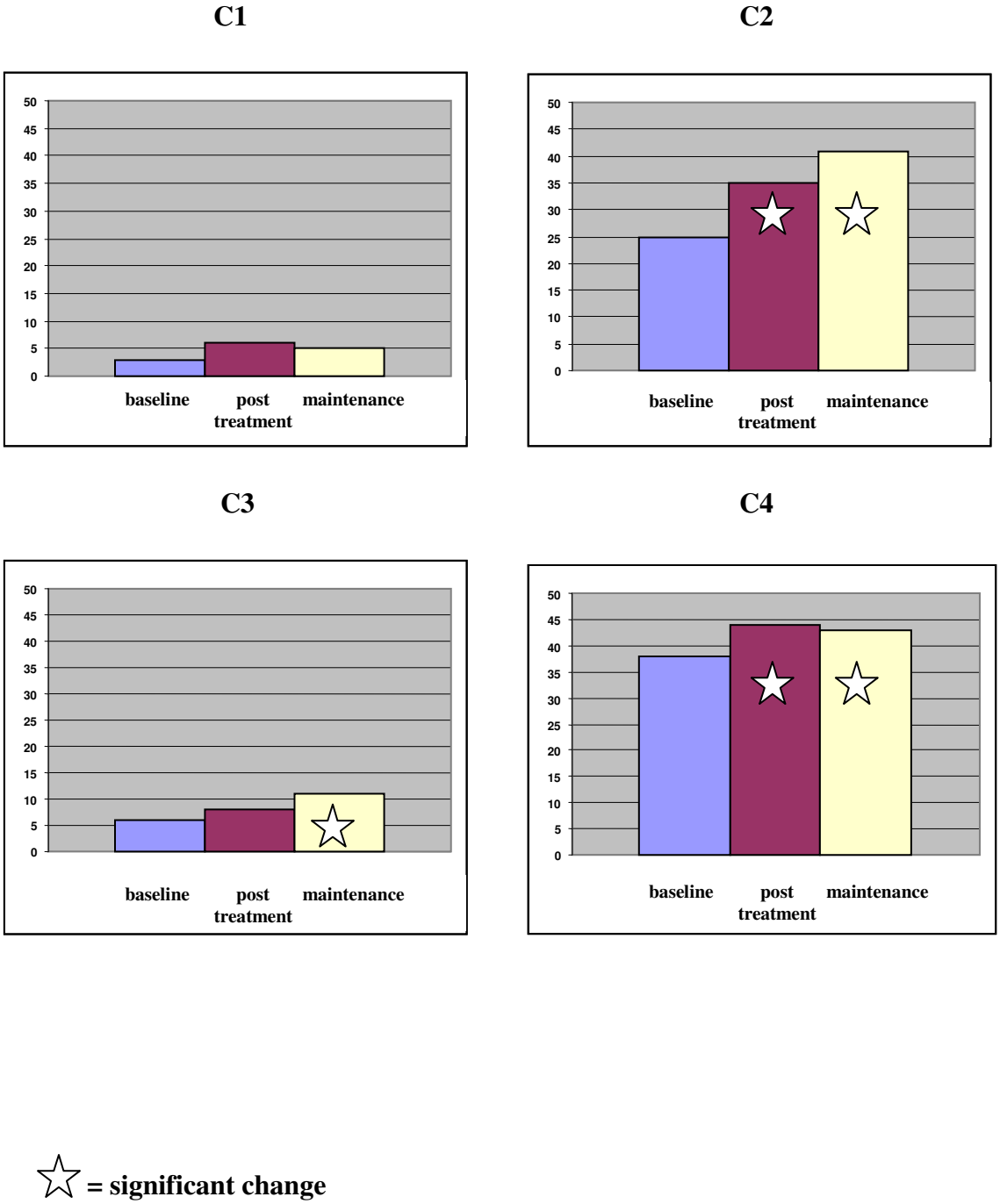


Table 7

Actual scores, change scores, and effect sizes for participants on O&A and ANT scores

Treatment	Participant	Pre-Tx	Post-Tx	Post – Pre change (d)	Follow up	Follow up –pre change (d)
o-CILT	P1	38	47	**9*****(1.78)^^^	39	1(.20)^
(O&A)	P2	37	41	**4***(.78)^^	48+	**11******(2.15)^^^
g-CILT	P3	34	34	0	45	**11******(2.15)^^^
(O&A)	P4	46	40	**-6****(1.17)^^^	49	3(.59)^^
	C1	3	6	3	5	2
Maher, et	C2	25	35	**10*****(.61)^^	41	**16******(.97)^^^
al. (2006)	C3	6	8	2	11	**5***(.30)^
(ANT)	C4	38	44	**6***(.36)^	43	**5***(.30)^
	Mean	18	23.25		26	

Maximum score for O&A =50 Maximum score for ANT= unknown

Note: the text for g-CILT results is presented in section 3.2.3.
Effect sizes not listed were calculated to be below .20

Legend:
Bold- significant according to McNemar's change test
Italics- significant according to Maher, et al. (2006) criteria
* significant at $p \leq .05$ ** significant at $p \leq .02$
*** significant at $p \leq .01$ **** significant at $p \leq .001$
+ gains preserved between post treatment and maintenance testing
Magnitude of Effect Sizes (as per Cohen, 1988): ^ small effect size (d=.2);
^^medium effect size (d=.5); ^^^large effect size (d=.8)

unavailable, so it is not possible to determine how well Maher, et al. participants performed on this measure.

3.1.4 Summary of results for research question #1

Overall, Maher, et al.'s participants demonstrated lower scores on measures of severity, indicating higher severity, than individuals receiving o-CILT in this study. While it is important to take into consideration the dramatic differences in average scores, participants in this thesis produced different scoring trends than Maher, et al.'s CILT participants. Participants in this study showed significant changes from baseline to post-treatment testing on three of six possible measures (50%) and CILT participants in the Maher, et al. study showed significant baseline to post-treatment changes on seven of twelve baseline measures (58%).

Positive maintenance changes were evidenced by participants receiving o-CILT in two of six opportunities (33%), while Maher, et al.'s CILT participants demonstrated significant positive maintenance changes in eleven of twelve opportunities (92%). While the differences between these participants appear significant, small sample size, high baseline variability of individual test scores, inability to directly compare and calculate significance for verb naming measures, and ceiling effects may be impacting this data. These possibilities are further discussed in the Discussion section.

3.2 Research question #2: Determining the efficiency of a grammatical modification to CILT for individuals with agrammatic aphasia

In addition to the severity tests described in the previous section, additional measures of grammaticality were used to compare g-CILT and o-CILT: Verb Inflection

Test (Faroqi-Shah, unpublished), Spontaneous Speech Score- Fluency, Grammatical Competence, and Paraphasia score from the Western Aphasia Battery (Kertesz, 2007), and the following grammaticality measures in discourse samples: proportion of sentences, proportion of grammatical sentences, accuracy of tense, and variety of tense.

Aphasia Severity Measures

3.2.1 Western Aphasia Battery-Revised Aphasia Quotient

The Aphasia Quotient (AQ) composite scores for the Western Aphasia Battery-R (Kertesz, 2007) are given along with the change scores and effect sizes for individuals receiving o-CILT and g-CILT in Figure 5 and Table 5 (on pages 57 and 59, respectively). The mean baseline AQ for o-CILT participants was 81.7 (scores: 69.7-93.7) while the mean baseline was 69.35 for g-CILT participants (scores 62.9-75.8). Hence there was a 12.35 point difference in severity between the participants of the two conditions. While participants were matched between the two treatment groups, matching was conducted on the basis of recency since individual treatment and VIT score, not necessarily aphasia severity. Post-treatment scores improved by an average of 2.1 points for o-CILT (83.8; scores: 74.8-92.8), while the improvement was nearly 17 points for g-CILT participants (86.15, scores: 85 to 87.3). Maintenance scores remained relatively similar to pre-treatment levels for o-CILT participants (81.5, scores: 70.7-92.3), while maintenance scores remained stable from post treatment gains for individuals receiving g-CILT (86.65, scores: 86.5-86.8).

To summarize, while one participant receiving o-CILT improved in post-treatment testing, this was not maintained, and participants receiving g-CILT evidenced

significant post-treatment gains which were maintained over time. As previously noted, P2 may have been exhibiting a ceiling effect on this task.

3.2.2 Boston Naming Test (BNT)

The BNT (Kaplan, et al. 2000) scores varied between study participants, and these results can be seen in Figure 7 and Table 6 (on pages 60 and 62, respectively). The mean baseline BNT score for o-CILT participants was 51.5 (scores: 47-56), while the mean baseline was 40 for g-CILT participants (scores 28-52). There was an 11.5 point difference in severity between the participants in the two groups. Post-treatment scores declined by an average of 4.5 points for o-CILT (47; scores: 36-58), while g-CILT participants declined by an average of 3 points (37, scores: 20 to 54). Maintenance scores remained relatively similar to pre-treatment levels for o-CILT participants (52, scores: 46-58), while maintenance scores improved by nearly 8.5 points for g-CILT participants (45.5, scores: 9-48).

To summarize, no participants receiving g-CILT or o-CILT demonstrated or maintained significant post treatment gains. One pair of matched participants (P1 and P3) both demonstrated significant downward trends immediately following treatment and both participants reverted back reach their original scores at maintenance testing. P4 also significantly improved at post-treatment testing, and possible explanations of the gains made by all of these participants are discussed in the next chapter.

3.2.3 Object and Action Naming Battery

The Object and Action Naming Battery (Druks & Masterson, 2000) results for individuals receiving o-CILT and g-CILT can be seen in Figure 9 and Table 7 (on pages 64 and 66, respectively). The mean baseline O&A score for o-CILT participants was 42

(scores: 39-45), while the mean baseline was 43.5 for g-CILT participants (scores 40-47). There severities were essentially evenly matched between participants in the two conditions. When averaged, post-treatment scores were essentially the same as baseline scores for o-CILT (43; scores: 41-45) and g-CILT (43, scores: 38 to 48). Maintenance scores remained relatively similar to pre-treatment levels for o-CILT participants (43.5, scores: 39-48), while maintenance scores improved by 4 points for g-CILT participants (47 scores: 45-49).

To summarize, overall, participants did not make significant gains from baseline to post-treatment testing, but on an individual basis, P2 in the o-CILT group did produce and maintain significant gains on this measure. As on the BNT, the two matched participants with lower VIT scores (P1 and P3) demonstrated a decrease in scores immediately following treatment. Factors which may underlie these changes are further explored in the discussion. Results of P4 may have been impacted by a ceiling effect

Measures of Grammaticality

3.2.4 Verb Inflection Test (VIT)

Only participants who received the g-CILT treatment demonstrated significant change on this measure, and this can be seen in Figure 11 and Table 8. The mean baseline VIT score for o-CILT participants was 5.5 (scores: 4-7), while the mean baseline was 7.5 for g-CILT participants (scores 7-8). There was a two point difference in severities between participants in the two conditions. Post-treatment scores averages increased by a point for o-CILT (6.5; scores: 5-8), and g-CILT scores increased by 5.5 points (13, scores: 12 to 14). Maintenance scores remained at baseline for o-CILT

participants (6, scores: 5-7), while maintenance scores declined by 2 points for g-CILT participants (11 scores: 10-12).

To summarize, individuals who utilized the grammaticality constraint during CILT (g-CILT participants) showed significant increases and some maintenance in correct verb morphology usage as defined by the VIT, and individuals in the o-CILT group did not demonstrate significant changes in performance on this measure. This is a crucial finding of this thesis and will be revisited in the Discussion chapter.

3.2.5 WAB Spontaneous Speech- Fluency, Grammatical Competence, and Paraphasias

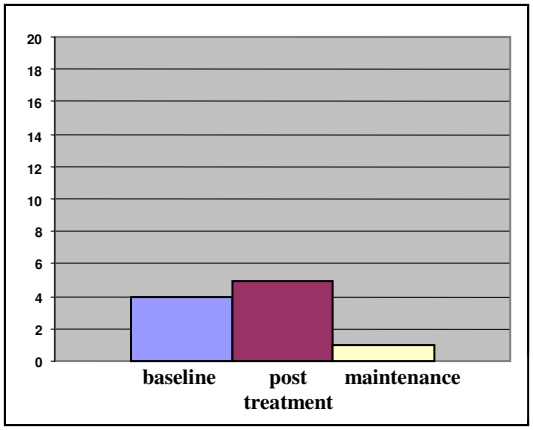
Only participants who received g-CILT demonstrated significant changes on this measure (see Figure 12 and Table 9). The mean baseline VIT score for o-CILT participants was 7 (scores: 5-9), while the mean baseline was 5 for g-CILT participants (scores 5-5). There was a two point difference in severities between participants in the two studies. Post-treatment scores averages increased by 1 point for o-CILT (8.25; scores: 6.5-10) and g-CILT scores increased by almost 4 points (8.75, scores: 8.5 to 9). Maintenance scores remained at essentially at baseline for o-CILT participants (7.5, scores: 6.5-9), while maintenance scores remained the same for g-CILT participants (9, scores: 9-9).

To summarize, individuals who received g-CILT produced significant and consistent improvements on the grammatical competence score of the WAB-R spontaneous speech subtest, while individuals who received o-CILT did not demonstrate significant gains on this subtest. This is another crucial finding of this thesis and will be addressed in the Discussion chapter.

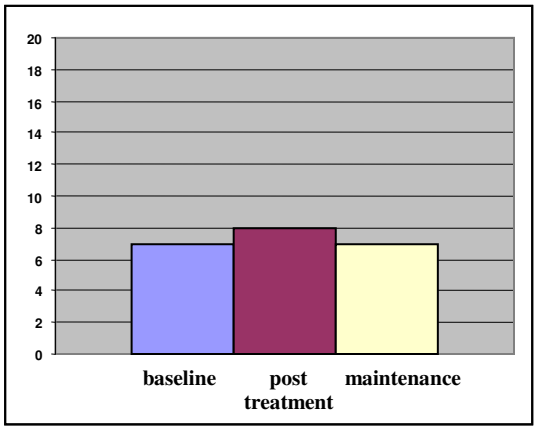
Figure 11. Verb Inflection Test (Faroqi-Shah, unpublished) raw score changes at each assessment point (maximum score=20).

o-CILT

P1

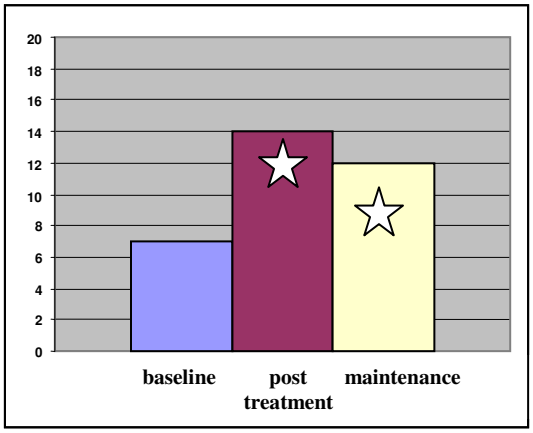


P2

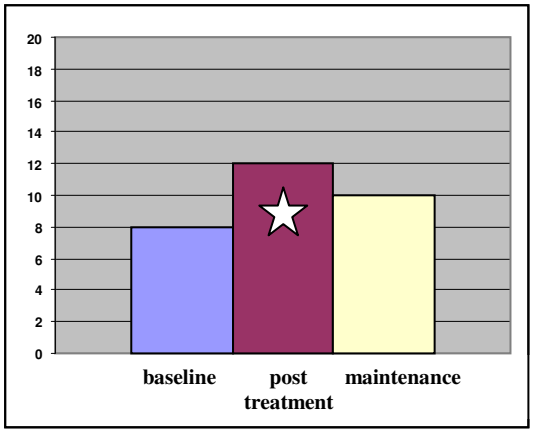


g-CILT

P3



P4



☆ **= significant change**

Table 8

Actual scores, change scores, and effect sizes for participants on VIT scores (maximum score=20)

Treatment	Participant	Pre-Tx	Post-Tx	Post – Pre change (d)	Follow up	Follow up –pre change (d)
o-CILT	P1	4	5	1(.58)^^	5	1(.58)^^
	P2	7	8	1(.58)^^	7	0
g-CILT	P3	7	14	**7*****(4.04)^^^	12+	**5***(2.89)^^^
	P4	8	12	**4***(2.31)^^^	10+	2(1.16)^^^

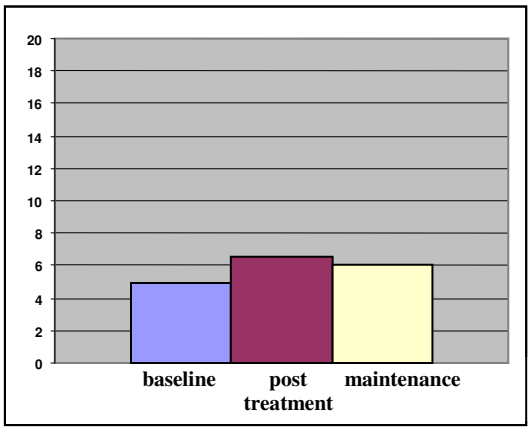
Note: Effect sizes not listed were calculated to be below .20

Legend:
Bold- significant according to McNemar's change test
Italics- significant according to Maher, et al. (2006) criteria
* significant at $p \leq .05$ ** significant at $p \leq .02$
*** significant at $p \leq .01$ **** significant at $p \leq .001$
+ gains preserved between post treatment and maintenance testing
Magnitude of Effect Sizes (as per Cohen, 1988): ^ small effect size (d=.2);
^^medium effect size (d=.5); ^^^large effect size (d=.8)

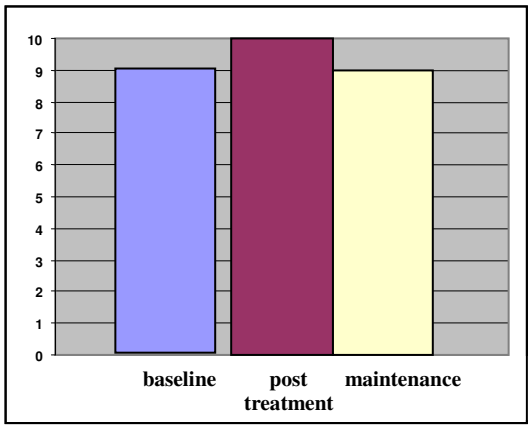
Figure 12. Western Aphasia Battery Spontaneous Speech Fluency, Grammatical Competence, and Paraphasia Score (Kertesz, 2007) raw score changes at each assessment point (maximum score= 10).

o-CILT

P1

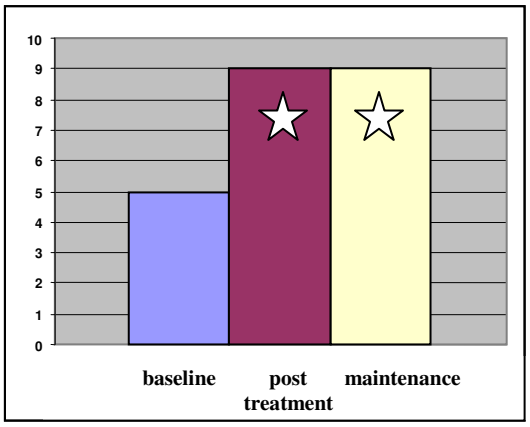


P2

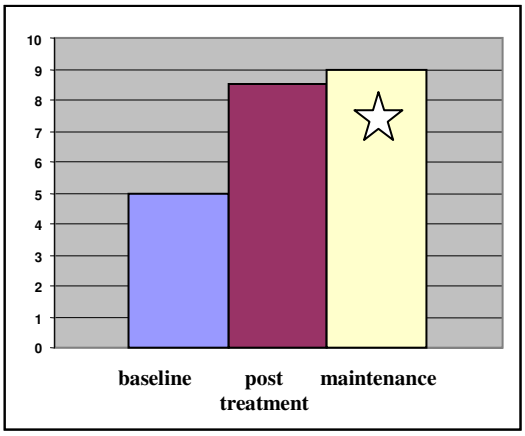


g-CILT

P3



P4



☆ **= significant change**

Table 9

Actual scores, change scores, and effect sizes for participants on WAB spontaneous speech grammatical content scores (maximum score=10)

Treatment	Participant	Pre-Tx	Post-Tx	Post – Pre Change (d)	Follow up	Follow up –pre change (d)
o-CILT	P1	5	6.5	1.5 (.75)^^	6	1(.5)^^
	P2	9	10	1(.5)^^	9	0
g-CILT	P3	5	9	**4***(2)^^^	9+	**4***(2)^^^
	P4	5	8.5	3.5(1.75)^^^	9	**4***(2)^^^

Note: Effect sizes not listed were calculated to be below .20

Legend:
Bold- significant according to McNemar's change test
Italics- significant according to Maher, et al. (2006) criteria
* significant at $p \leq .05$ ** significant at $p \leq .02$
*** significant at $p \leq .01$ **** significant at $p \leq .001$
+ gains preserved between post treatment and maintenance testing
Magnitude of Effect Sizes (as per Cohen, 1988): ^ small effect size (d=.2);
^^medium effect size (d=.5); ^^^large effect size (d=.8)

3.2.5 Grammatical analyses of discourse measures As described in section 2.3.2, a variety of discourse samples were used to calculate several measures of grammatical well formedness. The results are given in Tables 10 and 11. No data are reported for the BDAE Narrative Discourse subtest because no measures reached significance criteria. No post-treatment scores are reported for P3 narrative measures due to experimental error. P4's scores may be impacted my mild apraxia and a reluctance to elaborate during narratives and engage in conversation. Overall, participants who received o-CILT produced more significant changes in discourse grammaticality measures than participants who received g-CILT.

P1, who received o-CILT, demonstrated an increase in proportion of well-formed sentences at maintenance testing during the Cinderella narrative retelling and a significant increase in accuracy of tense at post treatment testing which was maintained while describing the Tom and Jerry cartoon video. P2 produced significant gains in proportion of grammatical sentences which was maintained during the retelling of the Cinderella story, and no significant changes were produced on grammaticality measures for other narrative samples. Conversational data were significant for an increase in accuracy of tense at the maintenance testing session relative to baseline status. No improvements in tense accuracy in a conversational context were noted immediately after treatment with o-CILT.

P3 received g-CILT and significantly improved on the proportion of grammatical sentences used at maintenance testing during the retelling of the Cinderella story; however no improvement was noted immediately post treatment on this measure in narrative discourse. In conversational samples, P3 demonstrated an increase in tense

accuracy immediately post-treatment, and this gain in conversational accuracy was maintained. P4 showed no statistically significant gains in narrative or conversational grammaticality measures. It is worthwhile to note that P4 was extremely reluctant to have his speech recorded during testing sessions, and this may have impacted his performance. To summarize, it appears that participants who received o-CILT demonstrated qualitatively more significant increases on discourse grammaticality measures in comparison to individuals who received g-CILT.

3.2.6 Summary of results for research question # 2

To summarize, a variety of significant changes were observed in both aphasia severity and grammaticality test measures and on linguistic analyses of discourse for individuals receiving o-CILT and g-CILT. In terms of aphasia severity measures (WAB-R AQ, BNT, O&A), participants receiving o-CILT showed positive significant changes from baseline to post treatment testing on two of six possible measures (33%) and one of the significant change scores was maintained. g-CILT participants demonstrated positive significant changes from baseline to post treatment testing on two of six possible measures (33%), and both of these significant changes were maintained.

On measures of grammaticality (VIT, WAB grammatical competence score, discourse measures of grammaticality), individuals receiving o-CILT produced three significant change scores of a possible twenty-four change scores (12.5%), and two of these changes were maintained over time. Participants who received g-CILT demonstrated four significant positive change scores of a possible sixteen change scores (25%), and three of these changes were maintained. Upon closer examination of the data, overall, it appears that individuals receiving o-CILT demonstrated more significant

Table 10. Discourse sample grammatical analyses reported in proportions (proportion of sentences (S), proportion of well-formed sentences (WFS), tense accuracy (TA), and tense variety (TV) for individuals who received o-CILT.

Key: ***Bold and italicized*=** Significant by a criterion of a two standard deviation change score (note: Standard deviations calculated on the basis of o-CILT and g-CILT participant baseline data for each measure and sample)

ID	Discourse sample	Grammaticality measure (proportions)	Pre TX	Post TX	Pre-Post TX	Maintenance TX	Pre-Maintenance change
P1	Cinderella	S	.94	.84	-.10	.79	-.15
		WFS	.25	.24	.01	.37	***.12***
		TA	.75	.87	***.12***	.86	.11
		TV	.15	-.08	.07	.08	-.07
	Cartoon	S	1	.80	-.20	.92	-.08
		WFS	.29	.20	-.09	.13	-.16
		TA	.63	.93	***.30***	.83	***.20***
		TV	.08	.07	-.01	.17	.09
	Conversation	S	.93	.83	-.10	.74	.19
		WFS	.41	.39	.02	.26	.15
		TA	.79	.95	.16	.76	-.03
		TV	.07	.05	-.02	.12	.05
P2	Cinderella	S	.92	.90	-.02	.86	.06
		WFS	.23	.45	***.22***	.50	***.27***
		TA	.86	.71	.61	.94	.08
		TV	.29	.19	-.10	.24	-.05
	Cartoon	S	.88	.83	-.05	.80	-.08
		WFS	.13	.33	.20	.40	.27
		TA	.67	.80	.13	.78	.11
		TV	.33	.40	.07	.33	0
	Conversation	S	.46	.33	.13	.38	.08
		WFS	.17	.11	-.06	.08	-.09
		TA	.67	.72	.05	1	***.33***
		TV	.20	.22	.02	.27	.07

Table 11. Discourse sample grammatical analyses reported in proportions (proportion of sentences (S), proportion of well-formed sentences (WFS), tense accuracy (TA), and tense variety (TV) for individuals who received g-CILT.

Key: ***Bold and italicized*=** Significant by a criterion of a two standard deviation change score (note: Standard deviations calculated on the basis of o-CILT and g-CILT participant baseline data for each measure and sample)

ID	Discourse sample	Grammaticality measure (proportions)	Pre TX	Post TX	Pre-Post change	Maintenance TX	Pre-Maintenance change
P3	Cinderella	S	1	X*	X*	1	0
		WFS	.15	X	X	.46	***.31***
		TA	.76	X	X	.87	.11
		TV	.16	X	X	.10	-.06
	Cartoon	S	1	X	X	1	0
		WFS	.53	X	X	.27	.26
		TA	.73	X	X	.85	.12
		TV	.20	X	X	.20	0
	Conversation	S	.34	.60	.26	.73	.39
		WFS	.07	.40	.33	.38	.31
		TA	.55	.94	***.33***	.92	***.37***
		TV	.18	.18	0	.12	-.06
P4	Cinderella	S	.73	.80	.07	.67	-.06
		WFS	.20	.10	-.10	.22	.02
		TA	.71	.81	.10	.74	.03
		TV	.14	.12	-.02	.13	-.01
	Cartoon	S	.19	.88	.69	.73	.54
		WFS	.14	.25	.11	.18	.04
		TA	.86	.71	.15	.67	-.19
		TV	.29	.18	.11	.33	.04
	Conversation	S	.35	.26	-.09	.23	.12
		WFS	0	.10	.10	.13	.13
		TA	.75	.93	.18	.53	-.22
		TV	.17	.13	-.04	.18	.01

*Note- Data are unavailable for time points indicated by an “X” due to experimental error.

changes in narrative and conversational sentence level morphosyntactic measures and individuals receiving g-CILT produced more significant and better maintained results in all test contexts for aphasia severity (e.g., WAB-R, BNT, O&A) and sentence level grammaticality (e.g., VIT, WAB-R Grammaticality quotient). While all grammaticality tests and morphosyntactic measures were analyzed at the sentence level, participants in o-CILT and g-CILT performed differently in these contexts. It is important to note that the morphosyntactic measures require participants to use an increased cognitive load, as it is necessary to consider contextual elements. Several other factors may have impacted these results, such as small sample size, ceiling effects, and experimental error resulting in missing discourse data, and these are discussed in the following section.

CHAPTER 4: DISCUSSION AND FUTURE IMPLICATIONS

In this thesis, constraint-induced language therapy (CILT) as described by Pulvermuller, et al. (2001) and Maher, et al., (2006), (o-CILT) was administered to two individuals with agrammatic aphasia to determine the effects of CILT on this population and compare these results with efficacy found in other CILT studies. Grammatically modified CILT (g-CILT) was administered to two individuals with agrammatic aphasia, and outcomes were compared with data from two individuals receiving o-CILT to determine the effect of adding a grammaticality constraint. Participants were tested on measures of aphasia severity and measures of grammatical ability before treatment, immediately after treatment, and within three to four months after treatment. The treatment outcomes revealed that, in some ways, individuals with agrammatism respond differently to CILT than individuals with aphasia recruited in previous studies. Further, g-CILT and o-CILT produced differential outcomes for the individuals with agrammatism recruited in this study. The implications of these findings with reference to the research questions are addressed below.

4.1 Response to o-CILT by individuals with agrammatic aphasia

The first research question examined the efficacy of constraint-induced language therapy for individuals with agrammatic aphasia. Participants with agrammatism in this thesis study did not respond in the same way as previous CILT study participants which were conducted with heterogeneous (and often more severe) aphasia classifications. In the following paragraphs, results from post treatment and maintenance testing are discussed in the context of previous research.

4.1.1 Post treatment performance of individuals with agrammatism in o-CILT

On the basis of theoretical viewpoints on the CILT (Taub, 2006; Pulvermuller, et al., 2001), it was hypothesized that, if massed intervention and constraint of compensatory strategies is sufficient to overcome any kind of learned non use, participants with agrammatic aphasia in this study would show improvement in aphasia severity scores after receiving constraint-induced language therapy. Outcome measures of aphasia severity used in this thesis (WAB AQ, Kertesz, 1982; BNT, Kaplan, et al., 2000; O&A, Druks & Masterson, 2000) showed that individuals receiving o-CILT produced only a few positive significant changes from baseline to post-treatment testing, and these gains were not typically maintained. This finding partially supports the alternate hypothesis that massed intervention and constraint are not sufficient to evidence significant therapeutic outcomes for individuals with agrammatism. The next paragraphs review differences between the findings of this study and previous CILT literature at post-treatment and maintenance testing and potential reasons for the differences in performance are discussed.

In a direct comparison of results from the Maher, et al.'s (2006) study with participants receiving o-CILT in this thesis, post treatment gains are inconsistent. Maher, et al. CILT participants produced several positive significant gains in aphasia severity testing contexts, and participants receiving o-CILT in this thesis evidenced only two positive significant changes post treatment. While variability within individual participants on test measures is evident in the Maher, et al. study, an overall trend of significance can be reported from the data. In contrast the minimal significant gains at post-treatment testing evidenced by the present study are not consistent enough to assert

that individuals with agrammatism are able to benefit from o-CILT on standardized test measures. There are several possible explanations for this finding.

First, as it was hypothesized in this thesis, it may be that o-CILT is not sufficient to address the complex problem of agrammatism. As mentioned earlier, individuals with agrammatic aphasia have relatively good comprehension, and their primary difficulty is with morphosyntax. While a combination of massed intervention and constraint of compensatory strategies have been useful for participants in previous CILT studies, this treatment in its original form may not effectively target the primary area of difficulty for individuals with agrammatic aphasia. Barthel, et al. (2008) proposed that massed practice and constraints were not sufficient to evidence significant changes in comparison to a treatment with a deficit-specific focus in a time-intensive protocol. A lack of deficit-specific focus may explain the absence of a significant positive overall trend in data from participants in this study.

Second, it is important to consider differences in aphasia severity of participants in this study in relation to participants in previous CILT studies. Participants in this study can be classified as moderate (P1) and very mild (P2) individuals with aphasia. While previous studies have not classified participants as agrammatic, an examination of participants with Broca's aphasia or nonfluent aphasia reveals that the majority of participants in previous studies are classified as moderate to severe. Maher, et al. (2006) did not describe their participants with a specific aphasia classification, but baseline WAB AQ scores (scores: 40.6- 70) suggest that these participants had more severe aphasia than participants in this thesis (WAB AQ scores: 69.7-93.7). This difference in participant profiles could be an explanation as to why participants in this thesis, who

demonstrated less severe aphasia, did not produce as many significant changes as individuals in the Maher, et al. study. Previous CILT studies have suggested that effects of CILT may be related to pre-treatment aphasia severity (see Pulvermuller, et al., 2001). Specifically, individuals who demonstrate increased initial severity appear to produce more significant changes with CILT. Work by Meinzer, et al (2007); however, argues against this explanation. Meinzer, et al. (2007) reported including six participants with mild Broca's aphasia and five participants with moderate Broca's aphasia, and all but one of these participants produced significant gains on standardized aphasia test measures after completion of the CILT protocol. Meinzer, et al. did not provide associated test scores to support these classifications, but this finding suggests that it is possible for participants with mild and moderate nonfluent aphasia to benefit from this treatment protocol. CILT also demonstrates relative limitations for use with individuals with mild aphasia as assessed by tests. Participants with mild aphasia do not appear to make significant gains with CILT. P2 and P4 in this study showed no gains on test measures at which they approached the ceiling.

A third noteworthy point is individual performance variations. Previous studies have cited individual differences and variation in gains from participants receiving CILT, but in each of these studies, a general trend of improvement was identified. For example, Pulvermuller, et al. (2005) reported behavioral data for four individuals with mild or moderate Broca's aphasia which showed variability among participant test scores, in that one participant produced significant improvements on some subtests and significant declines in scores on other subtests. When examined in conjunction with gains produced by individuals with other aphasia classifications in the participant population; however,

Pulvermuller, et al. found a positive significant trend post-treatment. This type of conclusion has been reported in several studies (see Meinzer, et al., 2005; Brier, et al., 2006; Szaflarski, et al., 2008), and despite the individual variation in scores, these authors suggested that overall changes were significant. It may be that particular classifications of aphasia respond better to original CILT, and the inclusion of those types of participants in a heterogeneous cohort of participants may assist data analyses in producing overall positive gains, but it is still unclear what determines prognosis for CILT. Presence and severity of apraxia may play a role in recovery potential, as individuals in this study and the Maher, et al. (study) with apraxia did not make gains to the same extent as participants without apraxia. CILT is a general language stimulation protocol, and there is also a chance that it does not properly target certain characteristics of aphasia (e.g., agrammatism).

The small sample size in this thesis (n=2) also limits the extent of possible benefits that could be observed. A small cohort may have diminished possible significant effects as a result of an overall variation in significance scores. In previous CILT studies, with samples sizes ranging from three to twelve participants, not all participants improved on all test scores, but most participants improved in at least one test score (see Szaflarski, et al., 2008; Maher, et al., 2006; Brier, et al., 2006; Meinzer, et al., 2005; Pulvermuller, et al., 2005). If more participants had been included in the study, an overall positive trend may have appeared.

Finally, participants receiving o-CILT in this study may not have produced significant results mirroring those produced by Maher, et al. in post treatment testing because of the sensitivity of the measures used for individuals with mild aphasia. P2 did

not produce significant gains on two of the three aphasia severity test measures, but this may have been because he was at ceiling for both of these tests, and it would have been very difficult for this participant to evidence positive significant changes on these tests. It is widely recognized that there are a lack of sensitive tests to detect progress for individuals with mild aphasia, and the tests used in this thesis may not have adequately assessed P2's gains.

4.1.2 Maintenance of gains from o-CILT by individuals with agrammatic aphasia

Maintenance of gains demonstrated by participants receiving o-CILT in this thesis was poorer than the maintenance shown by Maher, et al.'s (2006) participants. Although thesis participants evidenced two significant gains at post treatment testing, only one of those gains was maintained. Maher, et al.'s participants maintained all but one of the gains produced in post treatment testing. While the small sample size of this thesis makes it difficult to determine the relevance of these results, gains in previous CILT studies contrast these results, as these gains are typically stable or demonstrate an increase over time (Meinzer, et al., 2005; Pulvermuller, et al., 2001). There are several possibilities which may account for this performance.

First, according to the thesis hypothesis, individuals with agrammatic aphasia may not benefit from o-CILT alone, and a deficit-specific modification (as proposed by Barthel, et al., 2008) may be necessary for this treatment to benefit this population. Maher, et al. (2006) demonstrated the importance of the constraint as a component necessary for maintenance of skills, but it appears as this constraint may not be sufficient to maintain significant changes for individuals with agrammatism.

A second explanation for this difference in findings is the time post treatment at which the maintenance testing was conducted. Maher, et al. (2006) administered maintenance testing at one month post treatment, and this thesis administered maintenance testing at three to four months post treatment. Hence, it may be the case that maintenance of gains diminished over time and may have been more salient at one month post testing as opposed to three months later. However, this may not be the entire cause because Meinzer, et al. (2005) performed a CILT study in which maintenance was assessed at six months post treatment, and gains were still present for the majority of participants.

Third, it is likely that participation in individual therapy following treatment may have impacted maintenance scores. P1 received one hour of individual treatment a week throughout the entire course of the study, and the focus of that therapy may have contributed to a change to her severity scores. For example, if P1's individual clinician chose to concentrate on comprehension and diminish focus on word retrieval, this may have assisted in changing her scores from post treatment to maintenance testing. This might affect participants on an individual basis. P2, however, did not receive individual or group therapy between the completion of treatment and maintenance testing, and he demonstrated improvement on one measure. None of the prior CILT studies report the amount of individual/ group therapy between post-treatment and maintenance testing. It is important to document intervening therapeutic participation, given that several studies have reported an increase in scores from post-treatment to maintenance testing. Other patient variables that can affect maintenance of skills are individual motivation and attitudes towards communication, presence of opportunities to practice verbal output,

personal events such as changes in health family dynamics etc. For example, it was reported that P1 experienced some personally stressful events just prior to the maintenance testing period. This is the potential disadvantage of single subject studies where such factors have a significant impact on outcomes

Another factor that needs further investigation in interpreting maintenance outcomes is the susceptibility of the test to practice effects, especially in studies where improvements have been observed from post- to maintenance. Unlike many neuropsychological tests where alternate equivalent forms are available to control for practice effects, this is not so commonly the case for aphasia tests. None of the severity tests used in this thesis had alternate forms.

4.2 Response to g-CILT by individuals with agrammatic aphasia

Individuals with agrammatism produced better outcomes on tests of aphasia severity and grammaticality after receiving the g-CILT intervention, but individuals receiving o-CILT demonstrated more significant and maintained gains in measures of grammaticality in discourse. The following paragraphs discuss possible explanations for these findings.

4.2.1. Post-treatment outcomes of aphasia severity measures

Individuals receiving the g-CILT protocol evidenced more significant gains on aphasia severity measures post treatment when compared to individuals receiving o-CILT. The significant performance of individuals receiving g-CILT is similar to the results which have been reported in previous CILT studies (Maher, et al., 2006; Meinzer, et al., 2005; Pulvermuller, et al., 2001). As previously discussed, performance of individuals receiving o-CILT did not replicate findings of previous CILT studies, and there are

several possible explanations for the significant post treatment results of individuals from the g-CILT protocol.

First, this thesis hypothesized that a grammatically modified CILT protocol would produce significant benefits for individuals with agrammatic aphasia on aphasia severity measures because of the deficit-specific focus to the time intensive protocol. A deficit-specific focus was initially proposed by Barthel, et al. (2008), and it appears that these results support this modification as a means of increasing performance on test measures of aphasia severity.

A second explanation of the findings could be related to the complexity of structures trained for production in each group. Individuals in g-CILT were required to produce more complex sentence structures during treatment relative to individuals in o-CILT. Several treatments for agrammatism, including treatment of underlying forms (Thompson & Shapiro, 2005), have supported the idea that training of more complex sentence structures in treatments results in generalization to similar sentence structures and less complex sentence structures. This theory is called the Complexity Account of Treatment Efficacy (Thompson, Shapiro, Kiran, & Sobecks, 2003), and it may be applicable to this study. More complex structures were trained during g-CILT, and these may have generalized to measures of aphasia severity more effectively than the less complex structures shaped in o-CILT. This explanation would suggest that findings in this study resulted from increased complexity of stimuli used in g-CILT relative to o-CILT.

The third possibility for the difference in significant changes on measures of aphasia severity is an effect from the grammaticality judgment. Participants in the g-

CILT group were asked to judge the grammaticality of the requesting participant's utterance, thereby completing an additional comprehension task which was not included for individuals in the o-CILT group. While it is possible that this task could have increased the scores of g-CILT participants on comprehension subtests of aphasia severity measures, this is highly unlikely, as the theory of learned non use claims that it is the actual practice of the affected speech process that produces significant effects (Pulvermuller, et al., 2001). There is no mention in this theory of the effect of a metalinguistic task, such as the grammaticality judgment. This suggests that learned non use may not be the only contributing factor to therapy success.

A final factor which may have contributed to the decrease in scores at post-treatment testing was a change in research assistants administering the treatment. All post-treatment testing assessments were administered by research assistants who were unfamiliar with the participants to ensure no administration bias, and this may have impacted the outcomes.

4.2.2 Maintenance of aphasia severity measures

A comparison of o-CILT and g-CILT participants on maintenance of significant aphasia severity measure scores revealed that participants receiving g-CILT maintained significant gains on these tests more frequently than individuals who produced significant gains in o-CILT. The maintenance of scores demonstrated by individuals receiving g-CILT is comparable to what has been reported for maintenance in previous CILT research (see Maher, et al. 2006, Meinzer, et al., 2005; Pulvermuller, et al., 2001). Possible explanations for these results are discussed below.

First, this thesis hypothesized that, based on the work of Barthel, et al. (2008), massed intervention and constraint of compensatory strategies would be insufficient to produce significant changes on aphasia severity measures, and a grammaticality modification to the CILT protocol would produce more significant and well maintained benefits for individuals with agrammatic aphasia. The addition of a grammaticality constraint is one possible explanation for the findings of this thesis, which show more frequent retention of gains for individuals who received a grammaticality modification to CILT.

Second, Maher, et al. (2006) demonstrated that the function of the constraint in CILT may be to facilitate maintenance of gains over time. This study added a grammaticality constraint to the pre-existing compensatory strategy constraint, and it is possible that maintenance occurred because of the addition of an additional constraint. If another aspect of the treatment carried the emphasis on grammaticality, the outcome may have been different. For example, if it was not required that all sentences be grammatical during treatment. In the shaping process, participants were given feedback and encouraged to produce grammatical sentences and it would be questionable as to whether generalization would still occur. In contrast to this idea, an individual in o-CILT did show maintenance of gains on one aphasia severity measure, so further experimentation is needed to establish the role of constraints in CILT for individuals with agrammatic aphasia.

Third, the maintenance effect may be a factor of group therapy, which both participants in the g-CILT group received regularly from post treatment to maintenance testing. In comparison, individuals in the o-CILT protocol did not receive individual or

group treatment as regularly, as P1 did not attend therapy for approximately one month during the period following treatment for scheduling and personal reasons, and P2 decided to take a break from group therapy for the entire duration of post treatment to maintenance testing. Participation in therapy following CILT may have been a factor in maintenance gains, but Meinzer, et al. (2005) showed that even when participants received individual treatment for an hour a week for six months between treatment and maintenance testing, 85% of scores remained almost exactly the same. This corroborates the Pulvermuller, et al. and others findings that a massed intervention, such as CILT, is more effective than therapy administered for an hour a week spread over a longer duration. These data (Meinzer, et al., 2005; Pulvermuller, et al. 2001) make an argument against the significant effects of less intensive therapeutic interventions following CILT.

4.2.3 Post-treatment outcomes of grammaticality measures

Post treatment scores of grammaticality measures revealed differential outcomes for individuals receiving o-CILT and g-CILT. On grammatical analyses of discourse samples, individuals receiving o-CILT surprisingly demonstrated more significant improvements at post testing when compared to individuals receiving g-CILT. Individuals receiving g-CILT achieved more significant scores on test measures of grammaticality. These findings are interpreted in the following paragraphs with two major considerations. The first consideration is that no CILT studies, to date, have reported using individuals with agrammatism in their studies, and no test measures of grammaticality have been reported thus far; thus it is not possible to compare scores from tests of grammaticality with previous research. Additionally, there is a dearth of CILT literature addressing how CILT gains translate to functional contexts, as measured by

linguistic analyses. The majority of studies, to date, have used a subjective communication assessment for participants or caregivers to rate changes in a participant's speech before and after treatment. Some studies have also used raters blind to condition and time of testing to determine which participants utilize speech more effectively following CILT. Two studies utilizing CILT have published data which include some linguistic analyses of narratives, and these findings can be used for comparison to data from this thesis. The studies which have explored the transfer of CILT into speech have looked at summary linguistic measures (e.g., total number of utterances, total number of words, total number of word roots, mean length of utterance, and type token ratio) in narrative discourse contexts only (Szaflarski, et al., 2008; Maher, et al., 2006). Of these measures assessed, the majority of participants demonstrated increases on all measures of linguistic analyses from baseline to post treatment testing. While it appears that participants in this thesis study performed differently from narrative data of previous CILT participants, the data from previous research must be carefully considered. Both studies reporting the narrative data listed large changes in scores from baseline to post testing, but there was no significance criteria associated with the data. That is to say, that the numbers were discussed if they appeared to be significant. Discourse data from this thesis were only considered significant if they were above a two standard deviation change from baseline to post treatment testing and, because of the inter-individual production variation for these participants, these numbers were typically very large. It is possible that with a less stringent significance criterion, participants in this study would have evidenced more frequent significant changes, therefore resembling previous narrative data which have shown larger changes between treatment conditions. If

morphosyntactic measures were considered significant at a one standard deviation criterion, 17 measures (3 changes with 2 SD criterion) would be significant for the g-CILT condition and 22 (7 changes with 2 SD criterion) measures would be significant from participants in the o-CILT condition. Also, previous studies (e.g., Rochon, et al., 2000; Webster, et al., 2001) have used the variability of normal participants on discourse measures as a reference criterion. This may be a useful standard in future studies. There are several possible explanations for the post-treatment grammaticality findings.

First, this thesis hypothesized that individuals receiving the g-CILT protocol would make more significant changes on measures of grammatical ability (e.g., VIT (Faroqi-Shah, unpublished), WAB Spontaneous Speech Grammatical Competence Subtest Score (Kertesz, 1982), and discourse analysis measures) when compared with individuals receiving the o-CILT treatment, which did not feature grammaticality. Results showed a different effect, with participants receiving g-CILT performing more significantly on the test measures of grammaticality and o-CILT participants demonstrating more significant transfers to discourse. While this effect was unexpected, it could be explained by the levels of difficulty utilized by the complexity rule constraint. This constraint required participants to increase the length and difficulty of the sentence they produced. These participants may have used a combination of compensatory strategies of pointing or gesturing and agrammatic speech characteristics such as single word output and reduced syntactic complexity to communicate before treatment, but o-CILT required them to systematically increase the complexity of their output. In this way, CILT may have assisted these participants in overcoming learned non use of grammatical speech. These results may be indicative of CILT's ability to generalize treatment effects

to functional situations, but it is more likely that discourse measures of increased length need to be collected in a more naturalistic and/or varied setting to observe the true effect of the treatment on spontaneous speech.

Secondly, these findings do not support the adaptation theory of agrammatism. According to the adaptation theory of agrammatism, agrammatic speech is a strategy used to decrease the computational workload necessary for formulating a grammatical sentence. In consideration of this theory, participants receiving g-CILT should have demonstrated more gains in conversational and narrative speech, as practice should theoretically help individuals with agrammatism overcome a learned non use of grammatical speech.

While the lack of transfer to functional speech demonstrated by g-CILT participants was unexpected, it is possible that the grammaticality constraint was too structured to produce gains which could transfer to functional speech. Within the context of the game, production of sentences was highly structured, while conversation requires a combination of comprehension and production demands to develop grammatically-correct output. According to the complexity account of treatment efficacy (CATE; Thompson, et al. 2003), training more complex utterances should allow for generalization to similar and less complex structures, but in g-CILT participants were asked to produce a temporal adverb followed by a grammatically-correct sentence (e.g., “Yesterday, he walked). The temporal adverb was used to assist in the formulation of a grammatically correct sentence, but use of this sentence structure is less likely in narrative and conversational speech. For example, it would be more natural to say “Cinderella ran down the stairs” as opposed to “Yesterday, Cinderella ran down the stairs”. Henceforth,

it is possible that g-CILT trained sentences that were less complex than the sentences trained in o-CILT because of the use of the temporal adverb. Individuals in the o-CILT group often tried to use grammatical speech, even if they relied on the same grammatical structure during each sentence production. They attempted to do this without a temporal adverb. If the temporal adverb assisted in simplifying the sentence, it would be expected that generalization would not occur to discourse, as the protocol trained a less complex sentence. According to CATE, generalization would not occur as readily with less complex sentences, as used in this study, in comparison with training utilizing more complex sentences. This explanation is also supported by performance by g-CILT individuals on the VIT. Sentences, such as the ones elicited by the picture description task, were trained in the g-CILT protocol, and participants who received g-CILT demonstrated significant increases on this measure. Individuals with agrammatism who received g-CILT performed similarly to participants in treatments such as linguistic-specific treatment, mapping therapy, the sentence production program for aphasia, and response elaboration training in that these approaches demonstrate significant acquisition of trained structures, but generalization to untrained structures on narrative discourse measures is limited (see Ballard & Thompson, 1999; Fink, Schwartz, Rochon, Myers, Socolof, & Bluestone, 1995; Helm-Estabrooks, Fitzpatrick, & Barresi, 1981).

A second and more practical possibility which could account for the differences between g-CILT and o-CILT participants on linguistic analyses of discourse is experimental error. Two of four narrative discourse measures were not collected at post treatment testing for P3, and as a result, it is unknown if any significant changes occurred at that time point. Fourth, P4 was diagnosed with a mild apraxia, which primarily

impacted his production of free and bound grammatical morphemes as opposed to content words. Also, he was often reluctant to speak in narrative and conversational contexts because of his quick frustration level. This participant had the most difficulty adjusting to the constraint of speech only, and he was the only participant who did not reach the highest level of the task complexity hierarchy. It is possible that this participant's apraxia and/or reluctance to speak impacted his willingness to produce discourse, and in turn, his discourse grammaticality measures. Participants in the Maher, et al. study (2006) were also reported to have mild to severe apraxia, and this potential confound was noted in the discussion of that paper.

A third explanation of the differential performance between conditions may be the amount of cueing provided. Increased cueing is accompanied by an increased number of trials to produce a correct utterance. Therefore, individuals who received more cueing received a greater number of trials to use speech. These extra trials using speech could provide participants who receive more cueing with an advantage. For participant groups utilizing less cueing, more frequent conversational exchanges can occur, thus making the interactions more natural. This could aid in the transfer of gains to naturalistic speech contexts. Although it was not explicitly measured, participants in the g-CILT group received more cueing because of the increased complexity of their constraints (e.g., inclusion of temporal adverb) and the presence of mild apraxia (P4). Participants in the o-CILT condition did not receive as much cueing, and interacted more naturally with each other. This may have influenced the response of participants to the assessment measures. It could explain why g-CILT participants appeared to perform more accurately

on test measure contexts of sentence production and o-CILT participants produced more significant morphosyntactic measures in discourse contexts.

4.2.4 Maintenance of grammaticality measure outcomes

Maintenance testing measures of grammaticality demonstrated significant retention or improvement of skills at three months post-testing for both individuals in o-CILT and g-CILT. The hypothesis for this study stated that if massed practice and constraint of compensatory strategies were inadequate to evince significant changes and maintenance of those changes, a grammatical modification to the CILT protocol would be able to produce significant benefits for individuals with agrammatism and these gains would be maintained. Post-treatment gains indicated that individuals who received o-CILT produced more significant treatment gains on discourse measures of grammaticality and participants who received g-CILT demonstrated more significant gains on grammaticality test measures, unexpectedly challenging this hypothesis. Gains in both of these respective contexts were maintained, suggesting that an element common to both treatments is the source of gain maintenance. Participants in g-CILT demonstrated more overall maintained gains on measures of grammaticality. Possible explanations for this finding are discussed below.

The first possible explanation is that these results have replicated the findings of Maher, et al. (2006) in that constraint of compensatory strategies contributes to the CILT protocol by promoting maintenance of gains. The use of only verbal communication was enforced for participants in o-CILT and g-CILT, and all participants showed maintenance of gains. It is possible; however, highly unlikely, that the verbal constraint, material difficulty constraint, complexity hierarchy constraint, or a combination thereof produced

maintenance of gains, since all three constraints were present in each condition. Interestingly, gains were maintained without the use of the physical barrier with participants in g-CILT, thus adding efficacy to the argument against the use of a physical barrier. When comparing participants in each group; however, g-CILT participants demonstrated more overall maintained gains. This may be a result of the lack of physical barrier, as Szaflarski, et al. (2008) suggested that the barrier was an unnatural context to use within a therapy protocol which focuses on working within a naturalistic interaction setting.

A second possibility for the maintenance of gains would be the participants' involvements with individual and group therapy between the end of treatment and maintenance testing. Participants may have continued to work on skills tested on grammaticality measures during the months following therapy. As was previously discussed, while this is an argument worth noting, it is unlikely that therapy provided in a less time intensive format would provide additional benefits (see Meinzer, et al., 2005). Increased gains from post treatment were displayed by several participants on test measures, including P2, who did not receive individual or group therapy between post testing and maintenance sessions. While individual variations may impact performance between these two testing sessions, it is not likely that individual therapy produced increases in performance. It is possible that individual therapy may have helped maintain these gains.

4.3 Study Limitations

The findings of this study suggest that, for individuals with agrammatism in this study, CILT in its original form increases grammaticality of narrative and conversational

speech and a grammatically modified CILT decreases aphasia severity. Overall, these benefits were maintained at testing three months following treatment. While these results are interesting, it is important to consider limitations of this study to assist in appropriate interpretation of these results and to help in designing future studies.

4.3.1 Limitations of participant cohort and testing measures

This study was performed using a convenience sample of individuals with agrammatic aphasia who had also received individualized treatment focusing on grammaticality from this research center. As a result, these participants may respond differently to CILT than individuals with agrammatic aphasia who have not participated in previous intensive research in grammaticality. This was also an extremely small study cohort (n=2) and this made it difficult to determine reliable statistical significance. Additionally, participant characteristics may have impacted the results, as all participants enjoyed and were willing to participate in intensive social interactions. P4's mild apraxia may have impacted his ability to produce free and bound grammatical morphemes and the health of P1 was questionable at maintenance testing, as she had just had a root canal.

Although this study attempted to create a thorough testing protocol to assess changes in a variety of language domains, there were several limitations to the testing battery. The shortest time provided between testing intervals was approximately six weeks (baseline to post testing), and there was potential for a repeated exposure effect. Statistics were chosen in an attempt to elicit the effect, but because of the small sample size, it was difficult to explicitly address repeated exposure effects. Two participants (P2

and P4) were also near ceiling on several of the testing measures used at baseline, which made it difficult to properly assess change on these measures.

While this thesis aimed to collect naturalistic speech samples from participants, asking the participants to retell stories and converse with a researcher in the presence of a camera and audio recorder may have influenced the speech of the participants. This study was also limited in its ability to establish true baseline variability for participants in discourse contexts. Because of the significant variability of individuals with agrammatic aphasia in their speaking abilities as a function of context, conversational partner, or general health issues, baseline measures of conversation would be more effective if collected over several time points to obtain a more realistic representation of the participants' abilities.

4.3.2 Limitations of protocol and analysis measures

Group treatment studies require a significant amount of cooperation in terms of time commitment on the part of all participants involved in treatment. Although treatments were arranged to accommodate all participants, variables such as severe weather conditions and illness impacted the treatment schedule. All participants still received 24 hours of therapy a week over a two week time span, but there were sessions in which a volunteer unfamiliar with the aims of the study played with a study participant. This has been done in previous CILT studies (see Meinzer, Streiftau, & Rockstroh, 2007) with demonstrated efficacy, but it may have added variability to the treatment protocol. Analyses measures also may have introduced some variability into the results, as investigators found it difficult to agree on the definition of a sentence. Also, as it was mentioned in the methodology, in order for a change score on a discourse measure to be

considered significant, the score had to be two standard deviations from baseline, and the standard deviation was calculated on the variability of all participants on that measure at baseline. Verbal productions of individuals with aphasia vary on a daily basis, and formulating a baseline standard deviation from discourse samples taken at one point in time may not be wholly representative of the participants' abilities. This criterion may have been too specific to identify significant changes on linguistic analyses measures.

4.4 Implications and Future Directions

4.4.1 Implications

To our knowledge, this is the first study to examine the efficacy of CILT for agrammatic aphasia. Hence the findings add crucial information to the existing body of literature on CILT as well as on treatment for agrammatism. The findings of this study highlight the importance of deficit specific modifications within the context of constraint-induced language therapy for increased efficacy of this treatment for individuals with agrammatic aphasia to reduce severity of aphasia. Results suggest that CILT in its original form may not as effective as a deficit-specific version of CILT targeting grammatical sentence production for ascertaining and maintaining significant aphasia severity and grammaticality test scores. Participants receiving o-CILT, which did not focus on grammaticality, demonstrated more significant changes on grammaticality measures in discourse, suggesting that an intensive treatment with massed intervention, constraint of compensatory strategies, and shaping may be beneficial for functional transfer of therapeutic gains. There is also a relative limitation of CILT for individuals with mild aphasia, as assessed by test measures and demonstrated by P2 and P4.

4.4.2 Future directions

This thesis study raised several important questions regarding the benefit of constraint-induced language therapy for individuals with agrammatic aphasia. Future investigations should attempt to replicate the study and results with a larger participant cohort and use other grammaticality measures to more specifically determine the effect of this treatment on grammaticality. This line of research should target carryover of CILT in natural contexts, for example, describing activities performed during the day or conversation through the use of lengthy discourse samples and thorough and alternate linguistic analyses than those used in this study. The addition of other constraints which address the comprehension and production demands for grammaticality with a larger group of individuals with agrammatic aphasia would provide additional information on the role of the constraint within CILT for agrammatism. Also, a post-hoc analysis analyzing patient characteristics/demographics which predict success with CILT would be very beneficial for this area of research. As research builds to support a deficit-specific model of CILT, investigators might also address the response of other deficits associated with aphasia to a deficit specific CILT protocol, for example word retrieval deficits targeted within the context of CILT. As several studies begin to isolate the necessary components of an effective CILT protocol, clinical field trials should be conducted to assess the feasibility of utilizing this treatment in a real world setting. Finally, in order to truly ascertain the benefit of CILT for individuals with agrammatism, further research should directly compare CILT with other treatments for agrammatism.

It is important to address this study's position in terms of the developing database of aphasia treatment efficacy data (ANCDS, 2008). This was a single-subject design with a weak class of evidence examining multiple baseline studies over several behaviors, and it is a phase 1 study, as it is a pre-treatment study designed to examine if a treatment has any potential efficacy. It is essential that future studies focus on producing well controlled treatment studies to increase the strength and phase of studies. This study was a starting point for future examination of a deficit-specific approach to CILT. As investigators begin to better understand constraint-induced language therapy and its underlying mechanisms for change, it is the hope that this treatment will significantly impact the delivery of significant and long lasting treatment effects for individuals with chronic aphasia.

REFERENCES

AllPsych online (2008). Depression Screening Test. Retrieved on 10/20/2007 from http://allpsych.com/tests/diagnostic/depression.html.

American Heart Association (AHA) (2008). Know the facts, get the stats. Retrieved on 6/16/08 from http://www.americanheart.org/downloadable/heart/ 116861545709855-1041%20KnowTheFactsStats08_loRes.pdf.

American Psychiatric Association (APA) (1994). *Diagnostic and Statistical Manual of Mental Diseases*. American Psychiatric Publishing.

American Speech, Language, and Hearing Association (ASHA) (2007). Aphasia. Retrieved on 5/10/07 from http://www.asha.org/public/speech/disorders/ Aphasia_info.htm.

Academy of Neurological Communication Disorders & Sciences (ANCDS) (2008). Overall language performance studies. Retrieved on 6/15/2008 from http://www.u.arizona.edu/~pelagie/ancds/overall.html.

Ballard, K. & Thompson, C. (1999). Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech, Language, and Hearing Research, 42,* 690-707.

Hanna, W. & Barbera, J. (Producers). (2000) *Tom and Jerry's greatest chases.* [Motion picture]. United States: Warner Home Video.

Barthel, G., Meinzer, M., Djundja, D., & Rockstroh, B. (2008). Intensive language therapy in chronic aphasia: Which aspects contribute the most? *Aphasiology, 22*, 408-421.

Bastiaanse, R. (1995). Broca's aphasia: A syntactic and/or a morphological disorder? A case study. *Brain and Language, 48,* 1-32.

Brier, J., Maher, L., Schmadeke, S., Hasan, K., & Papanicolaou, A. (2007). Changes in language specific brain activation after therapy for aphasia using magnetoencephalography: A case study. *Neurocase, 13,* 169-177.

Breier, J., Maher, L., Novak, B., & Papanicolaou, A. (2006). Functional imaging before and after constraint-induced language therapy for aphasia using magnetoencephalography. *Neurocase, 12, 322*-331

Bhogal, S., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke, 34,* 987-993.

Bhogal, S., Teasell, R., Foley, N., & Speechley, M. (2003). Rehabilitation of aphasia: More is better. *Topics in Stroke Rehabilitation, 10,* 66-76.

Chapey, R. (2001). *Cognitive Approaches to Impairments of Word Comprehension and Production.* Brooklyn, NY: Lippincott, Williams, & Wilkins.

Chapey, R. (2001). *Language intervention strategies in aphasia and related neurogenic communication disorders.* Brooklyn, NY: Lippincott, Williams, & Wilkins.

Doyle, P., Goldstein, H., & Bourgeois, M. (1987). Experimental analysis of syntax training in Broca's aphasia: A generalization and social validation study. *Journal of Speech and Hearing Disorders, 52,* 143-155.

Druks, J. & Masterson, J. (2000). *An Object and Action Naming Battery.* London: Psychology Press.

Faroqi-Shah, Y. (in press) A comparison of two theoretically-driven treatments for verb inflection deficits in aphasia. *Neuropsychologia.*

Faroqi-Shah, Y. & Thompson, C. (2007). Verb inflections in agrammatic aphasia: Encoding of tense features. *Journal of Memory and Language, 56,* 129-151.

Faroqi-Shah, Y. (unpublished). *Verb Inflection Test.*

Fink, R., Schwartz, M., Rochon, E., Myers, J., Socolof, G., & Bluestone, R. (1995). Syntax stimulation revisited: An analysis of generalization of treatment effects. *American Journal of Speech-Language Pathology, 4,* 99-104.

Francis, W. & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Friedmann, N. & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: pruning the syntactic tree. *Brain and Language, 56*, 397-425.

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination.* Philadelphia: Lea & Febinger.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language, 27,* 135-159.

Heeschen, C & Schegloff, E. (1999). Agrammatism, adaptation theory, conversation analysis: on the role of so-called telegraphic style in talk-in-interaction. *Aphasiology, 13,* 365-405.

Helm-Estabrooks, N., Fitzpatrick, P., & Barresi, B. (1981). Response of an agrammatic patient to a syntax stimulation program for aphasia. *Journal of Speech and Hearing Disorders, 46,* 422-427.

Holland, A., Fromm, D., DeRuyter, F., Stein, M. (1996). Treatment efficacy: Aphasia. *Journal of Speech and Hearing Research, 39,* S27-S36.

Isserlin, M. (1985). On agrammatism. *Cognitive neuropsychology 2,* 308-345.

Kaplan, E., Goodglass, H., & Weintraub, S. (2000). *Boston Naming Test.* Philadelphia: Lea & Febiger.

Kearns, K., & Scher, G. P. (1989). The generalization of response elaboration training effects. *Clinical Aphasiology, 18*, 223-245.

Kegl, J. (1995). Levels of representation and units of access relevant to agrammatism. *Brain and Language, 50,* 151-200.

Kertesz, A. (2007). *The Western Aphasia Battery.* New York: Grune and Stratton.

Kolk, H. & Heeschen, C. (1992). Agrammatism, paragrammatism, and the management of language. *Language and cognitive processes, 7,* 89-129.

Kolk, H. & Van Grunsven, M. (1985). Agrammatism as a variable phenomenon. *Cognitive neuropsychology, 2,* 347-384.

Lillie, R., & Mateer. C.A. (2006). Constraint-based therapies as a proposed model for cognitive rehabilitation. *Journal of Head Trauma Rehabilitation, 21*, 119-130.

Maher, L., Kendall, D., Swearengin, J., Rodriguez, A., Leon, S., Pingel, K., Holland, A., & Gonzalez-Rothi, L. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society, 12,* 843-852.

Meinzer, M., Streiftau, S., & Rockstroh, B. (2007). Intensive language training in the rehabilitation of chronic aphasia: Efficient training by laypersons. *Journal of the International Neuropsychological Society, 13,* 846-853.

Meinzer, M., Elbert, T., Djundja, D., Taub, E., & Rockstroh, B. (2007). Extending the constraint-induced movement (CIMT) approach to cognitive functions: Constraint-induced aphasia therapy (CIAT) of chronic aphasia. *NeuroRehabilitation, 22,* 311-318.

Meinzer, M., Djundja, D., Barthel, G., Elbert, T., & Rockstroh, B. (2005). Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. *Stroke, 36,* 1462-1466.

Meinzer, M., Elbert, T., Weinbruch, C., Djundja, D., Barthel, G., & Rockstroh, B. (2004). Intensive language training enhances brain plasticity in chronic aphasia. *BMC Biology,2,* 1-9.

Miller, J.F., & Chapman, R.S. (2000). SALT: A computer program for the Systematic Analysis of Language Transcripts. Madison, WI: University of Wisconsin.

Mitchum, C., Haendiges, A., & Berndt, R. (1993). Model-guided treatment to improve written sentence production: A case study. *Aphasiology, 7,* 71-109.

Morris, D., & Taub, E. (2001). Constraint-Induced therapy approach to restoring functioning after neurological injury. *Stroke Rehabilitation, 8,* 16-30.

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

Pedersen, P., Jorgensen, H., Nakayama, H., Raaschou, H., & Olsen, T. (1995). Aphasia in acute stroke: incidence, determinants, and recovery. *Annals of Neurology, 38*, 659-666.

Pulvermuller, F., Hauk, O., Zohsel, K., Neininger, B.,& Mohr, B. (2005). Therapy-related reorganization of language in both hemispheres of patients with chronic aphasia. *Neuroimage, 28,* 481-489.

Pulvermuller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke, 32,* 1621-1626.

Richter, M., Miltner, W., & Straube, T. (2008). Association between therapy outcome and right-hemispheric activation in chronic aphasia. *Brain, 131,* 1391-1401.

Rochon, E., Laird, L., Bose, A., & Schofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15,* 1-36.

Rochon, E., Saffran, E., Berndt, R., Schwartz, M. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language, 72,* 193-218.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain, 129,* 1371-1384.

Salis, C. & Edwards, S. Adaptation theory and non-fluent aphasia in English. *Aphasiology, 18*, 1103-1120.

Shewan, C. & Donner, A. (1988). A comparison of three methods to evaluate change in the spontaneous language of aphasic individuals. *Journal of Communication Diroders, 52,* 171-176.

Shewan, C. & Kertesz, A. (1980). Reliability and validity characteristics of the Western Aphasia Battery (WAB). *Journal of Speech, Language, and Hearing Research, 45,* 308-324.

Sterr, A., Szameitat, A., Shen, S., & Freivogel, S. (2006). Application of the CIT concept in the clinical environment: Hurdles, practicalities, and clinical benefits. *Cognitive Behavioral Neurology, 19*, 48-54.

Sunderland, A., & Tuke, A. (2005). Neuroplasticity, learning, and recovery after stroke: A critical evaluation of constraint-induced therapy. *Neuropsychological Rehabilitation, 15*, 81-96.

Szaflarski, J., Ball, A., Grether, S., Al-fwaress, F., Griffith, N., Nells-Strunjas, J., Newmeyer, A., & Reichhardt, R. (2008). Constraint-induced aphasia therapy stimulates language recovery in patients with chronic aphasia after ischemic stroke. *Medical Science Monitor, 14,* 243-250.

Szekely, A., Jacobsen, T., & D'Amico, S. (2004). *Journal of Memory and Language,51,* 247-250.

Taub, E. (2006). The learned nonuse phenomenon; implications for rehabilitation. *Europa medicophysica, 42*, 241-255.

Thompson, C., Kearns, K., & Edmonds, L. (2006). An experimental analysis of acquisition, generalization, and maintenance of naming behaviour in a patient with anomia. *Aphasiology, 20,* 1226-1244.

Thompson, C. & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of underlying forms. *Aphasiology, 19,* 1021-1036.

Thompson, C. (2000). Neuroplasticity: Evidence from aphasia. *Journal of Communication Disorders, 22,* 357-366.

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

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

Wambaugh, J., Martinez, A., & Alegre, M. (2001). Qualitative changes following application of modified response elaboration training with apraxic-aphasic speakers. *Aphasiology, 15,* 965-976.

Webster, J., Franklin, S., & Howard, D. (2001). An investigation of the between thematic and phrasal structure in nonfluent agrammatic subjects. *Brain and Language, 78,* 197-211.

Virion, C. (unpublished). *Article Contrastive Judgment Test.*