The Impact of Computer-based Cognitive Treatment on Language Skills in an Individual with Aphasia

by

Caitlin Malli

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Speech-Language Pathology

Department of Communication Sciences and Disorders
University of Alberta

Abstract

The cognitive theory of aphasia, which purports that the language impairments found in people with aphasia are due to underlying cognitive impairments, rather than to interruption of linguistic-specific areas of the brain, has been gaining clinical and research interest in recent years. Indeed, treatments targeted towards remediating cognitive impairments have resulted in improvements in cognitive and language functioning. Further, computer-based interventions have shown promise as a means for increasing therapy intensity without increasing workloads for therapists. The purpose of this study was to investigate the efficacy of a commercially-available, computer-based cognitive training program (*BrainFitness*) as an intervention for an individual with aphasia.

Following 8 weeks (approximately 40 hours) of treatment, the participant demonstrated gains in language functioning as measured by the *WAB-R* and certain subtests of the *Alberta Language Function Assessment Battery* (*ALFAB*). The impact of the training on cognitive functioning was less clear. The results of this study suggest computer-based cognitive training may potentially benefit people with aphasia, but continued research is warranted. This type of treatment is not expected to replace therapy with speech-language pathologists, but to supplement the therapy already available to increase intensity of treatment without increasing workload for therapists. Examining a model addressing the neural connections underlying improvements resulting from auditory-based cognitive treatment may explain the mechanisms of recovery and the aspects of cognitive computer training that are most beneficial to recovery.

Preface

This thesis is an original work by Caitlin Malli. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "The Impact of Computer-Based Language Training in Chronic Aphasia", No. Pro00017432, March, 2011.

Acknowledgments

First of all, I would like to thank my amazing supervisor Dr. Kim for all of her support and guidance throughout this process. I can honestly say that this experience would not have been as fulfilling without her. She was always available to answer any questions I had, whether it was research-based or technical support during the writing process. Thank-you Dr. Kim for always replying quickly to my constant emails and overly detailed questions and for always being around to support me however I needed.

I would also like to thank my thesis committee for their time and efforts spent reviewing my document. Their comments and suggestions helped to make my thesis much stronger.

Additionally, I would like to thank R.D. for agreeing to participate in this research and for his continued motivation to complete the training across the course of the study. It was a pleasure to work with him and I am very fortunate that he agreed.

Finally, I would like to thank my incredible family and friends for their constant support. I would particularly like to thank my mother and father for constantly cheering me on in this process and for providing constant reminders that I was capable of succeeding. I could not have done this without their love and support.

Table of Contents

List of Tables

List of Figures

Introduction and Literature Review

Aphasia is an acquired language disorder, which can affect a person's ability to speak, understand, read and write. There are over 100,000 Canadians living with aphasia today (Aphasia Institute, 2012). The most common cause of aphasia is stroke, although it can also be caused by traumatic brain injury, tumours or dementia. One of the greatest risk factors for stroke is age (Heart and Stroke Foundation, 2013). Given that the segment of the population aged 65 and over is rapidly increasing, this means the prevalence of aphasia is also going to increase in the upcoming years. This increase in individuals who require therapy for the communication deficits in aphasia will place a large economic burden on the Canadian health care system. Therefore, more research needs to be done on effective and cost-efficient therapies to treat this condition.

Cognition and Aphasia

Individuals with aphasia have impairments using and understanding language. Naturally, these impairments can have significant consequences on many areas of functioning, including social, emotional and occupational domains. Although researchers agree that these language impairments exist in aphasia and affect communicative functioning, there is some debate over the underlying cause of this language disruption. Some believe that deficits in aphasia are due to interruption of linguistic-specific areas in the brain (Grodzinsky, 1990). Others believe that these linguistic processes are intact and that underlying impairments in cognitive functions are contributing to the observed language deficits (McNeil, Odell & Tseng, 1991). There are two cognitive domains in particular, attention and working memory, which have been shown to be impaired in people with aphasia (Hula & McNeil, 2008; Friedmann & Gvion, 2003). Notably, auditory language comprehension has been shown to be related to both attention and working memory (see Murray, 2012a for a review); therefore, it is reasonable to infer that individuals with aphasia who have impairments in these cognitive domains will also have impairments in auditory language comprehension.

Attention Deficits in Aphasia. Attention is defined as a "multidimensional cognitive construct… [that] include[s] the functions of vigilance or sustained attention…selective attention…attention switching…and divided attention" (Murray, 2012a, p.S51). The amount of attention focused on a task depends on the task demands and there is some evidence to suggest individuals with left hemisphere damage may show a different profile of attention deficits compared to individuals with right hemisphere damage. In a study by Sturm and colleagues (1997), a link between left hemisphere damage, which is the most common cause of aphasia, and greater difficulty with higher level attention tasks was found. In this study, Sturm et al. administered a variety of attention tasks to individuals with left and right hemisphere damage. The 22 participants with a left-hemisphere stroke (19 with aphasia) varied in their performance, but all demonstrated impairments in at least one area of attention, including higher level functions such as divided attention. This differed from the 16 right-hemisphere stroke patients, who displayed deficits in more basic areas of attention, such as alertness and vigilance. This study provides evidence for the link between left-hemisphere lesions and higher level attentional problems. Individuals with left hemisphere damage had more severe attention problems than individuals with right hemisphere deficits. However, one methodological issue in this study is that Sturm et al. did not differentiate between subjects with aphasia and those without in their sample of individuals with left hemisphere damage. Nor did they report the severity or types of aphasia, so it could not be determined from the study if any aphasia-specific patterns were evident with regard to attention impairments.

On the other hand, other investigators have studied individuals with aphasia specifically to determine to what extent language difficulties are due to underlying cognitive impairments. Murray (2012a) examined the relationship between language and attention in persons with aphasia compared to persons without. Subjects were a group of individuals with varying types of aphasia (n = 39) and a group of age and education-matched controls (n = 39) for a total of 78 adults. All participants were given a range of cognitive and language tests; not surprisingly Murray reported that subjects with aphasia performed significantly worse across all measures of attention. In addition, performance on these attention measures was significantly correlated with performance on the language and communication measures. The *Aphasia Diagnostic Profiles* (*ADP;* Helm-Estabrooks, 1992) Aphasia Severity score was significantly correlated with performance on the attention test. ADP auditory comprehension and lexical retrieval scores also correlated with many of the attention measures. Murray's results support the notion that cognition and aphasia are strongly linked and both cognition and language should be examined with regards to therapy for individuals with aphasia. However, not all participants with aphasia scored within the impaired range on formal attention measures, suggesting a weak model of attention where cognitive impairments may not *cause* aphasic symptoms but *exacerbate* them.

In contrast, Hula and McNeil (2008) propose a strong model of attention, which states that the deficits found in aphasia are due solely to underlying cognitive deficits. They argue that the language impairments observed in aphasia are the direct result of deficits in attention and resource allocation. They used dual task paradigms to assess attention allocation abilities in these individuals. These tasks require subjects to simultaneously perform two cognitive operations, such as naming a picture as quickly as possible and then pressing a button indicating the pitch of a tone played along with the picture stimulus. Individuals with aphasia tend to exhibit significantly greater disruption of auditory processing skills in dual task paradigms compared to control individuals. That is, individuals with aphasia demonstrate much slower response times when presented with tasks requiring a division of attention allocation (Hula, McNeil & Sung, 2007). Hula and McNeil (2008) propose that this slowed auditory processing ability is caused by competing language and non-language processes sharing limited capacity parallel processing resources. The greater disruption of processing found in individuals with aphasia could be due to a reduction in processing capacity or ability to allocate that capacity. Additionally, group differences were larger if both tasks required the use of language, such as a semantic identification task (e.g., naming the category of the stimulus presented) rather than a tone naming task, which suggests a language-specific reduction in attentional abilities for individuals with aphasia. These studies have provided evidence for a possible link between the linguistic deficits present in individuals with aphasia and underlying attention deficits. Specifically, it has been proposed that attention and working memory deficits contribute to the auditory comprehension deficits present in individuals with aphasia.

Working Memory Deficits in Aphasia. General agreement exists in the literature that impairment in working memory can contribute to a decrease in language processing abilities, just as a deficit in attention does (Christensen & Wright, 2010; Hula & McNeil, 2008; Sung et al., 2009; Wright & Fergadiotis, 2012). The working memory system is most often conceived of as a limited capacity "verbal or nonverbal buffer or working space within which specialized linguistic or non-linguistic operations or computations occur" (Hula & McNeil, 2008, p. 173). Similar to the research supporting the role of attention in aphasia, working memory impairments in people with aphasia are found to be specific to processing linguistic information. Evidence for working memory deficits in aphasia come from research on a variety of tasks, such as phonological working memory tasks, *n*-back tasks and sentence comprehension, digit span and word span tasks (Caspari, Parkinson, LaPointe & Katz, 1998; Christensen & Wright, 2010; Friedmann & Gvion, 2003; Gvion & Friedmann, 2012; Martin, Kohen, Kalinyak-Fliszar, Soveri & Laine, 2012; Wright & Fergadiotis, 2012).

Christensen and Wright (2010) used visual *n*-back tests to measure how individuals with aphasia perform on working memory tasks with varying levels of linguistic processing demands. An *n*-back test is one where a participant is presented with a continuous stream of stimuli and is required to respond (e.g. press a button) when the stimulus presented is the same as the one *n*-back (i.e. 1-back or 2-back). All participants (12 individuals with aphasia and 12 controls) performed better when the stimuli used in the *n*-back task had a high linguistic load (e.g. well known, easily nameable items) compared to a low linguistic load (e.g. abstract blocks in different shapes). The authors suggested that the poorer performance on the more abstract tasks was due to a decreased ability to use linguistic strategies. The individuals with aphasia also performed worse in general than controls on the *n*-back tasks, reflecting their impaired working memory abilities. This study supports the notion that people with aphasia not only have a smaller capacity for working memory, but also that linguistic factors seem to have a greater impact on the performance of people with aphasia relative to controls.

Another source of evidence of working memory deficits in aphasia comes from Friedmann and Gvion (2003), who assessed the relationship between sentence comprehension and phonological short-term memory in individuals with conduction aphasia. Participants were required to make plausibility judgements for 180 sentences and to paraphrase the sentence as completely as possible. Eighty of the sentences contained an ambiguous word and context was biased towards a certain incorrect interpretation of that sentence (e.g. “The *pen* is always packed with woolly sheep"). When the subject realized that his/her initial interpretation of the sentence was incorrect, he/she had to be able to re-access the other meaning of the ambiguous word (i.e. phonological reactivation) in order to correctly judge the sentence. Individuals with aphasia were significantly worse at comprehending sentences than controls, especially when the sentences were long. Individuals with conduction aphasia were not able to comprehend sentences when there were 7-9 words between the ambiguous word and the resolution, rating half of these sentences as implausible (not significantly different from responses expected from chance alone). However, subjects were able to make correct judgements for the same ambiguous words when the ambiguity was exposed within 2-3 words. Friedmann and Gvion (2003) concluded that the sentence comprehension failures of individuals with aphasia could be attributed to working memory limitations.

Gvion and Friedmann (2012) were able to replicate this finding in a later, larger study, using 12 individuals with conduction aphasia and 296 control participants. Individuals with aphasia were able to comprehend fairly long sentences, except in cases where phonological reactivation was required. This suggests separate components in working memory for comprehension of syntax, semantics and phonology, with only phonological working memory playing a role in the impairments that accompany aphasia. The authors point out that these results indicate a specific impairment of phonological working memory and that daily sentence understanding may not be impaired in most cases. The specific system of working memory can become overloaded and lead to the decrease in ability. The fact that in both studies individuals with aphasia were able to comprehend the sentences- except when phonological reactivation, a language-specific working memory process, was required - supports the role of working memory limitations contributing to language impairments found in people with aphasia.

Treatment of Cognitive Impairments in Aphasia

The studies reviewed above support the theory that deficits in underlying cognitive abilities may contribute to language processing impairments in aphasia. These impaired cognitive processes are good candidates for treatment, provided they are treated within the context of linguistic operations (Gvion & Friedmann, 2012; Hula & McNeil, 2008). Although research on the treatment of cognitive impairments in aphasia is limited to mostly small scale, pre-efficacy studies according to Robey and Schultz's (1998) five-phase model of treatment research, findings from these studies suggest that working memory and attention in aphasia do respond to treatment and that these treatments may positively affect the language abilities of these individuals. Investigators have found improved word/non-word repetition abilities, sentence repetition abilities and reading comprehension as a result of training programs addressing cognitive processes (Kalinyak-Fliszar, Kohen & Martin, 2011; Koenig-Bruhin & Studer-Eichenberger, 2007; Sinotte & Coelho, 2007).

Murray, Keeton and Karcher (2006) used Attention Processing Training 2 (APT-2; Sohlberg, Johnson, Paule, Raskin & Mateer, 2001), a program designed to address attentional deficits for persons with mild cognitive dysfunction, as a treatment for mild aphasia. Daily communication abilities, attention and memory functioning were measured as outcomes of treatment. A single participant completed one hour/week in lab and one hour/week at home for a total of 50 hours of training. The training was found to enhance specific attention skills but positive changes in broader attention and untrained functions were less likely, which is consistent with the finding of Sturm et al. (1997). Neither the patient nor his wife reported any noticeable improvements in attention. These results suggest that although generalized changes in functioning may not occur from treatment, specific changes did occur in performance on trained attention tasks. The authors speculate that improvements in functioning may be observed if this training is made meaningful to the subject. This study also highlights the importance of considering practical and daily functioning outcomes of treatment, not just statistically measurable improvements on a standardized measure.

Investigators have also examined attention training and its role in treating reading comprehension impairments in mild aphasia. Coelho (2005) investigated an 8-week program of treatment using APT-2 in a single subject. In this study, Coelho found that the subject's reading comprehension improved and that the subject reported less effort required for reading tasks. These treatment effects mirrored the subject's improved ability to cope with distractions and maintain attention. The author attributed this change to an improvement in ability to sustain attention and to concentrate, which supports the idea that improvements in non-linguistic cognitive abilities lead to subsequent improvements in language.

In a follow-up study to Coelho (2005), Sinotte and Coelho (2007) also measured reading rate and comprehension before and after treatment based on the APT-2. The subject was an individual with mild anomic aphasia following a left hemisphere middle cerebral artery stroke. Post-treatment, the subject did not improve in reading rate and no change in comprehension accuracy was found. However, her overall reading comprehension accuracy stabilized, indicated by a decrease in variability of comprehension of longer complex reading passages. Additionally, the subject's aphasia quotient improved by a near clinically significant amount (4.7 points) and gains were noted on reading rate, accuracy and fluency on the *Gray Oral Reading Tests-4* (*GORT-4*). The subject also improved on 7 out of 10 subtests on the *Test of Everyday Attention* measure. The subject also reported functional attention skill improvement, which suggests that her improved attention abilities led to improved language processing in general. She also reported a better quality of life due to her finding that reading was less of a chore. Although the participant in this study did not improve reading comprehension as reported in Coelho (2005), this study nonetheless provides support that cognitive intervention can lead to an improved quality of life with regard to language activities. Importantly, this study also provides evidence that, for some individuals with mild forms of aphasia, improvements in language proficiency measures can occur without direct language intervention.

Studies have also been done that provide evidence for a directly measureable link between working memory intervention and language improvements. Koenig-Bruhin and Studer-Eichenberger (2007) sought to improve verbal short-term memory by examining repetition treatment for a single participant with conduction aphasia. The therapy consisted of repeating compound nouns and sentences in length from four to seven words, first immediately then with increasing delay. The subject's sentence repetition improved significantly, as well as the length of the subject's oral production of sentences in a picture description task (6.41 compared to 5.26 words pre-therapy), suggesting the improvement in verbal short-term memory, an underlying cognitive mechanism, led to language improvements.

Recently, a treatment study addressing working memory in a single person with aphasia used a word and non-word repetition protocol to strengthen semantic and phonological representations (Kalinyak-Fliszar et al., 2011). The single participant study involved repeating words and non-words with a progressively longer delay between repetitions. Results indicated that the participant improved her repetition in both words and non-words and maintained these improvements at follow-up. Comparison of pre- and post- test data suggest that generalization to subtasks such as rhyming and word-pair repetition also occurred. These findings suggest that working memory in at least some individuals with aphasia is responsive to training and that this training can improve language functioning.

To summarize, the studies reviewed have shown that many individuals with aphasia have trouble with working memory and attention. Additionally, these deficits affect language functioning, as indicated by significant correlations between measures of these cognitive factors and language ability (Murray, 2012a). Importantly, studies investigating the treatment of cognitive factors have shown promising results as a way to improve language functioning. A variety of language domains such as word and non-word repetition abilities, sentence repetition abilities and reading abilities have been targeted using cognitive-based treatment and promising language improvements have been found (Kalinyak-Fliszar et al., 2011; Koenig-Bruhin and Studer-Eichenberger, 2007; Sinotte & Coelho, 2007). Although the majority of the research in this area is single case studies, this research provides some evidence that treating working memory and attention could improve auditory language comprehension in aphasia.

Importance of Intensity of Therapy

Recent reviews on intensity of aphasia therapy have suggested that the more intensive a therapy is, the better the outcome, maximizing the brain's potential for neuroplasticity (Bhogal, Teasell & Speechley, 2003; Brady, Kelly, Godwin & Enderby, 2012; Cherney, Patterson, Raymer, Frymark & Schooling, 2008; Cherney, Patterson, Raymer, Frymark & Schooling, 2010). Bhogal and colleagues (2003) conducted a comprehensive search of five electronic databases, analyzing the data on aphasia therapy. By examining eight different studies on aphasia, Bhogal et al. found that there appeared to be a significant effect of the intensity of treatment. Of the eight studies, four showed a positive outcome. These studies had an average of 8.8 hours of therapy per week for 11.2 weeks. The four studies that did not show positive outcome only provided approximately 2 hours per week for 22.9 weeks. Bhogal et al. noted that both the greater number of hours of therapy and the greater intensity in number of hours per week were correlated with greater improvement on the Porch Index of Communicative Ability (PICA), the test most commonly used as an outcome measure in the studies included for analysis. This summary provides evidence that outcomes for clients are more positive when the treatment is provided in a more intensive format.

A more recent meta-analysis by Cherney and colleagues (2008) analyzed 10 studies examining constraint-induced language therapy (CILT) and intensity in treatment. Using five articles that directly addressed the role of intensity, Cherney et al. found that there was modest evidence for positive effects of more intensive treatment (various definitions for intensity depending on study, inclusion criteria for assessment was studies including 'intensity' or 'amount of treatment' in the abstract) for chronic aphasia. Language impairment outcome measures favoured more intensive treatment.

Outcome measures for community activity/participation were more mixed, with five subjects favouring more intensive therapy and four favouring less intensive. Only one study addressing acute aphasia was summarized that evaluated high-intensity treatment (Denes, Perazzolo, Piani & Piccione, 1996). In this study, investigators compared the efficacy of treatment sessions six or seven times a week to sessions three times a week in improving the language of a group of 17 individuals with global aphasia. Treatment was given by a trained speech-language pathologist using a stimulation approach with the ultimate goal being to improve language in a conversational setting. All 17 participants' outcomes in this study favoured high-intensity treatment, as measured using the Aachen Aphasia Test. Maintenance effects of intensity were mixed. Although this review is preliminary and more research is required for a more definitive answer, these reviews highlight the importance of examining the use of intensive treatment for improving communication outcomes for individuals with aphasia. However, if further research found that more intensive treatment was the best option for individuals with aphasia, this type of intensive therapy is often not available to patients, due to a lack of health care resources and heavy speech-language pathologist (SLP) caseloads.

Computer-based Training and Aphasia

Computer-based training that is easily accessible and that can be established as a valid treatment option for aphasia has important implications for the availability of intensive treatment options for individuals with aphasia. Clinical guidelines provided by the Royal College of Speech & Language Therapists state that "computer-based therapy offers the potential to provide intensive home-based therapy with minimal clinician input [and that] improvements in performance over a number of communicative modalities can occur" (Taylor-Goh, 2005, p. 110). SLPs can design and monitor these computer-based interventions, allowing the client access to more intensive treatment without significantly increasing the amount of time required by the SLP and without decreasing the quality of treatment the client receives. Thus, computer-based programs offer a potential solution to the problem of availability of resources.

Intensity of treatment is known to be an important factor in aphasia treatment outcomes and computer-based treatment programs offer a promising option (Bhogal et al., 2003; Taylor-Goh, 2005; Meinzer, Djundja, Barthel, Elbert & Rockstroh, 2005). There have been a number of studies investigating computer-based treatments for improving speech/language functioning in individuals with aphasia. Katz and Wertz (1997) designed a computer program to test the efficacy of computer-based reading treatment or computer-based non-language treatment (colours, movement and shapes used to focus on reaction time, memory and attention) compared to no-treatment. The two test treatments were compared to see if any improvements were due to specific language stimulation or to stimulation of basic cognitive constructs. Statistically significant improvement was found on five language measures for the computer reading group. Additionally, clinical significance was found on 10 out of 13 measures for the language treatment condition and on three measures for the computer treatment condition. These results demonstrate the efficacy of computerized treatment for chronic aphasia patients and that this treatment can be administered with minimal clinician assistance. The improvements found in the non-linguistic computer treatment condition provide support for these types of methods in treatment. Although subjects did not improve as much in the non-language-directed condition as in the language-stimulation condition, the tasks in the non-language condition targeted more generalized cognitive processes. Targeting these general processes may be less effective in treatment than more specialized processes. Cognitively-based computer treatment may still be an efficacious treatment technique in individuals with aphasia, especially if the treatment program was more specialized for certain types of cognitive processes (i.e. attention and working memory) and used linguistically-based stimuli.

A more recent study of computer-based intervention for aphasia focused on script training (Cherney, Halper & Kaye, 2011). Script training focuses on conversation practice oriented towards improving ability to function in daily activities. The authors developed their own program, *AphasiaScripts*, which provides “realistic conversational context for repetitive practice” (p. 493). Three participants with chronic aphasia practiced three individualized scripts over a period of nine weeks. All participants demonstrated aphasia after a left hemisphere stroke. 15 of the participants presented with nonfluent aphasia and 8 presented with fluent aphasia. After the nine weeks, a post-intervention interview was conducted with each participant to evaluate patient reported outcomes. The individuals with aphasia and the speech-language pathologist worked together prior to treatment to develop a meaningful script for each participant which the therapist then programmed into *AphasiaScripts*. Participants controlled how often and when they chose to complete their therapy sessions but were asked to practice at least 30 minutes per day, six days a week.

Post-treatment interviews were conducted using open-ended questions. Comments were coded into 10 themes, 5 related to communication behaviours of the participants and 5 related to the computer program and study procedures. The interviews were conducted with the person with aphasia and a spouse or other significant person. Twenty of the twenty-three participants stated that they noticed improvements in verbal communication over the course of the study, including initiation of more conversations and producing more words. Fifteen of the participants also noticed an improvement in self-confidence and attitude as a result of the improvements. Of the participants, 22 out of 23 stated that they were satisfied with the computer program. This study provides more evidence for the efficacy of self-directed computer programs in the treatment of aphasia. The positive perceived changes in subjects also strengthen the support for their use in a variety of aphasia treatments.

The above studies have provided support for the efficacy of computerized, client-directed treatment. Specifically, they found that computerized reading treatment and computer-based script training lead to improvements in verbal communication and quality of life for individuals with aphasia (Cherney et al., 2011; Katz & Wertz, 1997). Additionally, the positive view patients appear to have about this type of treatment could be motivating in its continued use (Cherney et al., 2011). However, the computer programs used were all designed by the research teams for their specific study and they were designed to address specific linguistic symptoms, as opposed to underlying cognitive impairments. Programs available to the wider public would be promising for increasing treatment availability, as clients would be able to access these programs outside of a clinic setting.

Commercially Available Computer-Based Treatment

Commercially-available programs for aphasia treatment would allow for wider access to computerized treatment resources for clinicians and clients. Despite the effectiveness of computer-based treatment shown by the studies above, there have been no studies directly examining the efficacy of commercially available programs targeting aphasia. There have, however, been a number of studies examining the effect of cognitive-based commercial computer therapy with other populations, such as healthy aging individuals and individuals with mild cognitive impairment.

BrainFitness Treatment in Healthy Elders. The present study examines the commercially available computer program *BrainFitness* (*BF*), which was originally designed to improve auditory processing speed, attention and working memory in healthy aging individuals. This program has been found to improve cognition in typically aging adults and is based on the principles of neuroplasticity, the idea that the brain can change due to experience. Smith et al. (2009) conducted a randomized controlled double-blind trial on the effects of *BF* on memory and attention in 487 healthy elders as part of the large Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. The active control condition consisted of watching digital video disk educational programs on history, art and literature to ensure participants would view the condition as a valid possible experimental condition. After 40 hours of intervention over a period of 8 weeks, participants in the experimental group showed significantly greater scores ($p < .05$) than the active control group on the *Repeatable Battery for* *the Assessment of Neuropsychological Status (RBANS)* subtests that use the auditory modality. These results suggest that the cognitive training program *BF* can create benefits that generalize to untrained measures of attention and memory. Also, these gains in attention and memory were larger for the *BF* program, targeting specific cognitive domains than for the program of general cognitive stimulation that was used in the active control condition.

In another study, Willis et al. (2006) examined long-term effects of cognitive training, such as that used in *BF*, in older adults. In a follow-up to a randomized controlled single-blind trial, Willis et al. found that cognitive improvements as a result of specific cognitive training were maintained up to 5 years post intervention. The study included three interventions, targeting the specific areas of memory, reasoning and speed of processing. Intervention involved 10 training sessions followed by booster training at 11 and 35 months post treatment. Although all groups reported improvements, only the reasoning condition showed a significant effect size post treatment. This indicates that only individuals in the reasoning condition reported significantly less difficulty in daily activities as a result of treatment. These results suggest, similar to Smith et al. (2009), that cognitive intervention can reduce age-related declines. This study also suggests that these results can be maintained well past treatment, especially if a booster intervention is applied.

BrainFitness Treatment in Individuals with Mild Cognitive Impairment. Recently, *BF* has also been examined as a treatment tool for adults with mild cognitive impairment (MCI). Barnes et al. (2009) completed a pilot study examining the effect of computer training on 47 subjects diagnosed with MCI. The intervention group performed exercises to improve auditory processing speed and accuracy while the control group performed passive computer activities (e.g. listening to audio books, reading online newspapers). Although the intervention group's total scores on the primary outcome, the *RBANS,* improved by a larger amount than the control group's (0.36 SD improvement from previous scores compared to 0.03 SD), this result was not significantly different between the groups (p=0.26). The effect size for verbal learning and memory measures was larger between pre- and post-conditions in the intervention group, while the effect size for the language and visuospatial measures was larger in the control group. These results suggest domain-specific effects. They also support the idea that computer training targeting cognitive functions may be useful for subjects with MCI.

BrainFitness has been found to improve cognition in typically aging adults and individuals with MCI. The cognitive improvements were found to generalize to untrained measures of attention and memory and in typically aging adults, gains were found to be maintained 5 years post-treatment (Smith et al., 2009; Willis et al., 2006). Auditory processing and speech accuracy improvements were also found for individuals with mild cognitive impairments as a result of intervention (Barnes et al., 2009). This research shows the efficacy of *BF* in addressing impairments in working memory and attention, areas which individuals with aphasia are known to have difficulty. These results, combined with the fact that computer-based interventions been found to be a good therapy option for individuals with aphasia, provide support for research into the use of this program for individuals with aphasia. Many clinicians have begun recommending commercially-available cognitive programs to their clients (Rabipour & Raz, 2012). However, to date, there have been no studies examining the efficacy of computer-based cognitive therapies for individuals with aphasia.

Research Question

The purpose of this study was to examine the efficacy of a commercially available cognitive training program (*BrainFitness*) as an intervention technique for individuals with aphasia. Using an exploratory single-subject case study design, the following research question was addressed:

- Does *BrainFitness* training improve language and cognitive functioning of an adult with stroke-induced aphasia?

Hypothesis: BrainFitness training will result in improved scores on measures of language and cognitive functioning in an adult with stroke-induced aphasia. The connection between cognitive functioning and language abilities in individuals with aphasia has been supported in many studies (Murray, 2012a; Friedmann & Gvion, 2003; Gvion & Friedmann, 2012). Additionally, research supports the idea that language can be improved following treatment of underlying cognitive deficits (Kalinyak-Fliszar et al., 2011; Murray et al., 2006; Sinotte & Coelho, 2007; Koenig-Bruhin & Studer-Eichenberger, 2007). Unlike previous research with individuals with aphasia, this study will use a commercially available cognitive training program (*BrainFitness*). As there is a lack of evidence on the efficacy of this type of training for individuals with acquired neurological disorders such as aphasia, this project will help us address this knowledge gap.

Method

Participant

R.D. was a 67-year-old right-handed male, who suffered a left-hemisphere stroke at the age of 63. He presented with aphasia and mild apraxia of speech as the result of an infarct in the left middle cerebral artery. As seen in figure 1, MRI scans revealed a left perisylvian lesion which covered cortex and sub-adjacent white matter primarily in the anterior portions of the middle temporal gyrus and the superior temporal gyrus, as well as the insula and subcortical structures of the brain (thalamus and basal ganglia).

Figure 1: Representative axial slice illustrating R.D.'s lesion

R.D. had no history of learning disabilities or neurological disorders prior to his stroke, and spoke English as his primary language. Although he had mild right hemiparesis, consistent with a left hemisphere stroke, he had no other presenting symptoms that interfered with treatment. He was very motivated and disciplined to participate in treatment; because a large aspect of this program involved daily use of the computer program in a home environment, this was an important factor. He passed screenings for vision (reading 18 point font text at a distance of 30 cm), hearing (pure tone hearing thresholds: 500, 1000, 2000 and 4000 Hz bilaterally at 30 dB) and an informal measure of manual dexterity and proficiency with computers (i.e., if he was able to navigate the computer mouse to open/close the program and select relevant targets). R.D. had 9 years of schooling and worked prior to retirement as a janitor. He had previously participated in group therapy, as well as an earlier version of the *BrainFitness* training program two years previous to this study. During the course of this study, R.D. did not participate in any additional treatment.

Design

A single-subject ABA design, using multiple baseline and generalization measures was employed. Behavioural measures were assessed at baseline, pre-treatment, post-treatment and follow-up time points (although neuroimaging measures were also assessed concurrently, only behavioural measures will be reported on in this study). A no-treatment baseline phase of four weeks was instituted prior to beginning the eight week treatment phase (Figure 2).

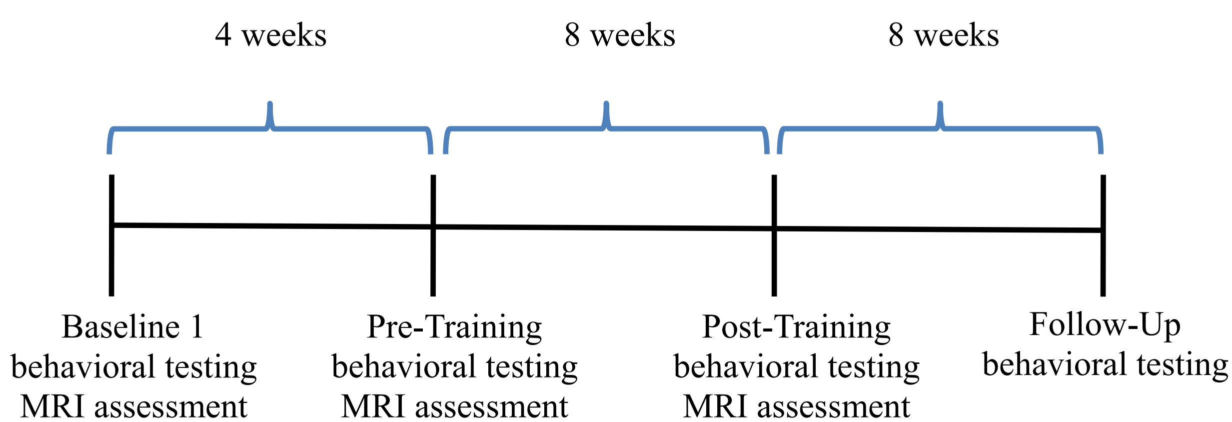

Figure 2: Treatment design

R.D. completed eight weeks of treatment after which post-training measures were administered. Finally, after another period of eight weeks, follow-up measures were conducted as a measure of maintenance. Over the course of intervention, we expected no changes in behavior between baseline and pre-test (demonstrating a stable baseline), changes from pre-test to post-test (demonstrating a treatment effect), and no change from post-test to follow-up (demonstrating maintenance of treatment gains).

Outcome Measures

Assessment of R.D.'s cognition and language were carried out at each of the four time points using a variety of measures. These measures were administered over a period of two to three days. The outcome measures used in this study are listed below (Table 1).

Table 1: Language and cognitive outcome measures and general descriptions

Outcome Measure	**Description**
Direct Treatment Measure	
BrainFitness Program	% improvement on each of the trained modules
Generalization Measures	
Western Aphasia Battery-Revised (*WAB-R*; Kertesz, 2006)	General measure of spoken language ability
Alberta Language Function Assessment Battery (*ALFAB*; Westbury & Sheldon, 2005)	Computerized battery of language assessment (Subtests: Phoneme Discrimination, Auditory Lexical Decision and Semantic Plausibility Judgment)
n-back test (Leung & Alain, 2011)	Cognitive measure of working memory decisions on continuous stream of auditory stimuli
Tompkins et al. (1994) modified span tasks	Measure of subject's ability to simultaneously retain and make judgments about auditory signal
Test of Everyday Attention (*TEA*; Roberson, Ward, Ridgeway & Nimmo-Smith, 1994)	Cognitive measure of auditory working memory and attention

Language Measures. The Aphasia Quotient from the *Western Aphasia Battery- Revised* (*WAB-R*; Kertesz, 2006) was used as a measure of R.D.'s spoken language ability. The *WAB-R* comprises a series of subtests designed to distinguish the presence, degree and type of aphasia. The subtests include measures of spontaneous speech production, auditory verbal comprehension, repetition and naming abilities. Following Katz and Wertz (1997), a clinically significant change on the Western Aphasia Battery was defined as an improvement of at least 5 points on the *WAB-R* Aphasia Quotient.

The *Alberta Language Function Assessment Battery* (*ALFAB*; Westbury & Sheldon, 2005) is a comprehensive computerized battery of tests designed to assess many areas of language. For the present study, the participant completed the auditory-based subtests of Phoneme Discrimination, Auditory Lexical Decision and Semantic Plausibility Judgment in order to assess language processing abilities at the phoneme, word and sentence levels, respectively. In the Phoneme Discrimination task, the participant heard pairs of words/non-words that differed by a single phoneme (e.g. *mop/mob)*, and indicated whether they were the same or different by a keyboard button press. There were 90 pairs of words or non-words presented in this subtest. The Auditory Lexical Decision task required the subject to indicate as quickly but as accurately as possible if a stimulus was a word (*c* key) or a non-word (*x* key). There were 145 items presented during this subtest. The Semantic Plausibility Judgment task required the subject to press the *c* key if a sentence was sensible and plausible (e.g. ‘A bear is a hairy mammal) and the *x* key if the sentence was either senseless or implausible (e.g. ‘He laughed out my joke’). There were 80 sentences presented in this subtest. Measures of reaction time (RT) and accuracy were obtained.

In addition to the *ALFAB* measures administered during behavioural assessment sessions, a subset of 60 words and 60 sentences were used to measure auditory lexical decision and semantic plausibility judgment abilities during neuroimaging assessment. Behavioural accuracy data from these two tasks were also examined as additional measures of response to treatment.

Finally, auditory lexical decision tasks were administered weekly as probes of language processing. Each probe was constructed using 30 words and 30 non-words. Words were regularly- and exceptionally-spelled words matched for length, frequency, phonologic and orthographic neighbourhood density. Non-words were constructed by changing one letter in the word to construct a non-word; pseudohomophones (i.e. non-words that sound like real words) were excluded.

Cognitive Measures. Changes in cognitive processing were assessed using two measures of working memory capacity and one test assessing attention. The working memory span task developed by Tompkins, Bloise, Timko and Baumgaertner (1994), adapted from Daneman and Carpenter (1980), requires the participant to judge the semantic accuracy of a set of sentences, by responding true or false, and to remember the last word of each sentence heard. Sentences ranged in length between three and five words. R.D. heard sentences in sets, and then was required to additionally repeat the last word of each sentence heard. The task began with two sentences in each set and increased in number as the participant successfully completed a level. This task required the participant to simultaneously process and store incoming information and yielded scores reflecting accuracy of semantic judgments and retention of sentence-final words.

The participant completed an auditory version of the *n*-back task designed by Leung and Alain (2011) as another measure of working memory through an auditory modality. The *n*-back is a test that requires the participant to judge from a continuous stream of stimuli whether a current stimulus matches one that occurred *n* places back in a sequence. Performance on this measure is calculated using a sensitivity index, the d prime statistic (d') (Haatveit et al., 2010). The d' statistic is used to compare a participant's success in terms of his/her Hit Rate, the proportion of successful responses when a signal is present, and his/her False Alarm Rate, the proportion of incorrect responses when a signal is absent. It measures the participant's sensitivity to the desired response. If a participant has 50% accuracy, the d' value is 0, indicating performance is no better than chance. A positive d' score indicates a better than chance performance; therefore, a higher d' indicates better performance on the measure.

The *Test of Everyday Attention* (*TEA*; Roberson, Ward, Ridgeway & Nimmo-Smith, 1994) subtests of Elevator Counting and Elevator Counting with Reversal were used as a measure of cognitive functioning, examining auditory and attention processes. On the Elevator Counting subtest, the participant listened to a series of tones and counted how many tones they heard. Each tone represented a floor passed in an elevator. The participant mentally tracked the tones and responded which floor they had reached at the end of the sequence of tones. The Elevator Counting with Reversal task was very similar to the Elevator Counting task but included both a low tone that signaled when the elevator was going down and a high tone for when the elevator was going up. The subtests from the *Test of Everyday Attention* were chosen as another measure of cognition due to the fact that they measured sustained and selective attention using an auditory modality and that the *TEA* itself has been demonstrated to be reliable and valid in a number of studies (Chen, Koh, Hsieh & Hsueh, 2013). Additionally, it has been shown to have good test-retest reliability in patients with chronic stroke, even across time frames as short as one week. These *TEA* subtests were also found to be most strongly correlated with auditory language comprehension scores by Murray (2012a). By completing only these subtests, we were able to get a measure of the participant's auditory comprehension without risking fatigue by administering the entire 45-60 minute test.

Treatment

The treatment program used in this study was *BrainFitness*, a subset of six auditory-based modules from the BrainHQ cognitive training program developed by Posit Science™. These modules are designed to improve processing, speed, attention and working memory in the auditory modality. The program is designed for participants to participate in intervention for a period of 8 weeks, for a total of approximately 40 hours of training (1 training set/day x 5 days/week x 8 weeks). The program is designed to be completed independently by the participant.

Modules. Each module responds to the user's performance by increasing in difficulty as the user progresses. Each module is designed to address one of the main skill areas (auditory processing, speed of processing, attention or memory) at different difficulty levels. The modules are completed in a specific order, starting with simple frequency sweeps and moving up to short clips of a conversation. Within each module, there are a specific number of levels of training for the participant to complete. Table 2 describes each of the six modules in the order that they were completed and the cognitive domain that each was designed to target.

Table 2: BrainFitness Modules

Module	Description	Domain Targeted	Number of Stages
Sound Sweeps	Participant listens to rising/falling tones and identifies whether they go up or down. As the level increases, the sweeps speed up.	Auditory processing speed	Twenty-one
Fine Tuning	Participant listens to a syllable and chooses which of two visually-presented syllables they heard (e.g. 'boo' and 'do'). As the levels increase, the syllables are digitally altered to decrease presentation time making the sounds harder to distinguish from one another.	Auditory perception and processing speed	Twenty-one
Memory Grid	Participant matches pairs of cards representing aurally-presented syllables. As the level increases, there are more sets of cards to match.	Memory	Twenty
Syllable Stacks	Participant listens to a series of syllables and clicks on tiles	Memory	Twenty

	representing each syllable in order. As the level increases, the participant is required to remember more syllables.		
To-Do List	Participant hears a set of instructions and is required to select tiles representing the instructions in the order indicated (e.g. “choose Bag A, then choose the hammer”). As the level increases, the number of instructions given also increases.	Working memory	Eighteen
In-the-know	Participant listens to a conversation then answers questions based on what they heard. As the level increases, conversation length and complexity of content increases.	Language in conversational context	Twenty

Procedure. BF was accessed through the BrainHQ website. The participant was given a personal log-in and password to access his specific training program. The program kept a record of daily progress through each stage for each of the modules. Three modules were completed each day such that the participant advanced to the next stage every second day of training. The Sound Sweeps and Fine Tuning modules had 21 stages, Memory Grid, Syllable Stacks and In-the-Know had 20 stages and To-Do List had 18 stages. The participant was expected to complete one stage in three of the six modules per day to achieve the expected 40 hours of training. For each new stage of a module, the participant was required to complete a baseline measure of pre-test ability and the passing criterion for that stage was determined in reference to this baseline ability. The participant was expected to either perform higher than his baseline or to complete a set number of items if baseline performance could not be passed.

The first session was conducted in the participant’s home environment, assisted by the SLP student. The SLP student explained the procedure and ensured that the participant was comfortable with the computer hardware and software. The participant then completed the remaining sessions on his own, using a laptop provided by the researchers. It was possible to track the participant's progress through access to an online researcher portal on the *BF* program website. This allowed the clinician to monitor treatment compliance and whether the participant was making any gains or progress in the levels of *BF*. The participant also met the clinician four times throughout the treatment process to ensure there were no concerns about the training procedure. During three of these four meetings, probes on the Auditory Lexical Decision task were administered as a measure of treatment progress.

Treatment Frequency. Treatment was conducted over an eight week period, during which R.D. completed approximately five hours a week of independent, online training using the *BrainFitness* program.

Post-treatment/ Follow-up Assessment. Post-treatment assessment measures were taken one week after completing the online *BrainFitness* training program. At this time, each of the pre-treatment language tests (*WAB-R* and *ALFAB*), as well as the cognitive measures (*TEA,* Tompkins et al. working memory measure and *n*-back) were conducted again to test for any improvements. Follow-up measures were also conducted eight weeks following the post-treatment measures. At this time point, due to client availability and time restrictions, only the language generalization measures (*WAB-R* and *ALFAB*) were conducted. These measures were chosen to re-administer as they were shown to be the most sensitive to small changes in R.D.'s accuracy and reaction time abilities at post-treatment. Although R.D. was not participating in any intervention during this eight week period, it was predicted that he would maintain any gains found post-treatment and that there would be minimal change between post and follow-up time point measures.

Results

Direct Treatment Effects: BrainFitness Training Program

R.D. was able to complete the training program in the expected 8 week time period. Percent improvement values were not available for each individual module; rather, only a general ability score was available. R.D. improved his overall ability on the *BF* program by 15.2%. This was calculated by comparing the mean percentage change between baseline ability and the all-time best scores for all levels done by the user. After completing all the levels of each stage, R.D. completed 536 levels total. Table 3 below displays the total levels completed and time trained on each of the six modules of *BF*.

Table 3: Time Trained on Brain Fitness Modules

Module	Time Trained (in hours)
Sound Sweeps	6.66 hours
Fine Tuning	5.75 hours
Memory Grid	7.32 hours
Syllable Stacks	4.34 hours
To-Do List	3.78 hours
In-the-know	5.49 hours
Total	33.33 hours

Time spent on a task was determined by the number of stages for each test and by how long the participant took to respond. As the table shows, the most amount of time was spent on the Memory Grid task (7.32 hours). The least amount of time was spent on the To-Do List task (3.78 hours). This can be explained by the fact that the To-Do List task only had 18 stages to complete compared to the 21 and 20 stages for the other modules. Total time trained amounted to 33.33 hours. This is lower than the 40 hour approximation as some days R.D. was able to complete the required daily training stages in less than an hour. Certain modules (e.g. sound sweeps) had more levels required for completion to account for the fact that these levels required less time to complete.

Generalization Effects: Language Measures

WAB-R. The *WAB-R* (Kertesz, 2006) was administered at baseline, pre-test, post-test and follow-up as a measure of R.D.'s spoken language ability. Changes in behaviour were examined between baseline and pre-test, between pre-test and post-test and between post-test and follow-up (Table 4).

Table 4: Western Aphasia Battery- Revised Subtest Scores and Aphasia Quotient

Subtest	Baseline (JAN/13)	Pre-test (MAR/13)	Post-test (MAY/13)	Follow-up (JULY/13)
Aphasia Quotient	74.4 (Conduction)	76.6 (Conduction)	81.4 (Conduction)	83.1 (Conduction)
Content Score	9	8	9	9
Fluency Total	6	8	8	9
Comprehension Total	9.7	8.7	9.4	9.45
Repetition Total	5.8	6.9	6.4	6.1
Naming Total	6.7	6.7	7.9	8

At baseline, the participant presented with an aphasia quotient (AQ) of 74.4 (out of a possible 100), and a profile of conduction aphasia, characterized by repetition difficulties, good auditory comprehension and relatively fluent output. At pre-treatment, the participant presented with the same profile of conduction aphasia, and an aphasia quotient of 76.6. Although the subtest scores varied slightly between baseline and pre-test, this was not a clinically significant difference (Katz & Wertz, 1997). Fluency increased slightly, as did repetition, while comprehension total and content decreased slightly. Overall, this resulted in a similar AQ between the two time periods.

Between pre-test and post-test, R.D.'s AQ showed a much larger increase than between baseline and pre-test. R.D.'s AQ increased from 76.6 to 81.4, 4.8 points which is a near-clinically significant increase. Between baseline and post-test, R.D. demonstrated a clinically significant increase of 7 points, improving from 74.4 to 81.4; this profile was still consistent with conduction aphasia. These results demonstrate that R.D. made overall language improvements as a result of treatment.

The *WAB-R* was conducted a fourth time at follow-up assessment to examine maintenance of the treatment effects. Between post-test and follow-up, R.D.'s AQ showed a slight increase from 81.4 to 83.1. This 1.7 increase was not clinically significant. This score increase was mainly due to a 1-point increase in fluency. As the above results show, R.D.'s score on the *WAB-R* was stable before and after treatment, with no significant change in AQ, and increased between pre- and post-treatment.

ALFAB. Measures for each subtest of the *ALFAB* were available for pre-treatment, post-treatment and follow-up sessions. There was no baseline measure available due to technical difficulties during administration. On the Phoneme Discrimination subtest, R.D.'s accuracy remained relatively stable across the course of the study, from 91.11% at the beginning of treatment to 92.20% at post-test and to 93.30% at follow-up (Figure 3). Visual analysis of this data suggested that this difference was not large enough to be significant across time points. Reaction time values, calculated from correct responses only, were adjusted to exclude outliers that were greater or less than two standard deviations from the mean. Removal of outliers resulted in removal of 4% of the items at pre-test, 8% at post-test and 7% at follow-up. RTs decreased from 1053.35 milliseconds (ms) to 898.35 ms from pre- to post-treatment. This decrease in latency suggests that R.D.'s ability to distinguish between phonemes became faster post-treatment. Notably, accuracy did not decrease with the decrease in reaction time, suggesting the reduction in reaction time was not due to a speed-accuracy trade-off. His reaction time also increased slightly between post-treatment and follow-up, suggesting that the improvement found between pre-test and post-test was not due to repeat administration of the subtest but rather to a treatment generalization effect. Data values for reaction time and accuracy are presented in Appendix D.

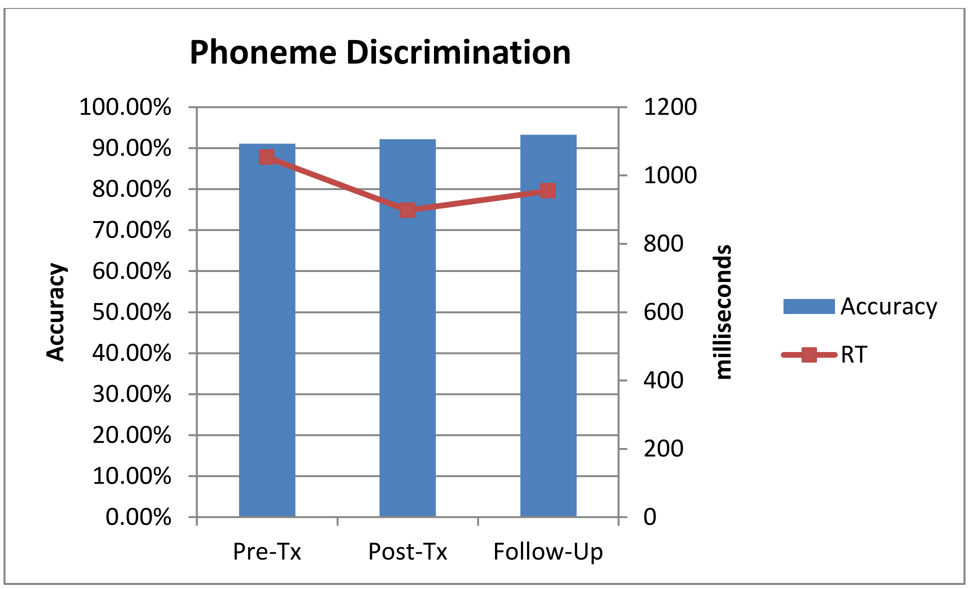

Figure 3: Phoneme discrimination accuracy values and average reaction time (RT) on the *ALFAB*

Cohen's *d* was calculated as a measure of the treatment effect size. Cohen's *d* is the most commonly used measure of effect size in single subject case designs and is especially well-suited for use with an ABA design due to the multiple time-point measures (Beeson & Robey, 2006).

Effect size was calculated by taking the mean reaction time at time point 1 and subtracting it from the mean reaction time at time point 2, then dividing this value by the pooled standard deviation of the two samples. Using the values proposed by Cohen, an effect size of 0.2 or less constitutes no effect, an effect size of 0.2-0.5 constitutes a small effect, 0.5-0.8 constitutes a medium effect and greater than 0.8 constitutes a large effect.

Between pre-test and post-test, effect size calculations revealed a medium effect of $d = -0.51$. This negative effect shows that between pre- and post-test, reaction time values decreased, suggesting a faster mean reaction time at post-test. Between post-test and follow-up, a small effect of $d = 0.33$ was observed. Notably, this effect size was positive, meaning there was an increase in reaction time between post-test and follow-up.

In addition to the overall performance reaction times, performance was also separated by lexical status (e.g. words versus non-words for Phoneme Discrimination subtest). For the Phoneme Discrimination sub-test, there did not appear to be a large difference in accuracy or reaction time when words and non-words were compared (Figure 4).

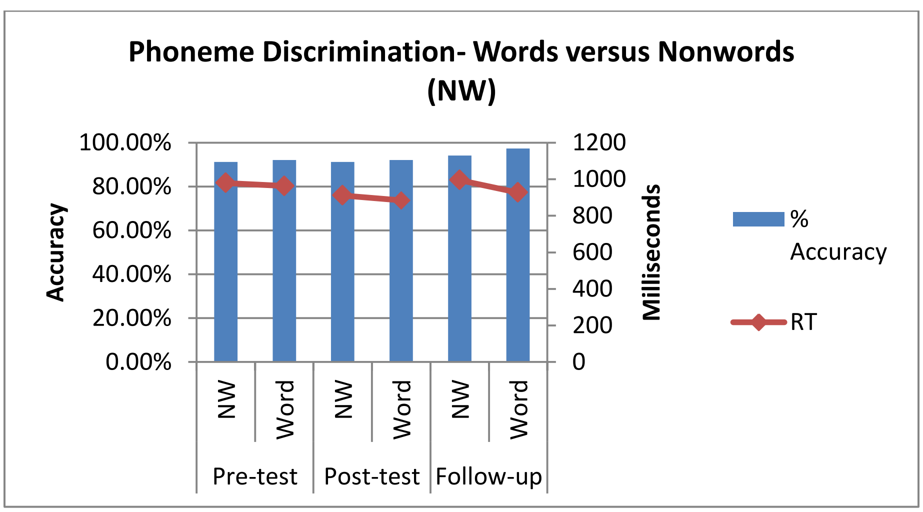

Figure 4: Reaction time and accuracy performance on words versus non-words Phoneme Discrimination Test

Effect sizes were also calculated to compare words versus non-words between time points. Both words and non-words showed a small effect size between pre-test and post-test mean reaction time values (d = -0.29 for words and d = -0.32 for non-words). Between post-test and follow-up, effect size for words and non-words were also small, (d = 0.22 and 0.35 for words and non-words, respectively. The positive direction of the effect reflects that reaction times became slower between post-test and follow-up, and that this change was larger for non-words.

The second subtest administered from the *ALFAB* was the Auditory Lexical Discrimination Subtest (Figure 5). Removal of outliers resulted in removal of 3% of the items at pre-test, 3% at post-test and 6% at follow-up. Percent accuracy decreased from 84.03% at pre-test to 79.86% at post-test, then held at approximately the same accuracy at follow-up with an accuracy percentage of 79.90%. Visual examination of the data revealed that the percent accuracy change was larger between pre-test and post-test than between post-test and follow-up. Data values for reaction time and accuracy on the Auditory Lexical Decision task are presented in Appendix E.

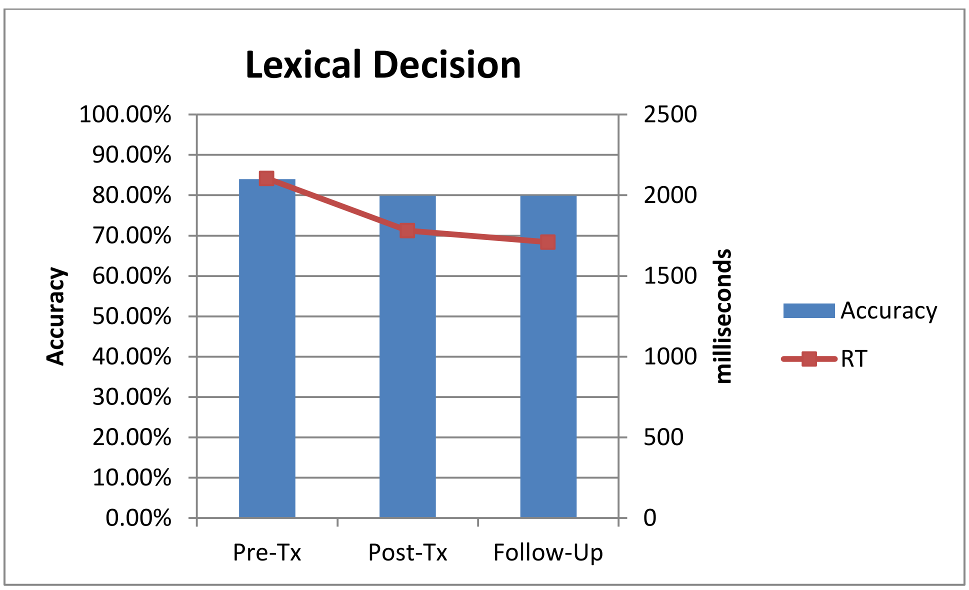

Figure 5: Auditory lexical decision task accuracy values and average reaction time (RT) on the *ALFAB*

At pre-test, mean reaction time was 2104.43 milliseconds (SD = 1154.05 ms). Reaction time decreased at post-test (M = 1780.14 ms, SD = 1145.96 ms) and remained relatively stable at follow-up (M = 1709.41 ms, SD = 902.63). Cohen's *d* effect size was used to examine the size of the difference between mean reaction times across time points. Between pre-test and post-test, effect size was calculated as *d*= -0.28, a small effect. Effect size between post-test and follow-up was *d*= -0.07, a null effect. These suggests a small change between pre-test and post-test in reaction time values and that this change remained constant between post-test and follow-up.

Reaction time and accuracy were also compared for words and for non-words on the Auditory Lexical Decision task (Figure 6). A difference in reaction time, as well as accuracy, was noted for words versus non-words for this task. Accuracy was higher and reaction times were faster overall for judgments made for words.

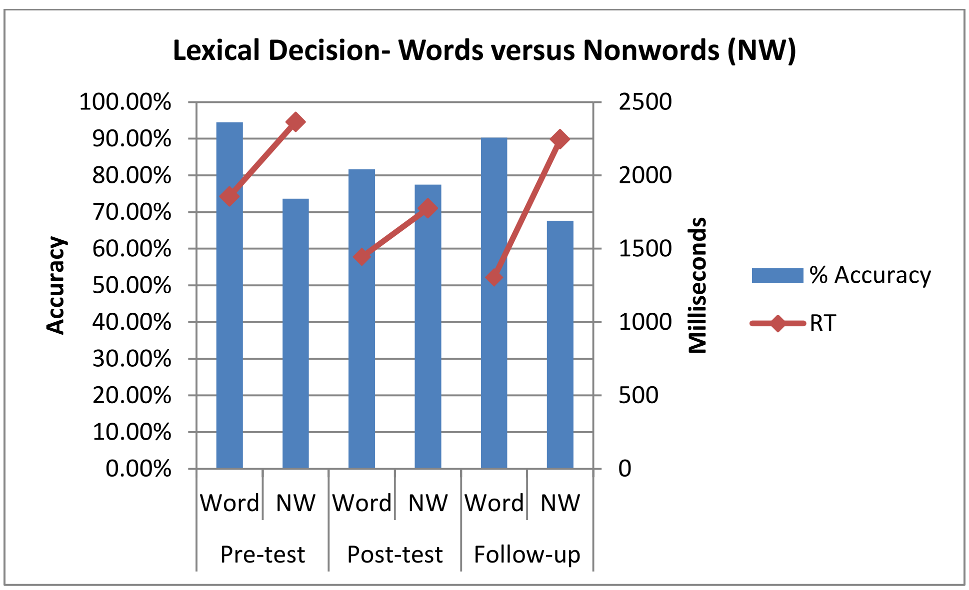

Figure 6: Reaction time and accuracy performance on words versus non-words (NW) Auditory Lexical Decision Task

Statistical analysis using a paired samples *t*-test revealed that reaction times were much slower for non-words than for words on this subtest ($t(147) = -4.499$, $p < 0.01$), a common pattern in this type of data (Ratclif, Gomez & McKoon, 2004). Accuracy was also lower across all time points for non-words, however the size of this difference appeared to decrease between pre-test and post-test. Effect size calculations were conducted to compare mean performance on words and non-words across the time points. Between pre-test and post-test, there was a small-medium effect size for words ($d = -0.45$) and a medium effect size for non-words ($d = -0.59$). Again, the negative effect size reflects the reaction times becoming faster across time points. The effect size for non-words was larger, reflecting the bigger decrease in reaction time for non-words. Between post-test and follow-up, reaction time continued to decrease for words ($d = -0.30$), while reaction time for non-words increased almost to pre-treatment levels ($d = 0.43$). These values suggest a difference in performance on words versus non-words across the treatment time points.

The third subtest used from the *ALFAB* as a measure of general language abilities was the Semantic Plausibility Judgment task (Figure 7). Again, performance was measured at pre-test, post-test and follow-up time periods. Removal of reaction time outliers resulted in removal of 3% of the items at pre-test, 3% at post-test and 5% at follow-up. Percent accuracy increased from pre-treatment to post-treatment, from 73.75% to 77.50%, and decreased at follow-up to 71.25%. Data values for reaction time and accuracy on the Semantic Plausibility Judgment task are presented in Appendix F.

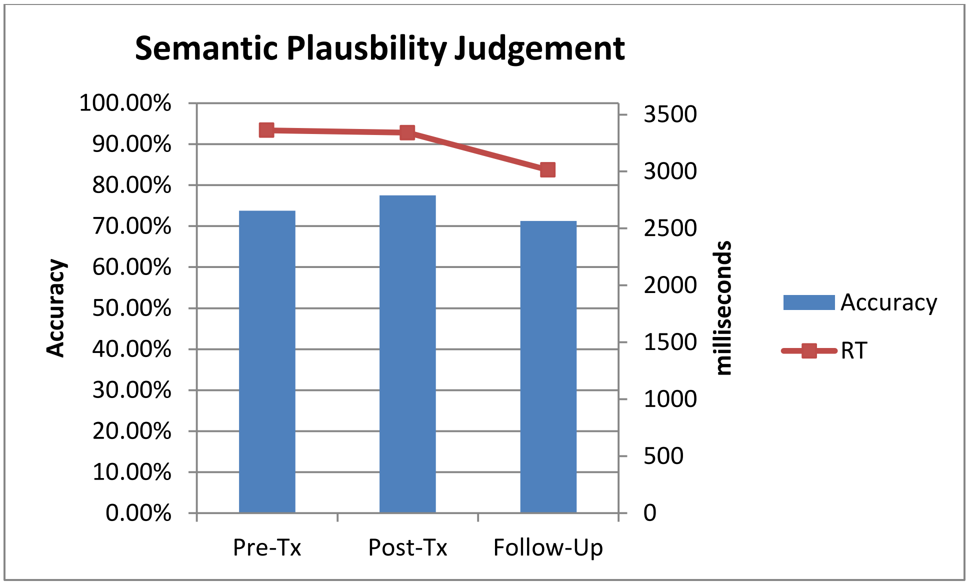

Figure 7: Semantic plausibility judgment task accuracy values and average reaction time (RT) on the *ALFAB*

Reaction time remained near-constant between pre-test (M = 3361.91 ms, SD = 1286.88 ms) and post-test (M = 3341.03 ms, SD = 1116.68 ms) and decreased at follow-up (M = 3013.15 ms, SD = 1044.49 ms). Effect size calculations for pre-test to post-test revealed a null effect (d = -0.02). Effect size between post-test and follow-up was small, suggesting a slight decrease in reaction time across time points (d = -0.30).

Reaction time and accuracy values are reported in Appendix D, Appendix E and Appendix F for each factor analyzed from each of the subtests. In addition to words versus non-words, concreteness, length and frequency were also compared in the Auditory Lexical Decision results. Statistical calculations could not be run due to the small *N*; instead, the data was analyzed visually. No strong patterns emerged overall, although high frequency words did appear to have faster reaction times overall than low frequency words (results are presented in Appendix E). The difference between these values is larger than the difference between high and low frequency words in the normed sample (1308.92 ms high frequency versus 1360.21 ms low frequency). On the Semantic Decision task, R.D. had slower reaction time on garden path sentences (i.e. sentences where the listener's most likely first interpretation is incorrect and they must reprocess to understand) and for sentences where the word in error was a preposition (e.g. "beside, across") (results are presented in Appendix F). This pattern is similar to those seen in the values in the normed sample for the ALFAB, with garden path and preposition sentence being the slowest. Accuracy was also lowest for preposition sentences (Figure 8). Due to the small sample size number of items per cell, additional analyses were not conducted.

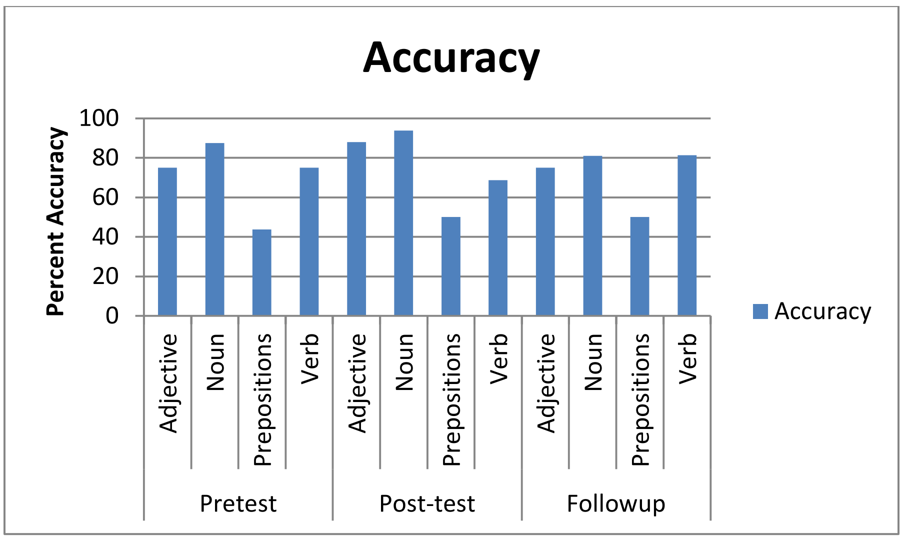

Figure 8: *ALFAB* Semantic Plausibility Judgement Subtest Accuracy Values

Lexical Decision Task Probes. Three different versions of a lexical decision task were administered to R.D. on three separate days as a measure of continuous change during the course of treatment (Figure 9). Measures of accuracy and latency (reaction time in milliseconds) were obtained. Accuracy varied over the course of the three weeks. At the first time point (3 weeks into treatment), accuracy was 76.67%, at the second 65.00% and at the third 75.00%. The difference between reaction times at time point 1 (M = 2350.18 ms), time point 2 (M = 2293.31 ms) and time point 3 (M = 2168.77) were null (d = -0.04 between time point 1 and time point 2, *d*= -0.09 between time point 2 and time point 3). These reaction time values were slower overall than the *ALFAB* lexical decision task scores found at pre-test, post-test or follow-up.

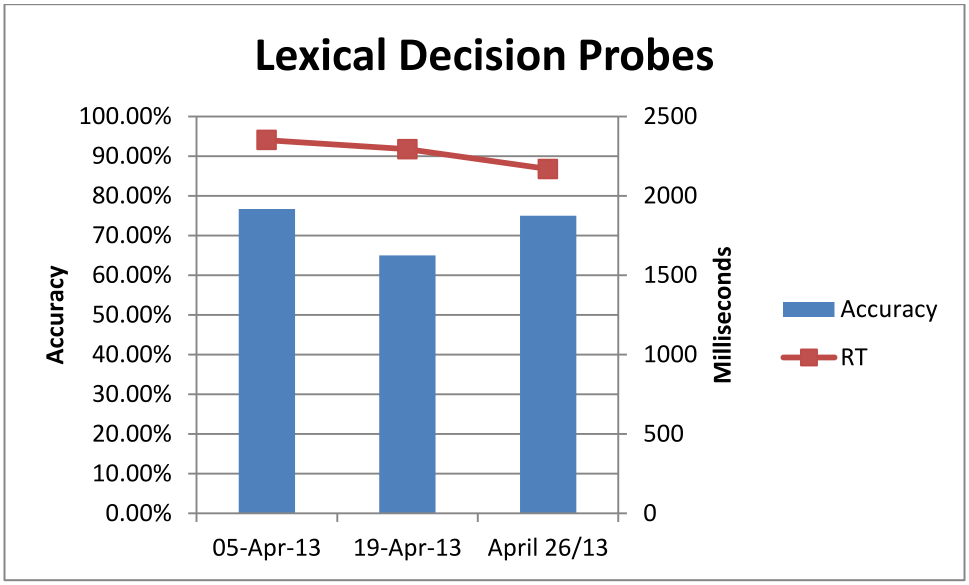

Figure 9: Lexical decision task probe accuracy values and average reaction time (RT)

MRI Behavioural Measures. A version of both the Auditory Lexical Decision task and the Semantic plausibility judgment task was administered to R.D. during fMRI scanning (Figure 10). Measures were taken at baseline, pre-test and post-test. Word identification accuracy remained stable between baseline and pre-test with 58.3% at baseline and 61.67% at pre-test, and then increased to 76.70% at post-test. Sentence identification performance was identical between baseline and pretest at 66.67% and then increased to 80.0% at post-test, suggesting a treatment effect on both identification of words versus non-words and identification of plausible versus implausible sentences. Reaction time data was not collected during these measures. This data supplements the results of the behavioural *ALFAB* measures and provide evidence of a stable baseline before treatment, as there was no baseline measure available for the behavioural *ALFAB* tasks.

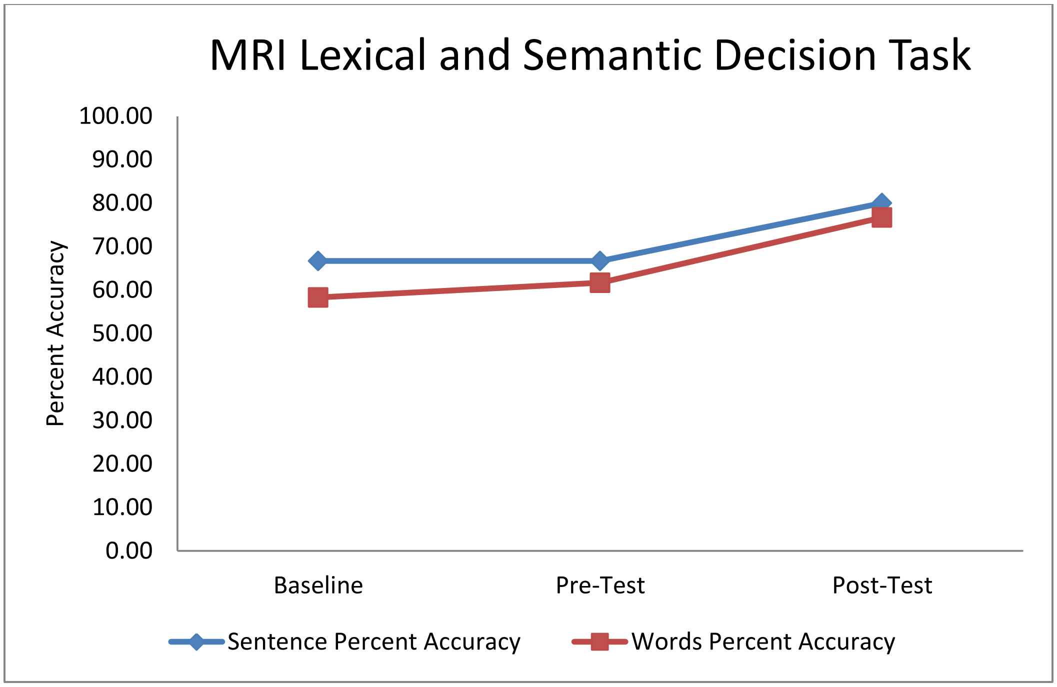

Figure 10: Lexical decision task (words) and Semantic plausibility judgment task (sentences) percent accuracy values from MRI scans

Generalization Effects: Cognitive Measures

Tompkins et al. Working Memory Measure. R.D. was able to judge accuracy (6/6 correct true/false judgments at all time points) but was only able to recall one word from the sentence pairs presented at baseline. At pre-treatment, he correctly remembered one pair of words from the sentence set and a single word from another sentence set. At post-test, R.D. could not correctly recall any pairs, only a single word from one of the pairs. Only Level 2 sets (two sentences at a time) were attempted as longer sets were too cognitively challenging.

n-back Auditory Working Memory Measure. This task required R.D. to monitor a continuous stream of letters/numbers and to make a response when the stimulus was the same as the one *n*-back. This measure was administered at baseline and at post-test. This measure could not be administered at pre-test or at follow-up due to computer program difficulties. The statistical calculation used for this task is the d prime (d') calculation, which gives the relative proportion of hits to false alarms (Haatveit et al., 2010). The version of the *n*-back given at baseline had only 150 items while the post-test *n*-back had 240 items, therefore analyses on the post-test measure were conducted in two ways: using all 240 items and using only the first 150 items. There appeared to be a length effect in that performance was worse at post-test when all 240 items were included. Due to this, post-test *n*-back scores were also calculated using only the first 150 items presented in the test. Both adjusted and un-adjusted *n*-back scores are reported (Figure 11).

The adjusted *n*-back scores showed that R.D. improved his performance post-test on the 2-back measure. R.D.'s score increased from 2.07 at baseline to 2.29 at post-test. This indicated a positive increase in performance and an increase in the proportion of hits to false alarms. The unadjusted score including all 240 items did not show this improvement, as R.D. obtained a score of 2.05 at post-test. R.D. decreased in performance on the 1-back measure, as shown in both the adjusted and unadjusted post-test scores. At baseline, R.D. had a d' score of 2.91 on the 1-back task. This decreased to a d' of 1.84 adjusted score and a 1.98 unadjusted score at post-test, indicating a lower ratio of hits to false alarms than at baseline.

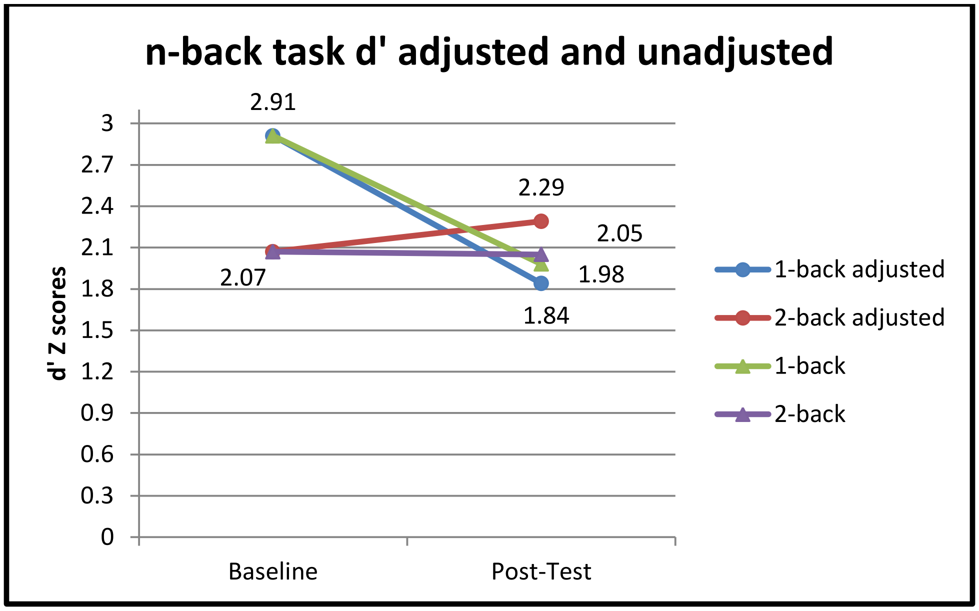

Figure 11: 1-back and 2-back n-back d prime (d') scores adjusted and unadjusted for number of items on post-test

TEA. The *Test of Everyday Attention* measures were conducted at baseline, pre-test and post-test. There was limited time at follow-up assessment; therefore, only select measures were chosen to be administered at that time. Performance on both subtests was variable across time points (Figure 12). On the Elevator Counting subtest, R.D. scored 5 at baseline, 6 at pre-test and 5 at post-test. The one-point difference did not indicate any significant change in performance across time points. On the Elevator Counting with Reversal subtest, R.D. scored 4 correct at baseline, 0 correct at pre-test and 3 correct at post-test. This did not indicate any significant effect, positive or negative, of treatment on behavioural performance on this measure. This suggests that this measure may not have been sensitive enough to detect any cognitive changes that occurred as a result of the online training program.

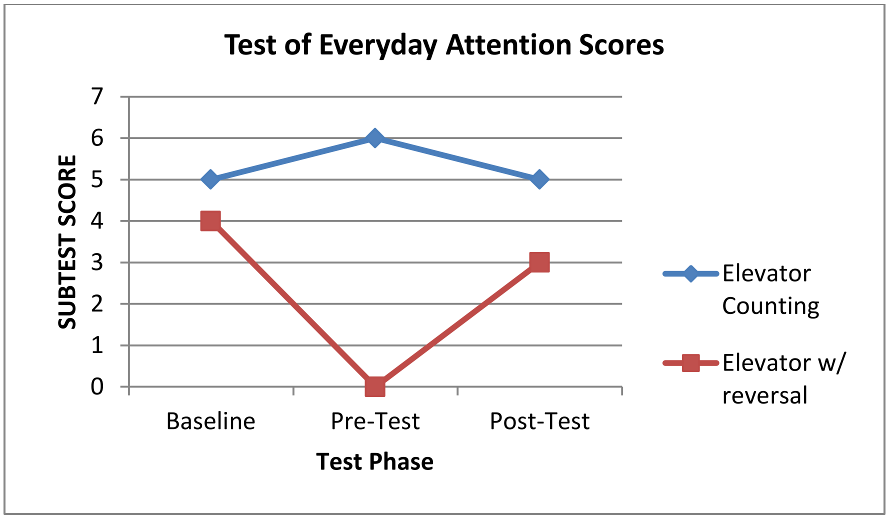

Figure 12: *Test of Everyday Attention* subtest scores for Elevator Counting and Elevator Counting with Reversal subtests

Summary

R.D.'s performance was examined with respect to the direct training measure, *BrainFitness,* as well as on a number of language and cognitive measures. R.D. spent a total of 33.33 hours training with the *BF* online program, with an improvement of 15.2% increase in his general performance. *WAB-R* scores indicated a clinically significant improvement in performance between baseline and post-test and a near-clinically significant (4.8 points) increase between pre-test and post-test. There was no clinically significant increase between baseline and pre-test or between post-test and follow-up. Although accuracy values on the *ALFAB* subtests remained fairly stable and did not fluctuate to a great degree across time periods, latencies decreased with treatment. This reaction time decrease, combined with the steady decrease in reaction time on the lexical decision task probes and the accuracy increase found in the behavioural data from the MRI at post-treatment, suggests that the *BF* training program had an effect on these language domains. The *n*-back adjusted d prime values indicated an increase in ability to respond correctly to a target 2 items back, which showed improvement on a more complex cognitive working memory task. However, a decrease in performance was found on the 1-back task, a task that is less cognitively complex. Descriptive examination of the Tompkins et al. working memory task and the *TEA* show a less clear pattern of changes in performance post-treatment.

Discussion and Conclusion

The goal of this study was to examine if computer-based cognitive treatment using the online *BrainFitness* (*BF*) program would improve cognitive and language functioning in an individual with aphasia. Cognitive training has resulted in positive outcomes for individuals with aphasia and previous use of the *BF* program was found to be effective for improving brain functioning in typically aging adults. Therefore, we sought to determine whether the *BF* program would be an efficacious treatment for certain individuals with aphasia. In addition to the direct training results obtained from the *BF* program itself, generalization measures for both cognitive and language functioning were taken.

To our knowledge, this is the first study of its type examining treatment of an individual with aphasia using a commercially-available computerized cognitive program. Although the study is limited to that of a single individual, positive results support further research into the use of the *BF* program its use and benefits with other clients.

Treatment Effects

BrainFitness. The *BF* program was chosen for intervention because it addresses working memory and attention through the auditory modality, two domains that are known to be impaired in aphasia. Further, the program follows principles of neuroplasticity in building up the complexity of stimuli from simple to more complex (e.g. tones to full conversations). Use of a program that employs the auditory modality is important as researchers have found that cognitive training must be in the modality being targeted (i.e. auditory comprehension) to see communication improvements (Kalinyak-Fliszar et al., 2011).

The client, R.D., showed general improvement in his performance on the *BF* program as treatment progressed. This is to be expected with repeated use of a program due to practice effects. However, this increase in performance is not very meaningful on its own as the results do not indicate which specific areas showed improvement. Only the overall percent increase was reported in the *BF* online training results. The direct treatment results would be more useful if additional information were available with regard to individual subtest improvements, which would allow researchers to examine which subtests were most beneficial for desired outcomes. The previous research that addressed the efficacy of the *BF* program did not report direct training results at all but instead used various measures of generalized cognitive functioning such as the *RBANS* (Barnes et al., 2009). Similarly, generalization measures were also used to assess the behavioural and functional implications of treatment on R.D.'s language and cognition. Therefore, it is more meaningful to examine outcomes with regard to generalization measures, which are more applicable to the real world, than to examine the BF results themselves. For this reason, in the following section I will discuss the generalization measures results and their implications in more detail.

Generalization measures-Language. When available, baseline, pre-treatment, post-treatment and follow-up measures were compared to determine if there was significant improvement. It was hypothesized that by the conclusion of treatment, the individual's scores on measures of working memory and attention would have increased from pre- to post-treatment, and maintained from post-treatment to follow-up time periods. As cognitive measures are thought to be related to improvements in language functioning in individuals with aphasia, it was expected that these score improvements would be mirrored in the measures of language functioning, specifically the *WAB-R* (Kertesz, 2006) and the *ALFAB* subtests (Westbury & Sheldon, 2005). This allowed for a quantitative analysis of language improvement based on treatment of cognitive mechanisms.

The largest improvements found post-treatment were in the language measures rather than in the cognitive measures. The *WAB-R* results showed promise for the efficacy of *BF* as a method for improving language. R.D.'s aphasia quotient increased slightly between baseline and pre-test. However, this was not enough of an increase to be clinically significant according to Katz and Wertz (1997). Between pre-test and post-test, the AQ increase neared clinical significance and between post-test and follow-up the AQ remained steady. This pattern suggests that although there are some possible minor practice effects due to re-administration of the same test, as evidenced by the small increases between time points during the no-treatment phases, this increase was not as large as the increase when treatment was applied. The difference in performance between the no-treatment and treatment phases supports that the improvements found were due to treatment rather than to repeat administration of the *WAB-R*. Additionally, the *WAB-R* has been found to have very high test-retest reliability for individuals with chronic aphasia, supporting temporal stability in scores across administrations (Shewan & Kertesz, 1980). If no change were present due to treatment, we would expect R.D.'s scores to reflect that fact and be stable across time points. However, to the awareness of the researchers, no studies to date have examined the validity of the *WAB-R* over closely-spaced repeat administrations.

The results of the other language generalization measure, the *ALFAB*, are less clear than the *WAB-R* results. In general, R.D.'s results still suggest an overall improvement in language using this measure. However, there were differences in performance across the three measures with respect to patterns of change. The three subtests varied in terms of their difficulty and the units of language processing (i.e. phoneme, word, sentence) that they measured. By comparing the pattern of results found in these subtests, we can hypothesis about how *BF* training generalized to various areas of language.

Overall, *BF* training appeared to have the largest effect on smaller units of language. This is likely due to a combination of factors. First, the Phoneme Discrimination subtest was the simplest, and most closely matched the tasks actually performed during the *BF* training. For example, the Fine Tuning module required the participant to discriminate between minimal pairs of syllables, giving the participant practice with the tasks used in the Phoneme Discrimination subtest of the *ALFAB*. The Lexical Decision task and Semantic Plausibility tasks, on the other hand, are more complex tasks and do not closely resemble any of the modules completed in the training. They required processing and integration of phonemes into words and sentences.

Previous research has found that although treatment of cognitive areas can have effects on language functioning, the results are at times stronger for less complex areas of language. Sinotte and Coelho (2007) found an improvement in their participant's AQ and reading rate, accuracy and fluency scores on the *GORT-4* due to treatment targeting attention but did not see a significant improvement in reading speed and comprehension of longer more complex passages used to measure functional improvements. Lexical decision and semantic plausibility judgments may be too complex to be treated through indirect cognitive training alone and must be treated directly in addition to the cognitive training. It is possible that, to see improvements in these more complex areas of language, direct therapist intervention is required with cognitive computer-based training as a possible supplement for additional at-home training.

Regarding performance on the Lexical Decision subtest, slower reaction times found in responses to non-words when compared to words is consistent with the established lexical effect found in previous research (Ratclif, Gomez & McKoon, 2004). Recognition of non-words required more time and, even with that additional time, was less accurate. Interestingly, post-treatment, reaction times for non-words improved to a level similar to pre-treatment reaction times for words. This could suggest that, as supported by the difference in performance on the subtests, *BF* training targets language at a sublexical level, addressing phonemes rather than the whole words. However, this improvement did maintain at follow-up.

Poor short term memory, as evidenced by R.D.'s poor performance on the Tompkins et al. working memory task, has also been found to be strongly related to performance on tasks that require syntactically-based comprehension (Caplan, Michaud & Hufford, 2013). As the Semantic Plausibility judgments required R.D. to process syntactic information along with the semantic information, R.D.'s poor short term memory may have led to poorer performance on this subtest than the other two due to the strong connection between short term memory and syntactic processing. If a larger improvement in short term memory capacity had been found as expected post-treatment, it is possible that performance on the Semantic Plausibility subtest may have been significantly improved as well.

Despite minimal post-treatment gains on the lexical decision and semantic plausibility subtests of the *ALFAB,* R.D. demonstrated improved performance on these tasks across fMRI scanning sessions (Figure 10). Although it is possible this may be due to practice effects (fMRI scanning sessions always took place within a few days to a week after behavioural assessment sessions), these results may simply reflect the inherent variability that is present in the language performance of people with aphasia. Multiple baseline assessment points (before and after treatment) may help in differentiating real treatment effects from spurious ones.

Complexity of stimuli may also be a possible explanation for the pattern of results. The Lexical Decision and Semantic Plausibility tasks may not have shown as clear-cut results because they are more complex and would require the intensive treatment to be applied for longer periods of time to show improvements. Bhogal, Teasell and Speechley (2003) found an average of 8.8 hours a week for 11 weeks in their high intensity studies showed the best results, so it is possible that if treatment were continued for a longer duration, generalization would have been found. Both semantic judgments and lexical decision require the participant to build on smaller learned segments, most likely phonemes. This explanation could account for why the more intensive Bhogal et al. studies found more promising results as these more complicated skills would likely require longer periods of time and higher intensity practice to show gains.

The results from the various language measures suggest that the *BrainFitness* program does not operate as generalized cognitive training for individuals with aphasia. Instead, it functions as cognitive training in a language context which in turn provides the participant with language training. The lexical decision task probes, as well as the lexical and semantic tasks completed during neuroimaging, support the *ALFAB* results that the *BF* training has a positive treatment effect on language. Although the results varied between the three measures in the performance across time points, all three showed a general trend of improvement post-treatment. These results support past research that cognitive-based training can have an effect on language of some individuals with aphasia. For example, Murray (2012b) reviewed a number of studies addressing treatment of working memory, attention and short term memory in individuals with aphasia and the language improvements found. Overall, the studies in the review support that these areas can be improved with treatment and that these improvements lead to at least some improvements in language, as found in this study.

Generalization Measures- Cognition. The results found regarding improvements in cognition post-treatment were much less straight forward than expected from previous research. Previous research predicted a much stronger effect of treatment on cognition that would mirror the effects found in language improvements. Either the cognitive measures used in this study were not sensitive enough to detect any changes present or *BF* training did not result in any changes in cognitive ability for R.D.. This is interesting, as the *BF* program was designed to target areas of cognition such as attention and memory, not language itself. In addition, previous research has shown how modifiable attention and working memory can be with treatment.

The Tompkins et al. (1994) working measure did not show any changes in performance post-treatment. R.D.'s baseline performance was 1 word recalled and pre-treatment performance was only 3 words recalled; therefore, floor effects may have precluded the usefulness of this measure as it did not seem highly sensitive to small changes in working memory for this individual and his presentation of aphasia. This task may be too complex to detect any cognitive changes that occurred as a result of treatment. Tompkins et al. found that individuals with left hemisphere damage had the highest number of errors on this task when compared to individuals with right hemisphere damage or controls. For individuals with left hemisphere damage, the range of responses was 1-36 words out of a possible 42 words recalled. In addition, Tompkins only used this measure at a single time point and it is possible that this measure was not suited to show small changes in working memory across repeated measures. In this study, the task was also stopped after administration of one set with no correct sentence pairs remembered due to the participant's frustration, whereas the Tompkins et al. participants completed the entire set, giving them more opportunity to produce a correct response.

The *n*-back results were used as a more statistically supported measure of R.D.'s cognitive performance baseline and post-treatment. Previous research supports the reliability and validity of the *n*-back tasks as a measure of working memory for individuals with aphasia. It has been shown to have test-retest reliability for as short a period as four weeks when administered to individuals with aphasia (Mayer & Murray, 2012). The statistical calculation of d prime used in the *n*-back is a more reliable measure of performance than simply reporting accuracy results. This measure takes into account how often the participant responds to incorrect items as well as correct items. In addition to this calculation, d' scores were adjusted to account for a possible length effect found in the post-test results. Performance decreased on both the 1-back and 2-back measures post-test and it was proposed that this decrease was due to a length effect as post-test included 90 more items than baseline. After accounting for this, the post-test results on the 2-back test showed an increase from baseline. It is possible that this is due to an increase in attention and working memory capacities post-treatment. This is in line with previous research that has found that language ability has an influence on working memory tasks when the task is linguistically related (Christensen & Wright, 2010). The decrease in performance on the 1-back task was not expected, especially as the more complex 2-back task showed improvements. This decrease is also in contrast to Cristensen and Wright's results, where all participants performed significantly better on the 1-back tasks than on the 2-back tasks. It is still unclear why these results differed to such a degree from previous research. In the future, usage of multiple time point measures could clarify the mechanism and reasons for this discrepancy.

Post-treatment, R.D. showed an improvement in his working memory through improvements on the adjusted 2-back task, which required the participant to hold information in working memory for a longer period of time than 1-back task. Previous research has shown that experience with repetition can lead to an improvement in working memory capacity (Kalinyak-Fliszar et al., 2011; Koenig-Bruhin & Studer-Eichenberger, 2007). During the observed sessions with the participant, it was noted that R.D. often repeated the stimuli out loud as he was trying to remember it during *BF* training. It is possible that repeating these stimuli throughout treatment lead to improvements in working memory such as those found in previous research and that this led to the improved ability to remember the stimuli required for the 2-back task. Additionally, as R.D. was repeating these stimuli back to himself, it is possible that he would be activating the dorsal language processing network (Hickok & Poeppel, 2007). This would lead to strengthening of this network, which is a possible mechanism of recovery as proposed in the Hickok and Poeppel model, which will be discussed further below.

On the *Test of Everyday Attention*, R.D.'s initial performance on the subtests could be used to provide additional evidence, in line with previous research, that individuals with aphasia can have concomitant impaired cognition in addition to their language impairment, as completion of the subtests required minimal language processing.

Results from the *TEA* measure were highly variable. As no pattern of change emerged across treatment, either no changes in attention performance occurred as a result of training or, similar to the results found in the Tompkins et al. (1994) task, it is possible that this measure was not sensitive enough to detect the changes in cognition that resulted from the use of *BF*. If there were changes post-treatment on this subtest, the measure was not sensitive enough to detect them. There are a couple possible explanations for the lack of pattern observed in the *TEA* results. In the manual of the *TEA*, it states that the subtest of Elevator Counting with Distraction may be too difficult for individuals with brain injury and should not be administered. The difficulty of this subtest for individuals with brain damage is one possible explanation for why performance was so variable. Research has also shown that there may be small practice effects on the *TEA* subtests, which may account for the small change found between baseline and pre-test on the Elevator Counting subtest (Chen et al. 2013).

Furthermore, previous research would lead us to expect that scores in R.D.'s *TEA* subtests would reflect improvements such as those found in R.D.'s aphasia quotient. Sinotte and

Coelho (2007) found an increase of 1 point on the Elevator Counting subtest and an increase of 7 points on the Elevator Counting with Distraction subtest, accompanied by a 4.7 point increase in AQ. However, the individual in their study was classified as having Anomic Aphasia, a type less severe than R.D.'s. The less severe type of aphasia could account for why their participant had less difficulty with the *TEA* subtests and why R.D.'s score was so variable on the Elevator Counting with Distraction test. The participant in Sinotte and Coelho's study could also have had a different distribution of subtest scores and improvements making up the AQ than R.D. did, which could affect the effect of treatment on performance. Additionally, individuals with aphasia are known for having highly variable performance, even within a single individual, which could also have played a role in the variable results of the *TEA* subtests. Any of these factors alone, or a combination of them, could possibly account for why R.D.'s performance differed from the results of previous studies.

The results of these cognitive measures reflect the poor working memory and attention abilities often present in individuals with aphasia. Interestingly, the cognitive measures used in this study did not show large improvements, even though the *BF* program is designed to target this area. As mentioned above, the lack of consistent improvement seen in R.D.'s performance either reflects that the cognitive measures used in this study were not sensitive enough to detect changes present or that, for this particular individual with aphasia, the training had more of an effect on language, as mediated through this cognitive domain.

Neural mechanism of speech processing.

BF targets processing speed, attention and working memory through the auditory modality. Therefore, models of auditory language comprehension and speech processing are particularly relevant to understand the improvement found on the language measures. A possible explanation of the role of neural mechanisms of speech processing is the dual stream model proposed by Hickok and Poeppel (2007).

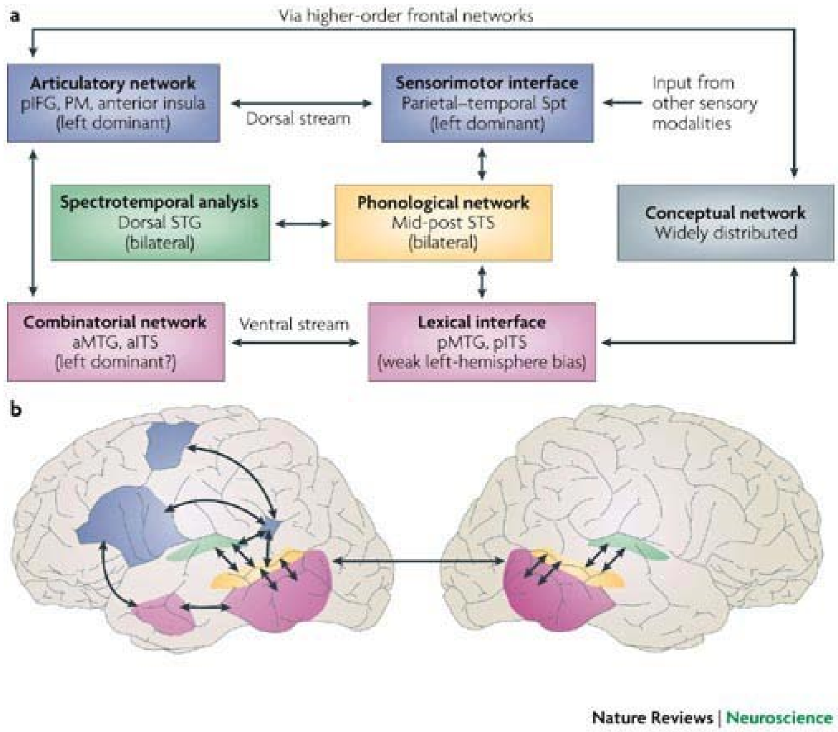

Figure 13: Hickok and Poeppel (2007) Dual Stream model of speech processing

This model (Figure 13) divides brain regions into a dorsal and ventral stream, each with a different role in speech-related processing. The ventral stream includes regions of the superior and middle temporal lobe and is bilaterally organized. Its functioning is more important for speech recognition and comprehension. This stream is proposed to perform a set of computations on the incoming acoustic signal that transforms this signal into a form that can be recognized by the mental lexicon as a specific word. The model assumes that the ventral stream is bilaterally organized, explaining why unilateral lesions in this area have not been found to strongly affect an individual's ability to recognize and assign meaning to incoming speech.

On the other hand, the dorsal stream is proposed to be involved in auditory-motor integration aspects of speech perception and involves the posterior frontal lobe, the posterior dorsal temporal lobe and aspects of the parietal lobe. This stream has a role in processing acoustic speech at a sublexical level. In contrast to the bilateral ventral stream, the dorsal stream is left-dominant. It is proposed that damage to this stream can account to some degree for language errors such as aphasia (Hickok & Poeppel, 2007). R.D.'s brain lesion includes the left superior temporal gyrus, overlapping with regions in the dorsal stream. Due to the location of this lesion, the presentation we would expect according to the dual stream model would be phonological production errors as a result of dorsal stream lesions.

It has been proposed that a major role of the dorsal auditory-motor integration circuit of the brain is to aid in speech development (Hickok & Poeppel, 2007). It has also been proposed that this stream is active even in adulthood. Learning to speak is primarily due to motor learning, but this learning also requires auditory input to guide fine-tuning and to ensure accuracy. If damage to this stream can lead to language errors in aphasia, recovery also likely involves the dorsal stream. This model proposes that auditory-motor interaction has two levels- one for segments and one for sequences of segments. Improvements in R.D.'s language would likely be due to changes in both the segmental and the sequence of segments level of the dorsal stream, which are involved in the "acquisition and maintenance of basic articulatory phonetic skills" and the "acquisition of new vocabulary, and in the online guidance of speech sequences" (p. 399) respectively. The *BrainFitness* program targets cognition through multiple tasks that require the participant to differentiate various segments, such as in the Fine Tuning and Memory Grid modules. By completing these tasks, R.D. was training and activating the auditory-based phoneme identification portions of his brain. According to Hickok and Poeppel's model, this would be activating the dorsal stream. In turn, this activation would lead to further auditory-motor integration.

According to this model, the increase in R.D.'s performance on various language production tasks, such as the *WAB-R*, and on speech perception tasks, such as the *ALFAB*, is due to auditory-motor interactions as a result of activation of the dorsal stream of the left hemisphere. The "sequences of segments" level of the dorsal stream would play a role in the improvements found with the higher levels of the *BF* training, such as the Syllable Stacks test where the participant hears a series of syllables and must repeat them back in order. By listening to these sequences as part of the *BF* program, according to Hickok and Poeppel, R.D.'s brain would generate a sensory representation that could later be used to guide articulation. Repeat activation of these networks can create a feed-forward system that would lead to increased efficiency in both processing and production. This activation could also account for an increased *WAB-R* score, as all of the tasks on this measure involve verbal responses.

The pattern found in the *ALFAB* results can also be accounted for by this dual-stream model. The Phoneme Discrimination task would operate at the simplest level, using acquired basic articulatory phonetic skills, while the Lexical Decision task would require these articulatory skills to be integrated into new words/non-words which naturally would require more processing as an extra step is involved. The Lexical Decision task would involve integrating the segments programmed at the segmental level of the dorsal stream into sequences. It is likely that the treatment effect was smaller than the effect for the Phoneme Discrimination subtest as the units the participant was required to attend to were larger and thus more complicated. Whereas the Phoneme Discrimination test was operating at the segments level, the Lexical Discrimination test was operating at the sequences of segments. The basic articulatory phonetic skills (segments level) would take less time to show an effect than the lexical level (sequence of segments) as more complicated skills generally take more time to be learned. It is possible that the Lexical Decision task only showed steady performance between post-test and follow-up as the skills learned in treatment would have had more time to consolidate since post-treatment measures. The Semantic Judgements subtest involves incorporating meaning and speech perception in one task. This would involve both the ventral and dorsal stream operating simultaneously. This entails the task being more complicated and, similar to the Lexical Decision task, would require more training to show a treatment effect.

Additionally, the model presented by Hickok and Poeppel can be used to account for the results found during analysis of words versus non-words on the Phoneme Discrimination and the Lexical Decision subtests. Previous researchers have noted that processing of words and pseudowords occur at different locations in the brain (Price, 2012). Researchers have found that many left lateralized areas of the brain are involved in processing meaning of words presented auditorily. Early research by Demonet et al. (1992 & 1994) using positron-emission tomography (PET) scans found that phonological processing (i.e., non-words) involves the supramarginal gyri and the left posterior inferior frontal gyrus, while semantic processing (i.e., words) more strongly activates the left middle and inferior temporal and angular gyri.

The inferior frontal gyrus is proposed to be part of the dorsal stream. This, in addition to the lexical effect mention above, could account for the overall performance difference between words and non-words. As R.D.'s lesion includes portions of the inferior frontal gyrus, this area would be less able to process the non-words presented at the segmental level (processing each phoneme individually) in order to determine if the presented stimulus was a word or a non-word. If the stimulus presented were a word, the somewhat preserved bilateral ventral stream could play a compensatory role through the activation of meaning associated with a real word. However, this explanation is contrary to the results demonstrated in the Semantic Plausibility subtest, as R.D. should have performed better on the semantic plausibility subtest if the bilateral ventral stream was intact. It is possible that this is due to the fact that cognitive factors such as working memory and attention interfered with processing of the sentences as they were longer and higher complexity than words alone.

This difference between performance on words and non-words did not exist on the Phoneme Discrimination task. This could be accounted for by proposing that the participant did not need to process words as a whole in this task and was only listening for individual sounds. This hypothesis is supported by Newman and Twieg (2001), where participants were asked to judge if an auditorily presented word or non-word ended in a /t/. Using fMRI, this study found no difference in activation of the inferior frontal gyrus. R.D. did not have a significant performance difference on this subtest between judgments on words and non-words as the task did not require activation of each individual segment to determine if it was a word or not.

As previously stated, the *BrainFitness* program targets processes mediated by the dorsal stream. The Lexical Decision Task results showed that, although R.D. performed best on words overall, the difference between pre-treatment and post-treatment was larger for non-words than for words. Pre-test to post-test for non-words showed an effect size of -0.59, a medium effect, with a mean difference between time points of 589.33 ms. Pre-test to post-test difference for words was also present, but to a smaller degree with an effect size of -0.45, a small effect, and a mean difference of 411.53 ms. This difference can be accounted for with the dual stream model. As the dorsal steam is targeted, treatment would show a greater improvement in non-words as the segmental level that controls detection of non-words is directly treated through this program.

It is possible that words showed less of an improvement because performance on them was initially better, but also because improvements of the impaired dorsal stream would have less of an effect on words than on non-words. However, the improvements found at post-test for non-words were not maintained at follow-up, although there was an effect between follow-up and post-test, follow-up was not largely different from pre-test. This indicates a return to pre-test performance without application of treatment. Future neuroimaging research could confirm or dispute these results suggested through behavioural data through comparison with brain imaging techniques. This would allow the researchers to determine if the changes due to *BrainFitness* that are proposed to be explained by the Hickok and Poeppel model could be supported through imaging data.

Intensity. R.D. was able to complete the *BrainFitness* module in the expected 8 weeks, completing a total of 33.3 hours of training. This is promising as it provides further evidence that independent computer-based treatment is feasible as a treatment method for individuals with aphasia. This program allowed R.D. to gain up to five hours of additional therapy a week on his own time. These results support those of the ASHA practice guidelines in Taylor-Goh (2005) discussing the therapy potential of a therapist-initiated at-home training program for a client. These guidelines discuss the potential of computer-based treatment as a method of increasing intensity. Although the program was only eight weeks in length, the daily set-up allowed for intensive application of treatment, closer to the high intensity 8.8 hours a week average found by Bhogal, Teasell and Speechley (2003) in their meta-analysis. Although this study did not compare the effects of *BF* based on intensity, past research provides support that more intensive treatment is more effective for individuals with chronic aphasia and that computer-based treatment is a very promising way to allow for this treatment when speech-language pathologists' time is limited. However, the individual in this study was highly motivated to complete additional training; therefore, future researchers and clinicians applying this program to therapy would need to be conscious of these factors to ensure client completion of the treatment program. If a client is not motivated to complete the modules on his or her own time with minimal clinician reminders, it would be more difficult to implement this program as a therapy option.

Limitations and Future Directions

This study is of an exploratory nature and examines a relatively new treatment method for individuals with aphasia. The measures used in this study have been well validated and a great deal of research supports the use of cognitive therapy as a possible treatment method for aphasia. However, there are certain factors that may have affected the results reported that must be taken into consideration when considering implications of this research.

Single case study designs are the most frequently used in the aphasia literature and have been shown to provide valuable information about this population when used properly (Thompson, 2006). However, there is a great deal of variability among people who have suffered a stroke and their presentation of aphasia. Because of this, the use of this type of design limits our ability to generalize results to the general community of individuals with aphasia. It would be near impossible to control for all possible factors to conduct a randomized controlled trial using *BF* for this population, especially at the early stages of analysis of this program as a treatment method. The results of this study present information regarding the efficacy of this program in treating one individual with a specific presentation which makes it difficult to determine if this treatment would be effective for treatment of individuals with aphasia in general. Given that this was a pilot study examining if further research into the use of *BF* as a treatment method was worthwhile, the lack of generalizability is less of a concern for the purposes of this study.

Another factor to consider related to the use of a single participant is the role that severity of aphasia can have on outcomes. R.D., similar to the individuals with aphasia chosen for many of the previous studies in the area of the role of cognition on aphasia, has relatively mild aphasia. It is unknown how this treatment would affect an individual with a more severe aphasia. Due to the promising findings with regard to *BF* treatment for R.D., further research will be required in the future to test applicability of the computer-based cognitive treatment method to a wider population.

Despite the fact that this was a single-subject design, many measures were taken to increase the rigor of the study. A multiple baseline across behaviours design was employed as, once training was applied, it would not be expected to have the same abilities as at baseline time point. The use of multiple measures across various time points allowed for measurement of behavioural changes over the course of treatment. Due to the necessity to re-administer multiple measures for the basis of comparison, an ABA design was used to account for this. The addition of baseline and follow-up time points increased reliability of results allowing for two periods of no-treatment to be compared to periods of treatment. If there was stability in behaviour between baseline and pre-test and between post-test and follow-up, but change between pre-test and post-test, this provides strong support that it was the intervention that caused these changes. This is shown most clearly in the *WAB-R* results. There was no clinically significant difference between time points in either of the non-treatment phases. After 8 weeks of *BF* intervention, though, there was a clinically significant change. In addition to these design factors, at the time of initial onset of treatment, R.D. was five years post onset of stroke, avoiding the period of maximal recovery. This means that any gains made can be attributed to the therapy and not to spontaneous recovery.

Another factor that readers should be conscious of is the technical difficulties present in administration of various measures during this study. Due to these difficulties, the *ALFAB* and the n-back could not be administered at every time point. This reduces the reliability of the results in describing patterns of change across each measure. However, the fact that multiple measures were used to assess language performance across the time points increases confidence in saying that the *BrainFitness* program did have an effect on language and that the changes were not due to random fluctuations in performance or to practice effects. The possibility of random fluctuations should be taken into account though for the *n*-back results as it was the only sensitive measure used to examine changes in cognitive abilities. As the measure is not able to show stability in performance with the use of a no-treatment phase, we are not able to reliably say that the improvements on the 2-back task were due to a treatment effect. Ideally, in future research, each of these measures would be taken at all four time points to provide a more reliable measure of stability and change over time.

Due to time constraints related to test administration difficulties, the amount of time between baseline and pre-test was only four weeks. This is shorter than the ideal length, which would be a no-treatment phase of the same length as the treatment phase. It is possible that, as this phase was shorter than the treatment phase, changes that occurred were simply due to the longer amount of time between measures rather than due to a treatment effect. However, this can be remedied by the fact that the post-treatment no-treatment phase was of equal length to the treatment phase. If the changes in R.D.'s performance were merely due to a certain length of time, there would be changes similar to the treatment phase present in the results of the second no-treatment phase. Although some differences were measured between post-test and follow-up on various measures, these changes were not as large as the changes seen between pre-treatment and post-treatment. This increases confidence that changes seen can be attributed to the treatment program rather than to an eight week time period.

During treatment, there was a period of five days where R.D. did not complete any online training. R.D. made up for this rest period by completing more training on subsequent days. Although this is likely not an area of large concern, as R.D. made up the total time at across the course of the study by training for than 5 hours the next week, this is a long period to go without training. However, as this was nearer the start of training rather than the end, it likely had minimal lasting effects on the final achievements in R.D.'s performance on the final measures.

Another area to consider for future application of the *BF* program is the neuroimaging aspect. The fMRI and MRI images taken during this study should be examined and compared to the behavioural data presented. Neuroimaging would allow an examination of the effect of *BF* training on the neurobiological structure and functioning of individuals with aphasia. It is unclear whether the modality in which treatment is applied to related to how recovery occurs. Further research would fill another gap in the literature, specifically the neural correlates of recovery due to auditory-based training. This would allow researchers to compare models such as that proposed by Hickok and Poeppel (2007) to actual physical results from neuroimaging pre- and post-treatment. Although many researchers agree with the concept of a dual-stream model, there is debate about the exact location of functions and recovery in these areas (see Gow, 2012 for discussion). Relating these imaging results to behavioural results would help to clarify mechanisms of recovery.

The effects *BF* training would have when combined with another treatment method are also unclear. As this program is commercially available and would likely be implemented as supplementary to treatment provided by a speech-language pathologist, it is important to examine the gains of *BF* training in this context. Intensity has been shown to be a very important factor in treatment and commercial computer programs have been shown in research to be a promising way to implement this (Taylor-Goh, 2005). This present study has shown that a participant can complete the training program independently and that this can lead to promising gains in language ability. However, this cannot fully replace the role of a speech-language pathologist, as evidenced by the fact that clinician monitoring of progress was important for continued compliance and by the fact that although improvements were seen in performance, clinician input would likely be needed to transfer these improvements into more functional everyday life situations. Therefore, further research on the exact outcomes that can be attributed to the program along with additional therapy should be examined. Related to this, as mentioned above, further research into which of the *BF* modules lead to the biggest gains in language abilities would be useful in maximizing patient time spent training outside of SLP treatment, by allowing them to spend more time on the most effective subtests.

Conclusion

This research has provided preliminary evidence for some potential benefits in using a commercially-available cognitive program as a treatment method for language impairments in an individual with aphasia. Results from our study suggest that training with the *BrainFitness* program can lead to improvements in a participant's spoken language and auditory discrimination skills of both words and sounds. This study did not show an effect of cognitive-based training on improving cognitive abilities as previous research has but it is possible that this is due to the lack of sensitivity of the measures used. The *n*-back measure of working memory did present evidence of some improvements post-treatment. These results provide support for future research regarding the implementation of *BrainFitness* as a treatment method for individuals with aphasia. Future research should work to further identify the aspects of *BF* that are most effective for producing the positive language improvements found, as well as on implementing the training with a larger number of participants to gain a more general impression of the program as a treatment method for individuals with aphasia. Future neuroimaging research could confirm or refute the results suggested through behavioural data through comparison with brain imaging techniques. This would allow the researchers to determine if the changes due to *BrainFitness* are mirrored in neuroanatomical and functional changes.

References

Aphasia Institute. (2012). *What is aphasia?* Retrieved from http://aphasia.ca/about-aphasia/what-is-aphasia/.

Barnes, D., Yaffe, K., Belfor, N., Jagust, W., DeCarli, C., Reed, B., & Kramer, J. (2009). Computer-based cognitive training for mild cognitive impairment results from a pilot randomized, controlled trial. *Alzheimer Disease & Associated Disorders, 23*(3), 205-210.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, *16*, 161-169.

Bhogal, S., Teasell, R., Foley, N., & Speechley, M. (2003). Rehabilitation of aphasia: More is better. *Topics in Stroke Rehabilitation, 10*(2), 66-76.

Brady, M., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews, 5*. doi: 10.1002/14651858.CD000425.pub3

Caplan, D., Michaud, J., & Hufford, R. (2013). Short-term memory, working memory, and syntactic comprehension in aphasia. *Cognitive Neuropsychology, 30*(2), 77-109.doi: 10.1080/02643294.2013.803958

Caspari, I., Parkinson, S. R., LaPointe, L. L., & Katz, R. C. (1998). Working memory and aphasia. *Brain and Cognition, 37*, 205–223.

Chen, H., Koh, C., Hsieh, C., & Hsueh, I. (2013). Test of Everyday Attention in patients with chronic stroke: Test-retest reliability and practice effects. *Brain Injury, 27*(10), 1148-1154.doi: 10.3109/02699052.2013.775483

Cherney, L. R., Halper, A. S., & Kaye, R. C. (2011). Computer-based script training for aphasia: Emerging themes from post-treatment interviews. *Journal of Communication Disorders*, *44*(4), 493-501. doi:10.1016/j.jcomdis.2011.04.002

Cherney, L. R., Patterson, J., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language & Hearing Research*, *51*(5), 1282-1299.

Cherney, L., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2010). Updated evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *ASHA.*

Christensen, S. C., & Wright, H. H. (2010). Verbal and non-verbal working memory in aphasia: What three n-back tasks reveal. *Aphasiology*, *24*(6/7/8), 752–762. doi:10.1080/02687030903437690.

Coelho, C.A (2005). Direct attention training as a treatment for reading impairment in mild aphasia. *Aphasiology, 19*(3/4/5), 275-283.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

Demonet, J.F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain: A Journal of Neurology, 115*(6), 1753-1768.

Demonet, J.F., Price, C., Wise, R., & Frackowiak, R.S. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. *Neuroscience Letters, 182*(1), 25-28.

Denes, G., Perazzolo, C., Piani, A., & Piccione, F. (1996). Intensive versus regular speech therapy in global aphasia: A controlled study. *Aphasiology, 10*, 385–394.

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language, 86,* 23–39.

Gow, D. W. (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language, 121*(3), 273-288. http://dx.doi.org/10.1016/j.bandl.2012.03.005

Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits [electronic resource].* Cambridge, Mass: MIT Press.

Gvion, A., & Friedmann, N. (2012). Does phonological working memory impairment affect sentence comprehension? A study of conduction aphasia. *Aphasiology*, *26(*3/4), 494-535. doi: 10.1080/02687038.2011.647893

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of d prime as a working memory index: Results from the 'Bergen n-back task'. *Journal of Clinical & Experimental Neuropsychology*, *32*(8), 871-880.

Heart and Stroke Foundation (2013). *Stroke prevention and risk factors.* Retrieved from http://www.heartandstroke.com/site/c.ikIQLcMWJtE/b.3483939/k.16FB/Stroke__Stroke_prevention_and_risk_factors.htm

Helm-Estabrooks, N. (1992). *Aphasia Diagnostic Profiles.* Austin, TX: PRO-ED.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*, 393-402.

Hula, W., & McNeil, M. (2008). Models of attention and dual-task performance as explanatory constructs in aphasia. *Seminars in Speech & Language*, *29*(3), 169-187.

Hula, W., McNeil, M., & Sung, J (2007). Is there an impairment of language-specific processing in aphasia? *Brain and Language, 103,* 240-241.

Kalinyak-Fliszar, M., Kohen, F., & Martin, N. (2011). Remediation of language processing in aphasia: Improving activation and maintenance of linguistic representations in (verbal) short-term memory. *Aphasiology, 25*(10), 1095-1131. doi:http://dx.doi.org.login.ezproxy.library.ualberta.ca/10.1080/02687038.2011.577284

Katz, R., & Wertz, R. (1997). The efficacy of computer-provided reading treatment of chronic aphasic adults. *Journal of Speech, Language & Hearing Research*, *40*(3), 493-507.

Kertesz, A. (2006). *Western Aphasia Battery- Revised.* San Antonio, TX: The Psychological Corporation.

Koenig-Bruhin, M., & Studer-Eichenberger, F. (2007). Therapy of short-term memory disorders in fluent aphasia: A single case study. *Aphasiology, 21*(5), 448-458.

Leung, A.W.S., & Alain, C. (2011). Working memory load modules the auditory “What” and “Where” neural networks. *Neuroimage, 55,* 1260-1269.

Martin, N., Kohen, F., Kalinyak-Fliszar, M., Soveri, A., & Laine, M. (2012). Effects of working memory load on processing sounds and meanings of words in aphasia. *Aphasiology, 26* (3/4), 462-493.

Mayer, J.F., & Murray, L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders, 45*(5), 325-339.

McNeil, M. R., Odell, K., & Tseng, C. (1991). Toward the integration of resource allocation into a general theory of aphasia. *Aphasiology Archive.* Retrieved from http://aphasiology.pitt.edu.login.ezproxy.library.ualberta.ca/archive/00000128/01/20-03.pdf

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.

Murray, L. (2012a). Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech-Language Pathology*, S51-64. doi:10.1044/1058-0360(2012/11-0067)

Murray, L. (2012b). Direct and indirect treatment approaches for addressing short-term or working memory deficits in aphasia. *Aphasiology, 26*(3/4), 317-337.

Murray, L., Keeton, R., & Karcher, L. (2006). Treating attention in mild aphasia: Evaluation of attention process training-2. *Journal of Communication Disorders, 39*(1), 37-61.

Newman, S.D., & Twieg, D. (2001). Differences in auditory processing of words and pseudowords: an fMRI study. *Human Brain Mapping*, *14*(1), 39–47.

Price, C.J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage, 62*(2), 816-847. doi: 10.1016/j.neuroimage.2012.04.062

Rabipour, S., & Raz, A. (2012). Training the brain: Fact and fad in cognitive and behavioral remediation. *Brain and Cognition, 79*(2), 159-179.

Robertson, I, Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *Test of Everyday Attention.* London, England: Thames Valley Test Company.

Robey, R.R., & Schultz, M.C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology, 12*(9), 787.

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

Sinotte, M., & Coelho, C. (2007). Attention training for reading impairment in mild aphasia: A follow-up study. *Neurorehabilitation, 22*(4), 303-310.

Smith, G.E., Housen, P., Yaffe, K., Ruff, R., Kennison, R.F., Mahncke, H.W., & Zelinski, E.M. (2009). A cognitive training program based on principles of brain plasticity: Results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *Journal of the American Geriatrics Society, 57*(4), 594-603.

Sohlberg, M.M., Johnson, L., Paule, L., Raskin, S.A., & Mateer, C.A. (2001) *Attention process training-II: A program to address attentional deficits for persons with mild cognitive dysfunction.* (2nd ed.). Wake Forest (NC): Lash & Associates.

Sturm, W., Willmes, K., Orgass, B., & Hartje, W. (1997). Do specific attention deficits need specific training? *Neuropsychological Rehabilitation*, *7*, 81–103.

Taylor-Goh, S. (2005). *RCSLT Clinical Guidelines: 5.12 Aphasia.* RCSLT Clinical Guidelines. Bicester: Speechmark Publishing Ltd.

Thompson, C. K. (2006). Single subject controlled experiments in aphasia: The science and the state of the science. *Journal of Communication Disorders*, *39*(4), 266.

Tompkins, C.A., Bloise, C.G.R., Timko, M.L., & Baumgaertner, A. (1994). Working memory and inference revision in brain-damaged and normally aging adults. *Journal of Speech & Hearing Research, 37*(4), 896-912.

Westbury, C., & Sheldon, S. (2005). *Alberta Language Function Assessment Battery.* Edmonton, AB: Westbury Lab.

Willis, S. L., Tennstedt, S., I., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., & Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of the American Medical Association, 296*(23), 2805-2814. doi:10.1001/jama.296.23.280

Wright, H., & Fergadiotis, G. (2012). Conceptualising and measuring working memory and its relationship to aphasia. *Aphasiology*, *26*(3/4), 258-278. doi:10.1080/02687038.2011.60430

APPENDICES

APPENDIX A

Participant Information Letter

Title of Research Study: The Impact of Computer Based Language Training in Chronic Aphasia

Principal Investigator(s): Dr. Jacqueline Cummine, Phone: (780)-492-3965

Dr. Esther Kim, Phone (780)-248-1542

Co-Investigator: Brea Chouinard, Phone: (780)-492-8759

Background: Persistent language difficulties in individuals who suffered a stroke are poorly understood. Recently, computer based treatments show promising improvements for such individuals. In addition, combining measures that look at behaviour and the brain provide a broad scope with which we can begin to understand these improvements. Together, these measures are important in the overall assessment and treatment of language difficulties in individuals who have had a stroke.

Purpose: Aphasia is an acquired impairment of language as a result of left hemisphere stroke, which can have significant consequences on the social, emotional, and occupational functioning of affected individuals. Participants with and without aphasia are being asked to participate in a research studyto examine the basic processes involved in language function, particularly speech comprehension.

Study Procedures:

For Individuals *without* Aphasia: You will have three visits to the University of Alberta, 8 weeks before beginning the computer program (baseline), as well as within one week before starting (pre-) and after completing (post-) the program. Each visit will involve a brief behavioral assessment and a visit to the NMR Center located in the University of Alberta Hospital. Each visit will take approximately 2 hours. At the first visit you will begin by reading this information sheet and signing a consent form. Next, a registered speech pathologist will administer a brief behavioral assessment in Corbett Hall. Following the assessment, we will go over to the NMR Center in the University of Alberta Hospital where you will complete a NMR screening checklist and get familiarized with the magnetic resonance imager (MRI scanner) room, and equipment. When we are ready to begin, you will be asked to lie in the scanner. While in the scanner, items will be shown to you on the computer screen or you will hear them through headphones, and you will be asked to do one or a combination of the following:

a. Name the picture (e.g., What is it?)
b. Make decisions about items (e.g., Do the picture and the word match? Is what you heard a real word? Could the sentence really happen?).

You will be removed from the scanner and the researcher(s) will answer any questions you have before you leave.

'Brain' Training: For 8 weeks following the "pre-" assessment session, (1 hour/day minus weekends), you will engage in a computer training program on a laptop that is provided to you by the researcher. This program is aimed at improving communication, hearing and memory through a variety of simple games. Following your last day of 'brain' training, the same procedures as visit 1 will be repeated including brief language assessment and a session in the scanner.

For Individuals *with* Aphasia**:** You will have 4 visits to the University of Alberta. At the first visit you will begin by reading this information sheet and signing a consent form. Next, a registered speech pathologist will administer a brief language assessment in Corbett Hall. Following the language assessment, we will go over to the NMR Center in the University of Alberta Hospital where you will complete a NMR screening checklist and get familiarized with the magnetic resonance imager (MRI scanner) room, and equipment. When we are ready to begin, you will be asked to lie in the scanner. While in the scanner items will be shown to you on the computer screen or you will hear them through headphones and you will be asked to do one or a combination of the following:

a. Name the picture (e.g., What is it?)

b. Make decisions about items (e.g., Do the picture and the word match? Is what you heard a real word? Could the sentence really happen?)

You will be removed from the scanner and the researcher(s) will answer any questions you have before you leave. Your first visit will take approximately 3 hours.

Treatment: For 8 weeks following the "pre-" assessment session, (1 hour/day minus weekends), you will engage in a computer training program on a laptop that is provided to you by the researcher. This program is aimed at improving communication, hearing and memory through a variety of simple games. Following your last day of treatment, the same procedures as visit 1 will be repeated including brief language assessment and a session in the scanner.

Follow-up: Three months following your last visit, you will be asked to return for a follow-up. The follow-up will be the same procedures as the other visits and will take approximately 2 hours.

Possible Benefits: For participants with aphasia, we are not sure if you will benefit from participating in this study. You may improve your language function, and you may learn about the basic processes involved in reading and/or speech perception. For individuals without language impairments, there is no direct benefit for participating in the study. The information we obtain may also help us develop better ways of helping people with language impairments in the future.

Possible Risks and Discomforts: You may feel claustrophobic in the MRI scanner. You will have an opportunity to see the MRI setting before participating in the study and if you feel you cannot do the specified tasks in the MRI setting you will be allowed to withdraw from the study.

Since the MRI is essentially a large magnet it is important that no metal be worn when near or in the MRI. You will be asked to complete and sign a separate document to ensure that you are able to have this test. When performing the specific study tasks, sometimes people are embarrassed of having made errors, but you should understand that making occasional errors is unavoidable (especially if you are asked to respond quickly). Thus, errors are normal and often expected in this type of research.

Confidentiality: Any data collected about you will be identified only by a code number, and will be protected at all times in secured computers which are behind locked doors when unattended. The researcher will safeguard and store the data, results, and associated material for a minimum of five years.

Voluntary Participation: You are free to withdraw from the research study at any time. If the study is not undertaken or if it is discontinued at any time all information/data collected will be destroyed. If appropriate, the researcher may choose to discontinue your involvement in the study in which case your data will be deleted and destroyed.

Compensation for Injury: If you become ill or injured as a result of participating in this study, necessary medical treatment will be available at no additional cost to you. By signing this consent form you are not releasing the investigator(s) or the institution(s) from their legal and professional responsibilities.

Contact Names and Telephone Numbers:

If you have concerns about your rights as a study participant, you may contact the Research Ethics Office at (780)-492-2615. This office has no affiliation with the study investigators.

Please contact the individual(s) identified below if you have any questions or concerns:

Dr. Jacqueline Cummine
Phone: 780-492-3965
Email: jacqueline.cummine@ualberta.ca

Dr. Esther Kim
Phone: 780-248-1542
Email: esther.kim@ualberta.ca

APPENDIX B

Participant Consent Form

CONSENT FORM

Title of Project: The Impact of Computer Based Language Training in Chronic Aphasia

Principal Investigator(s): Dr. Jacqueline Cummine Phone Number: 780-492-3965

Dr. Esther Kim Phone Number: 780-248-1542

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to withdraw from the study at any time, without having to give a reason?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐

Who explained this study to you?

____

I agree to take part in this study: YES ☐ NO ☐

Signature of Research Participant

____

(Printed Name)
____

Date:________________________________

Signature of Witness
____

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Person Obtaining Consent __________________________________ Date __________

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

APPENDIX C

Experimental Items: Tompkins, Boise, Timko, & Baumgaertner (1994) Working Memory Task

Level 2 Sets

Set 1

You sit on a chair. (T)
Trains can fly. (F)

Set 2

A table is an animal. (F)
Children like games. (T)

Set 3

Tigers live in houses. (F)
Milk is white. (T)

Level 3 Sets

Set 4

Sugar is sweet. (T)
Florida is next to Ohio. (F)
Horses run in the sky. (F)

Set 5

You ride on a bus. (T)
Cats can talk. (F)
Apples grow on trees. (T)

Set 6

Pumpkins are purple. (F)
Mice are smaller than lions. (T)
Roses have thorns. (T)

*Underlined words are words to be repeated

(T)= True; (F)= False

APPENDIX D

ALFAB Phoneme Discrimination Subtest Reaction Time and Accuracy Comparisons Across Time Points

Pre-test	**Reaction time (ms)**	**Accuracy (%)**
Overall	1053.35 (SD=381.77)	91.1
Non-word	979.46	91.2
Word	963.18	92.1
Post-test	**Reaction time (ms)**	**Accuracy (%)**
Overall	898.36 (SD=170.12)	92.2
Non-word	910.9	91.2
Word	883.82	92.1
Follow-up	**Reaction time (ms)**	**Accuracy (%)**
Overall	965.49 (SD= 231.63)	93.3
Non-word	994.93	94.12
Word	926.59	97.37

Note: These values do not include stimuli that included clusters; therefore, average reaction time values will not directly match values in Figure 3

APPENDIX E

ALFAB Auditory Lexical Decision Task Subtest Reaction Time and Accuracy Comparisons Across Time Points

Pre-test	**Reaction Time (ms)**	**Accuracy (%)**	**# items**
Overall	2104.43 (SD= 1154.05)	84.0%	
Word	1856.17	94.4%	
Non-word	2363.67	73.6%	
Highfreq*	1399.37	n/a	19
Lowfreq*	1870.15	n/a	34
Abstract*	1235.9	n/a	20
Concrete*	1167.12	n/a	26
Long*	1803	n/a	15
Short*	1661.26	n/a	38
Post-test	**Reaction Time (ms)**	**Accuracy (%)**	**# items**
Overall	1780.14 (SD= 1145.96)	79.9%	
Word	1444.64	81.7%	
Non-word	1774.34	77.5%	
Highfreq*	1370.56	n/a	18
Lowfreq*	1470.7	n/a	28
Abstract*	1515.64	n/a	22
Concrete*	1719.08	n/a	26
Long*	1314.08	n/a	13
Short*	1513.44	n/a	33
Follow-up	**Reaction Time (ms)**	**Accuracy (%)**	**# items**
Overall	1709.41 (SD= 902.63)	79.9%	
Word	1304.31	90.3%	
Non-word	2244.8	67.6%	
Highfreq*	1242.45	n/a	20
Lowfreq*	1577.18	n/a	33
Abstract*	1362.44	n/a	23
Concrete*	1367.07	n/a	27
Long*	1554.73	n/a	15
Short*	1317.14	n/a	36

*Calculated for words only; non-words excluded; accuracy not calculable as not all words were rated for frequency, concreteness and length.

APPENDIX F

ALFAB Semantic Plausibility Judgment Subtest Reaction Time and Accuracy Comparisons Across Time Points

Pre-test	**Reaction time (ms)**	**Accuracy (%)**	**# items**
Overall	3361.91 (SD= 1286.88)	73.8%	58
Adjective	3155.67	75.0%	12
Noun	2849.14	87.5%	14
Preposition	2523.17	43.8%	7
Verb	3634.67	75.0%	12
Gardenpath	4176.92	86.7%	13
Other	3123.47	71.9%	45
Post-test	**Reaction time (ms)**	**Accuracy (%)**	**# items**
Overall	3341.03 (SD= 1116.68)	77.5%	61
Adjective	3556.46	87.5%	14
Noun	3221.6	93.8%	15
Preposition	3155.25	50.0%	8
Verb	3052.36	68.8%	11
Gardenpath	3622	86.7%	13
Other	3263.32	75.0%	48
Follow-up	**Reaction time (ms)**	**Accuracy (%)**	**# items**
Overall	3013.15 (SD= 1044.49)	71.3%	57
Adjective	3751.42	75.0%	12
Noun	2658.46	81.3%	13
Preposition	2928.38	50.0%	8
Verb	2904.75	81.3%	13
Gardenpath	3376.64	68.8%	11
Other	3063.58	71.9%	46