Photocopy and Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission for extensive copying of my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

Measuring changes in lexical diversity of discourse samples to evaluate the effects of a Modified Intensive Comprehensive Aphasia Program

by

Martha Klousner

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Communication Sciences and Disorders

Idaho State University

Summer 2021

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Martha Klousner find it satisfactory and recommend that it be accepted.

_____

Victoria L. Scharp, PhD, CCC-SLP
Major Advisor

_____

Kris Brock, PhD, CCC-SLP
Committee Member

_____

Nancy Devine, PhD, PT, DPT
Graduate Faculty Representative

TABLE OF CONTENTS

List of Figures vi

List of Tables vii

Abstract viii

Chapter I: Introduction 1

Aphasia Background 1

Literature Review 1

Aphasia Assessment Process 1

Treatment of Aphasia, Frameworks, and Service Delivery Models 2

ICAP as an Effective Service Delivery Model 5

Discourse in Aphasia 7

Elicitation of Discourse Tasks. 9

Lexical Diversity 10

Methods of Measurement. 12

Purpose of this Study 13

Study Hypothesis 14

Chapter II: Methods 15

Experimental Design 15

Participants 15

Meridian Intensive Aphasia Program (MIAP) 16

Schedules and Procedures 17

Informed Consent Procedures 17
Program schedule 17
Diagnostics 18
Treatment Sessions 18
Discourse Tasks 19
Proposed Data Analysis 20
Chapter III: Results 21
Participants 21
Transcripts 22
Lexical Diversity Metrics 22
Descriptive Statistics 23
Number of Different Words (NDW) 23
Type-Token Ratio (TTR) 25
Moving Average Type-Token Ratio (MATTR) 26
Chapter IV: Discussion 28
Study Limitations 30
Future Directions 32
Conclusion 34
References 36
Appendix A - Measurements of Lexical Diversity 43
Appendix B - Discourse Tasks 45

Appendix C - Weekly MIAP Schedule 47

Appendix D - Individual Data per Discourse Task 48

List of Figures

Figure B1 Broken Window.....45

Figure B2 Refused Umbrella.....45

Figure B3 Cat Rescue.....46

List of Tables

Table 1 Participant Demographics 21

Table 2 Number of Different Words per Discourse Task 24

Table 3 Type-Token Ratio per Discourse Task.....26

Table 4 Moving Average Type-Token Ratio per Discourse Task.....27

Table D1 Number of Different Words for Broken Window Sequential Picture Description 48

Table D2 Number of Different Words for Refused Umbrella Sequential Picture Description.....49

Table D3 Number of Different Words for Cat Rescue Picture Description.....50

Table D4 Number of Different Words for Cinderella Familiar Narrative 51

Table D5 Number of Different Words for Puppy Love Unfamiliar Narrative.....52

Table D6 Number of Different Words for Stroke Story Personal Narrative.....53

Table D7 Number of Different Words for Letter Procedural Task 54

Measuring changes in lexical diversity of discourse samples to evaluate the effects of a Modified Intensive Comprehensive Aphasia Program

Thesis Abstract–Idaho State University (2021)

This study investigated lexical diversity as a parameter to examine the effects of an Intensive Comprehensive Aphasia Program (ICAP) service delivery model and its potential modifications. Lexical diversity was examined by analyzing discourse samples produced by people diagnosed with aphasia participating in Idaho State University's ICAP at three different time points (pre-treatment, post- treatment, and a follow-up therapy probe). The service delivery model used in this study is a modified version of the ICAP model where the primary difference is the length of time participants received services (1-week versus 2-6 weeks reported in the literature). Research referenced throughout this paper refers to studies using the typical ICAP model as the research for the modified version of this delivery model is in its infancy. Descriptive level statistical data from eight participants with non-fluent aphasia subtypes across seven discourse tasks was collected but no clear trends were noted in the group data though some emerging trends within the individual data were identified. Future studies are recommended with a larger population, incorporating different measures of lexical diversity such as the Word Information Measurement, and including a qualitative measurement to account for participant perspectives on the ICAP.

Key words: Aphasia, Lexical Diversity, Modified Intensive Comprehensive Aphasia Program, ICAP, Narrative Discourse Analysis

Chapter I: Introduction

Aphasia Background

Aphasia is an acquired language disorder that affects 25-40% of stroke survivors, equaling about 180,000 Americans every year (Aphasia FAQs, 2020). While the leading cause of aphasia is stroke, it can also be acquired after a trauma to the head, a tumor, or other neurological causes. According to the National Aphasia website, aphasia can affect people of all ages, genders, and races, though it is more common among the adult population. The symptoms of aphasia differ depending on where in the brain the trauma or tumor occurred, and the broad diagnosis of aphasia is broken up into fluent and nonfluent subtypes. Depending on the classification, a person with aphasia (PWA) can have impairments of expressive language (communicating a message to a communication partner), receptive language (receiving and understanding a message from a communication partner), or both. In addition, impairments cross all communication modalities such as reading, writing, or gestures (Hallowell, 2017).

Literature Review

Aphasia Assessment Process

Traditional aphasia assessments include standardized tests that assess both receptive and expressive domains of language. A comprehensive assessment also includes a case history to provide information on the patient's medical history and background as well as cultural and linguistic backgrounds, a self-report on their symptoms, progress, goals from either the patient themselves or a caregiver, and an oral-motor examination to determine if there are any co-occurring motor speech deficits such as dysarthria (muscle weakness) or apraxia (a motor planning difficulty) (American Speech-Language-Hearing Association, n.d.).

Once the severity and type of aphasia is diagnosed, clinicians can then begin to determine the strengths and weaknesses of the client's communication and create individualized goals and treatment plans to work on these areas. It has been noted in the literature (e.g., Andreetta & Marini, 2015; Marini, et. al, 2011) that traditional standardized assessments are not sensitive enough to capture the patient's pattern of linguistic deficits and recovery in PWA. To be truly representative of a patient's abilities, there needs to be an analysis of some sort of narrative or discourse sample as it is at this level that "linguistic skills interact with each other" (Andreetta & Marini, 2015, p. 706). In other words, there needs to be an assessment of everyday functional communication that cannot be captured through single word standardized assessments as everyday life is not comprised of just confrontation naming tasks. Kagan and Simmons-Mackie propose that assessments for this population should "be guided by 'real-life' outcome goals" that are as individualized as the patient and "related to life participation" in order to be the most beneficial as most interventions and decisions are heavily influenced by the original assessment (2007, p. 309).

Treatment of Aphasia, Frameworks, and Service Delivery Models

The treatment for all types of aphasia often focuses on functional communication whether that be speaking or through another modality such as gesturing, AAC devices, or writing and/or typing. Treatment techniques for people with aphasia (PWA) can be separated into restorative techniques and compensatory techniques. Restorative techniques (e.g., Semantic Feature Analysis; Boyle & Coelho, 1995) are where the clinician tries to rehabilitate a communication skill that the individual once had but has been impacted. In contrast, compensatory techniques (e.g., Visual Action Therapy; Helm-Estabrooks et al., 1982) are where the clinician helps the patient find strategies to work around deficits they may have so that they may achieve their

communication goals, whether that be returning to work or staying involved in their family and community.

One way to achieve this focus on functional communication for PWA would be to follow the World Health Organization's International Classification of Functioning, Disability, and Health (WHO-ICF) framework. The WHO-ICF is a holistic approach that incorporates more than just the medical diagnosis by including the functioning and disability, as well as other contextual factors of the individual undergoing treatment. In this context, functioning and disability refer to body structure and functioning (including its anatomical parts and both the physiological and psychological functions) in addition to the individual's activity and participation. The other contextual factors mentioned include such things as personal factors, or those characteristics that make up a person outside of their diagnosis such as education, gender identity, and life experiences, and environmental factors, which refer to factors outside of the person including physical surroundings and social contexts (Hallowell, 2017). This WHO-ICF approach allows SLPs to view the patient as more than their aphasia diagnosis and focus in on what the patient wants and their personal communication goals, as well as tailoring the treatment in order to cover all aspects of health and wellness (Bryant, 2017).

Following the WHO-ICF framework, goals should be focused on not only managing or eliminating unwanted symptoms resulting from the client's aphasia but be individualized and consider their mental and social well-being. Kagan and Simmons-Mackie recommend that goals for this population should integrate the client's "impairment, participation, environmental/barrier supports, and personal factors" that impact their overall quality of life (QOL) (2007, p. 312). Worrall et al. (2011), conducted a qualitative study that interviewed fifty people with aphasia to gain an understanding of what they perceived as critical goals for their treatment. The results of

this study included categories of goals related to communication, dignity and respect, and social/leisure/work, as well as others. Overall, the participants reported that they would like to be able to have social interactions such as "chatting with friends" and "read a night time story to the grandchildren", as well as be able to have a range in their communicative abilities, from communicating basic needs to expressing their own opinions and having the ability to connect with real-life and discuss "what goes on in the world" (Worrall et al., 2011, p. 314).

In order to provide treatment and work towards these goals, clinicians generally follow an aphasia-based framework. One of these frameworks, known as the Life Participation Approach to Aphasia (LPAA), was derived from the concepts of the WHO-ICF and emphasizes the everyday life of the patient and focuses on preparing them for "the communication needs of the community life" (Elman, 2016, p. 157). This framework was created due to the WHO-ICF proving to be too broad and not adequately addressing the life participation aspect central to PWA, as well as putting the PWA and their family in the center of the decision making (Chapey et al., 2000). The concept of life participation revolves around the "attainment of re-engagement in life" by participating in preferred daily activities (Chapey et al., 2000, n.p.).

From both the WHO-ICF framework and the LPAA, a new framework called the *Living with Aphasia: Framework for Outcome Measurement* (A-FROM) was developed to serve this particular population. This A-FROM framework places the client and their family at the center of the treatment, similar to the LPAA, as well as incorporates different aspects of their life, such as the client's personal identity, attitudes, and feelings, the severity of their aphasia, their participation in life situations, and their environment in order to create an intervention strategy that is personalized and dynamic (Kagan, et al., (2008). This framework expands upon the WHO-ICF framework by "explicitly incorporating aspects of identity and emotions" that directly

affect life participation as well as QOL (Kagan & Simmons-Mackie, 2007, p. 312). Within the A-FROM network are four domains: aphasia severity, participation/life habits, environment, and personal factors (including identity and emotions), which overlap to create the concept of "living with aphasia" (Kagan et al., 2008, p. 265).

ICAP as an Effective Service Delivery Model

The Intensive Comprehensive Aphasia Program (ICAP) is a relatively new service delivery model for aphasia that takes advantage of the brain's neuroplasticity via an intensive two-week (minimum) therapy course in which a cohort enters and exits the program together (Rose, et al., 2013). Neuroplasticity refers to the brain's ability to form new connections or neural pathways as a way to compensate for damage to an area of the brain. In other words, it is the way the brain is able to adapt to "internal and external influences" in order to "encode new experiences and learn new skills and behaviors" as well as learn these skills and behaviors anew through methods such as rehabilitation (Dignam et al., 2016, p. 256). As most strokes are generally in a localized area of the brain, these neural connections tend to be interrupted and some 'rewiring' may occur through the process of neuroplasticity. Crosson et al. (2019) asserts that treatment for PWA needs to be centered around restoring or reorganizing processes in order to capitalize on the brain's neuroplasticity and promote beneficial change. While research in this area is small and somewhat inconsistent due to the number of patient-specific variables such as lesion size and site, Mohr claims that "intensive, effective, and short-term aphasia therapy methods" are likely to facilitate neuroplasticity in PWA and encourage language recovery (2007, p. 4). Though it is important to note that depending on the location and degree of damage to the brain, as well as how the damage impacts the individual's functional systems, the extent of the neuroplasticity of the brain may have more limits (Crosson et al., 2019).

The basis for the ICAP model is highly individualized individual intensive treatment for the patient in conjunction with group therapy sessions, technology-based therapy, and counseling sessions for both the patient and the caregiver. In an ICAP service delivery model, the participant attends a minimum of 3 hours of daily treatment over at least a two-week period. The intensity that is incorporated into the ICAP philosophy reflects a number of factors associated with rehabilitative neuroplasticity, namely repetition, salience, and transference (Babbitt et al., 2015). Crosson, et al. (2019) suggests that in order to promote long-term improvements following aphasia treatment, one needs to have both intensive opportunities in which to produce target behaviors and plentiful repetition within the therapy sessions (otherwise known as saturated practice). In the ICAP model, communicative functioning and participation levels of language are targeted in both group and individual sessions. In addition, both the patient and family/caregiver are given education on the disorder and how they can best assist/adapt to the changes caused by the diagnosis (Rose, et al., 2013).

Differentiating the ICAP model from typical aphasia treatment is the intensity of the program and the comprehensiveness coupled with an emphasis on achieving personal goals (Babbit et al., 2015). In a study by Babbitt et al., the researchers found that while they were unable to determine any specific variables that indicated a participant would be able to make gains from treatment, the idea of "individualized, intensive, and comprehensive treatment" was beneficial for all different types of aphasia and various severity levels (2016, p. 182).

A noted difference in service delivery models is that group therapy is built into the ICAP model and is an integral component of the therapy program. Due to the ICAP model being based around a cohort that enters and leaves the program at the same time, participants are able to participate in group therapy (as well as the individualized and technology-based treatments)

which allows them to practice their skills in a supportive environment that encourages growth and generalization. Being introduced to others who share their struggles and have a deep understanding of their condition can provide motivation and promote further success in treatment (Griffin-Musick et al., 2020). In addition, these group therapy contexts can provide a crucial opportunity for the participants to engage in life participation in the form of typical discourse with their cohort and be less structured than in individualized therapy sessions, which can promote generalization of these skills. A number of studies have shown that PWA encounter a number of social consequences as a result of the communication difficulties arising from their strokes and subsequent aphasia such as a decreased desire to participate in social activities and negative feelings such as anxiety, shame, depression, a loss of self-confidence, and social isolation (Dalemans et al., 2010; Lima et al., 2020; Matos et al., 2014; Northcott et al., 2016). A study by Lima et al. (2020), asserts that along with showing evidence for improvement in functional communication, group therapy may also increase QOL for PWA through "provid[ing] opportunit[ies] for new friendships" and bolster "positive identity, independence, and self-confidence" (p. 2).

Discourse in Aphasia

Discourse is the "most naturally occurring and commonly used form of communication" (Fergadiotis, 2001, p. 1415). Discourse includes such things as conversations, picture or procedural descriptions, telling stories, etc. and as such, is used to complete many fundamental everyday tasks and activities. Linnik et. al, stated that "discourse is indispensable for human interactions" as well as noting that discourse is used to express an individual's feelings, ideas, and thoughts (2016, p. 765). With A-FROM in mind, a focus of treatment should be social participation through discourse, but in order to participate, one needs to have reliable

communication skills. PWA experience both "social and psychological strains" which negatively impact this participation (Dalemans, 2010b, p. 538). Dalemans et al. (2010a) describes how discourse is fundamental to human social interaction, and how "any defect affecting communication" can alter one's participation in these social events because these encounters depend "on the ability to engage with others, to interact, to share and to maintain equality" (p. 1678). Strong and Shadden (2020) describe how it is "widely accepted that humans are essentially storytellers" and how telling these stories demonstrates the "dynamic interaction between our shared stories and our identity" (p. 372).

The analysis of discourse is valuable to clinicians as this is a domain that is closely associated with real-life situations and employs communication participation (Cunningham, 2020). Because discourse reflects real-life situations, clinicians are able to tailor treatment to reflect the patient's interests or actual events and conversations they are likely to have and further emphasize the WHO-ICF approach for these individuals. Bryant, Spencer, and Ferguson (2017) describe discourse as "language in use" and identifies it as an "ecologically valid option for the evaluation of communication" (p. 1106). A main hallmark of aphasia is word-retrieval difficulties, PWA's discourse abilities may be impacted by the presence of fillers ("um" and "uh"), neoglisms (an invented word such as "gimble"), paraphasias (words in the same syntactic category "spoon" for "fork" or a word that is mispronounced "boon" for "spoon"), and/or non-referential terms (Fergadiotis, 2011). Through analyzing discourse samples, clinicians can study these and other cognitive and linguistic behaviors in as much of a naturalistic setting that is possible in a clinic in order to better understand what areas need treatment and what approaches may be best suited for the client. As evidenced by a study by Fromm, et al. (2017), discourse analysis can serve as a sensitive diagnostic marker for those who present with aphasia who might

slip through undiagnosed even after undergoing a standardized aphasia assessment. In addition, following along with the ideas of the A-FROM, it is important to take the time and energy to decipher discourse samples because of their correlation with everyday communication and their ability to allow individuals to complete fundamental activities such as giving instructions or sharing a story with a loved one.

Elicitation of Discourse Tasks. Traditionally, discourse tasks have been elicited by asking individuals to describe single pictures and/or picture sequences, provide story retells of familiar stories with or without picture stimuli (such as the story of Cinderella), engage in a personal narrative (such as recalling a time they were sick), and describe a procedural description (such as how to write and send a letter) (Nicholas & Brookshire, 1993). While still limited, there is growing evidence to suggest that the type of elicitation method impacts the information available from discourse-based language analysis in PWA (Fergadiotis & Wright, 2011; Stark, 2019). It is generally accepted that cognitive and linguistic demands may vary based on the elicitation technique of the task. For example, Stark (2019) found that narrative discourse may be "the most sensitive elicitation method for evaluating depth of vocabulary and content richness" when compared to procedural and picture description elicitation methods (pp. 1080-1081) in a sample of 90 PWA derived from the AphasiaBank repository (MacWhinney et al., 2011). Fergadiotis and Wright (2011) concluded that sequential pictures yielded more lexically diverse samples than single pictures, possibly due to the presence of additional information depicted within the images. As such variations occur based on elicitation method, Armstrong and Ferguson (2010) recommend utilizing numerous types of discourse tasks/elicitation techniques in order to best represent the client's language abilities in assessment and treatment capacities. While incorporating multiple types of discourse tasks is recommended, it is not always feasible

for clinicians in the field due to time constraints and the need to transcribe and code these samples, which has been estimated to require 6-12 minutes per sample (Stark, 2019). Clinicians do not always have the time to dedicate to transcribing and analyzing these samples so identifying and establishing certain parameters of interest for specific elicitation methods would prove clinically beneficial.

Lexical Diversity

Through the analysis of discourse samples clinicians can gain a plethora of information but the volume of information within a discourse sample could prove to be overwhelming. One component of a discourse sample that could be targeted to improve overall functional communication in an individual with aphasia is lexical diversity (LD). For the purposes of this paper, LD will be defined as, "the range of vocabulary deployed in a text by a speaker that reflects his/her capacity to access and retrieve target words from a relatively intact knowledge base (i.e., lexicon) for the construction of higher linguistic units" (Fergadiotis, 2011, p. 1415). Essentially, LD is the number of different words one has at their disposal to communicate with, including both how many different words types (nouns, verbs, adverbs, etc.) they employ and how many ways someone can say one thing (synonyms). The different word types refer to any words that convey meaning such as nouns, adjectives, verbs, and adverbs, whereas the different ways refer to synonyms of these words.

As described above, discourse is how one communicates with others in a typical context and is necessary in order to be able to functionally communicate and achieve everyday tasks and needs. LD also adds richness and personality to one's discourse through the use of figurative language and a dense vocabulary. If discourse is how one can express their thoughts, feelings, and ideas, then lexical diversity is a means for how to achieve this and personalize it. This is

important because language changes depending on one's communication partner and/or setting and being able to tailor one's expression helps form, build, and maintain relationships. One does not talk to their doctor the same way one might speak to a relative and having a substantial amount of lexical diversity allows one to communicate in these situations as expected. These different ways of speaking are known as different linguistic registers and are more clearly defined in foreign languages where they have various pronouns to denote these register changes. In American English, we don't have different pronouns to mark our shifts in register, but rather we use alternative words depending on our conversation partner and setting (Eaton, 2012).

In addition to decreased communicative abilities, aphasia can potentially cause damage to one's sense of self, or identity (Simmons-Mackie & Elman, 2011). One definition by Shadden and Agan (2004) states that identity is "a composite of roles, values and beliefs that are acquired and maintained through social interaction" (p. 175). Simmons-Mackie and Elman (2011) assert that identity is "intricately tied to communication" since one's word choice and own experiences "create an image for others" and provides "insights" into the individual's identity (p. 313). Strong and Shadden (2020) describe how telling stories is a dynamic process that reflects one's identity and changes depending on the environment and communication partner. If one is not able to engage in this dynamic process and utilize their LD, which is common in the case of PWA, they are unable to participate in these "meaning-making experiences" which can lead to a weakened identity and negative self-esteem (p. 373). When communicating, one does not only "convey factual information" to their conversational partners, but they also express their opinions and emotions in regard to the conversation in order to fully engage in social participation (Armstrong, 2005).

Methods of Measurement. There are several ways that lexical diversity has been measured in literature and each of these methods have their own strengths and weaknesses. A core method of measuring lexical diversity is the type-token ratio (TTR) where one divides the total number of different words (NDW) by the total number of words produced in the sample (Templin, 1957). If the ratio produced is closer to 0, the sample is considered to have less diversity within it while a ratio with a value closer to 1.0 denotes a sample with greater diversity. A number of alternatives based on the TTR have been created (such as the bilogarithmic TTR; Herdan, 1960) but like the true TTR measurement, they were dependent on sample size (Wright, 2003). Carroll (1964) asserts that as you increase the discourse sample size, the more likely you are to encounter the same words, decreasing the TTR, and as such may not allow for a direct comparison between discourse samples unless they are of equal length. This causes the discourse with a smaller sample size to "often appear richer" which is problematic when you are collecting samples from different individuals, and in the case of this study with a variety of aphasia types (Fergadiotis, 2013, pg. 2). Though attempts have been made to diminish the influence of differing discourse sample sizes by proposing a standardized discourse size, no conclusions have been made as there have been disagreements as to what this size should actually be (Cunningham & Haley, 2020; Fergadiotis, 2013). It is also acknowledged that even if a standardized discourse sample size was agreed upon, it may not always be possible to achieve for each individual. Because TTR is so dependent on discourse sample length, comparisons can only be completed on samples of equal length, which can be difficult to obtain. In addition, a number of computational methods have been created by building off of the TTR method to measure LD in a sample that account for differing sample sizes. These include the Moving Average Type Token Ratio (MATTR; Covington, 2010), the Measure of Textual Lexical Diversity (MTLD;

McCarthy, 2005), the D (McKee, Malvern, & Richards, 2000), and the Hypergeometric Distribution of D (HD-D; McCarthy & Jarvis, 2007). Appendix A discusses these different methods of measurement in detail.

Purpose of this Study

The purpose of this study is to expand the research base for the ICAP service delivery model and whether it can be an effective framework to treat PWA by using potential gains in lexical diversity of the participants as a measurement parameter. A study by Griffin-Musick et al. (2020), describes how traditional outpatient models, where the patient receives individualized therapy, twice a week, for about an hour, are often insufficient in their ability to provide a focused and comprehensive treatment of aphasia. With the ICAP model, patients are participating in a global rehabilitation process rooted in the ideas of neuroplasticity and A-FROM principles with a specific emphasis on social participation and the sense of identity.

When discussing a sense of identity and the importance of social participation in regard to aphasia, LD offers a potential link to the patient's goals. As mentioned above, to abide by the A-FROM and LPAA frameworks and take into consideration the wants and needs of the patients, it is crucial to tailor their goals to what is meaningful and functional for the client. One of the four main components of the A-FROM framework is centered around the participation in life situations, which includes relationships, communicating and conversations, and engaging in roles and responsibilities. Individualized goal setting connects to LD because when engaging in different social situations, you speak to people differently, depending on your relationship with them. For example, it is unlikely that you would use the same words to tell a story or delegate chores with both a loved one and a stranger. A diverse lexicon can help account for these different communication demands and experiences. In addition, LD can have a direct

relationship with one's sense of identity (incorporating another one of the A-FROM's central points), and with the growing emphasis on "interventions that not only improve communication skill, but also enhance quality of life", it is crucial that we "do not overlook the importance of a healthy sense of self" (Simmons-Mackie & Elman, 2011, p. 322). The comprehensiveness of the ICAP model, with a specific emphasis on group therapy, allows PWA to engage in practicing skills they have learned in a naturalistic, supportive environment and allow these individuals to break away from the typical focus of factual language (such as nouns and concrete verbs) and explore conveying emotions and expressing opinions commonly found in discourse.

Study Hypothesis

H_0: There is no difference in PWA's lexical diversity in discourse tasks following their participation in MIAP, based on pre-, post- and maintenance (1 month) data.

H_1: There is a difference in PWA's lexical diversity in discourse tasks following their participation in MIAP, based on pre-, post- and maintenance (1 month) data.

Chapter II: Methods

Experimental Design

This study was a retrospective, within-subjects cohort design, examining the application of an established quantitative discourse measure (lexical diversity) to analyze treatment effects of an evidence-based service delivery model (ICAP). The discourse tasks for the participants were recorded (through Video Audio Learning Tool [VALT]) so that they could be transcribed and analyzed by students in the Scharp Language and Brain Lab. Two trained students watched the video recordings and completed whole word transcriptions for the samples with coding for false starts, nonsense words, etc. The transcripts were subjected to an inter-reliability rating of a third trained student. Discourse samples were coded for total time, total number words, words per minute, and total discrepancies. For the purposes of this study, discourse samples needed to contain enough words to analyze (e.g., at least 10 words) and not be too heavily influenced by clinician/caregiver interjections or assistance. Viable samples were run through the Computerized Language Analysis program (CLAN; MacWhinney, 2000) in order to obtain the frequency counts needed to analyze the LD found in the samples. In this study, LD was measured by creating a frequency count of total number of words (TNW), the number of different words (NDW), and classifications of unique words and word types (TTR). Results from pre- and post- treatment samples, as well as data from a one-month treatment probe, were compared to look for significant changes or gains by utilizing specific LD metrics (i.e. MATTR).

Participants

The discourse samples analyzed in this study were collected from participants who attended the Modified Intensive Aphasia Program (MIAP) at Idaho State University (ISU) during the summer of 2019. Because of this, all of the participants were part of a population that was

suitable for ICAP and therefore may not have been truly representative of the entire population of PWA. Trained graduate students at the ISU clinic elicited the discourse samples with specific stimuli (picture description, video commercial, etc.) commonly used in clinical settings and aphasia research studies. Inclusion criteria for the participants included: 18-90 years of age; > 4 months post-stroke; fluent English speaker prior to stroke; documented medical diagnosis of aphasia, confirmed neurological damage via CT/MRI scan and/or evaluation report that indicated a diagnosis of aphasia or traumatic brain injury by a medical professional; corrected to normal hearing/vision; medically stable and able to tolerate activities for >6 hours per day per patient/family report for the intensive group and able to tolerate 5 hours of treatment per week for traditional/distributed group. Exclusion criteria included: lack of adherence to inclusion criteria; co-morbid neurological conditions; active chemical dependency per medical records, presence of severe cognitive impairments per the cognitive screener on the Comprehensive Aphasia Test (CAT).

Meridian Intensive Aphasia Program (MIAP)

MIAP refers to a modified ICAP (M-ICAP) located on the Idaho State University Meridian campus which provided 1080 minutes of treatment with a Treatment Intensity Ratio (TIR) of 75%, calculated by dividing the total number of therapy hours by the total number of possible treatment hours (Babbitt, et al., 2015). Treatment was delivered in a modified 1-week format compared to traditional ICAPs which range between 2-6 weeks. Treatment was provided by graduate student clinicians under the supervision of licensed speech-language pathologists and consisted of individual and group therapy, restorative and compensatory strategy training and practice, and a variety of evidence-based therapy approaches individualized to each PWA. Multiple individual therapy sessions lasting 50 to 75 minutes each were provided daily after the

first day and in addition to the individual sessions were 1 to 2 daily group therapy sessions (Gonzalez, 2019). The discourse samples used in this study were collected from eight PWA.

Schedules and Procedures

Informed Consent Procedures

All participants completed the ISU Clinic intake protocol which included an information sheet, authorization for the release of PHI (protected health information), and consent to receive treatment. In addition, the approved informed consent form (IRB-FY2018-184) was reviewed section by section with each potential participant. During and after the informed consent review process, the participants were asked if they have any questions or concerns. To confirm consent, the participants signed the documents in the presence of the student clinician as well as either the student's supervisor or the participant's family member.

Program schedule

Each client of MIAP had the same schedule structure throughout the week. Monday was a half day for clients and included a large group orientation and individual diagnostic sessions in which assessments were administered. From Tuesday through Thursday, clients participated in several 50-75-minute individual and group sessions. On Friday there was one last set of individual and group therapy sessions that included post-test measurements. That Friday afternoon, the clients presented a PowerPoint presentation about their stroke story with the support of their student clinician in front of all program participants and available family members. MIAP clients participated in approximately 1260 total minutes (30 hours) of treatment across the five-day program period (see Appendix C for daily schedule).

Diagnostics

Assessments were administered on Monday (pre-treatment) and Friday (post-treatment), in diagnostic sessions lasting about 75 minutes. Follow-up assessments (maintenance) were completed 10-12 weeks after participants' completion of MIAP. For each administration (pre-, post- and maintenance), the assessment battery consisted of the same standardized tests, chosen for their validity and reliability in measuring PWA's functional communication skills (Communication Activities of Daily Living (CADL-2); Holland et al., 1999), their word-finding ability (Boston Naming Test (BNT); Goodglass et al., 2001), the extent of their language impairment (Comprehensive Aphasia Test (CAT); Swinburn et al., 2004), and the level of their communication confidence (Communication Confidence Rating Scale for Aphasia (CCRSA); Babbitt & Cherney, 2010). This thesis project focuses on the discourse samples collected from MIAP participants.

Treatment Sessions

Individual treatment sessions were designed to support each participant's functional communication goals, which varied according to their strengths, deficits and individual profiles. Evidence-based approaches included Semantic Feature Analysis (SFA; Boyle & Coelho, 1995) and Verb Network Strengthening Treatment (VNeST; Edmonds et al., 2009) and other techniques suited to each PWA's unique needs and goals. Clinician-led group treatment sessions and informal lunches at the clinic site provided practical opportunities for clients to implement the skills and techniques learned in individual sessions within a social context, facilitating generalization. Caregivers, friends and family members were invited to join some group therapy sessions, which included counseling and education on topics pertinent to living with aphasia, as well as stroke prevention and aphasia advocacy. In keeping with ICAP principles, treatment at

MIAP targeted life participation, functional communication and language impairment goals for each PWA individually, while providing a range of opportunities for PWA to interact with other PWA in a social context supported by SLPs and graduate student clinicians.

Discourse Tasks

For the purpose of this paper, discourse will refer to the discourse tasks completed during MIAP, including 2 sequenced picture descriptions (*Broken Window* and *Refused Umbrella*), 1 single picture description (*Cat Rescue*), 1 procedural description task (describing how they would write and send a letter), 2 story retells (the story of *Cinderella* and *Puppy Love*, a commercial), and a personal narrative (the story of their stroke) (See Appendix B for task images and details). The variety of discourse tasks were chosen to reflect different elicitation methods in order to provide a more comprehensive analysis of spoken language that is representative of actual language use (Brookshire & Nicholas, 1994). Specifically, the picture descriptions (both sequenced and single) represent expositional narratives, the story retells and personal narrative represent narrative discourse, and the procedural task represents procedural discourse (Stark, 2019).

This study used these discourse tasks as a platform to analyze how discourse samples change over time between pre- and post- treatment. It should be noted that these tasks represent do not include a conversation, but monologues as elicited by different types of prompts. While most of these prompts have been documented in past studies, this study offers the novel prompt of *Puppy Love*, which includes the more dynamic elicitation technique of a video rather than still pictures.

Proposed Data Analysis

The independent variable in this study is the participation in MIAP for PWA and the dependent variables are possible changes in lexical diversity between pre- and post-treatment as well as a one-month therapy probe as measured by MATTR, TTR, and NDW. The original intent of this study was to measure PWA's LD through CLAN and analyze the NDW, TTR, and NDW for each of the seven discourse tasks. One-way repeated-measures ANOVAs were planned to determine changes in group performance from pre-treatment to post-treatment, as well as a one-month therapy probe. In order to account for the small sample size, it was proposed that a Kruskal Willis test may be appropriate in the case that the data results in an abnormal distribution in order to evaluate statistical significance using non-parametric methods if necessary. In addition, a Pearson's correlation coefficient (Pearson's *r*) was anticipated to show whether individual client variables (e.g. aphasia subtype, demographics, time post-onset) correlated with results.

Due to time constraints and complications due to the COVID-19 pandemic, adjustments were needed to the proposed data analysis. Data analysis for this project in the form of descriptive level statistics including mean, standard deviation, and range were used to describe trends throughout the range of discourse tasks.

Chapter III: Results

Participants

This study utilized discourse samples collected from 11 PWA. However, three PWA's samples were excluded, reducing the usable sample size to *n*=8 (4 men, 4 women). One excluded participant (C10) was ill during the week of MIAP and was only able to contribute 2 of 7 discourse samples. The other two (C3 and C18) required consistent student-clinician interjections for support during their discourse samples, rendering that data inadmissible due to deviation from elicitation protocols. Participant demographics for eight PWA are in Table 1. Participants had an average age of 54.8 (SD = 15.2; Range = 33-78) and 13.9 years of education (SD = 3.1; Range = 9-18). There were 4 males and 4 females in the participant sample and time post-onset ranged between 4-78 months (M = 20.6; SD = 25.5).

Table 1

Participant Demographics

Client	Age	Post Onset (months)	Education Level (years)	Gender	Aphasia Type
Client 1	78	78	Not reported	M	Nonfluent
Client 2	67	37	High School	F	Nonfluent
Client 5	57	14	Some college	M	Nonfluent
Client 7	60	4	Bachelor's degree	F	Nonfluent
Client 8	43	6	9th grade	M	Nonfluent
Client 17	33	6	College degree	F	Nonfluent
Client 19	61	11	Some graduate school	M	Nonfluent
Client 20	39	9	High school	F	Nonfluent

Participant demographics from MIAP 2019

Transcripts

Interrater reliability (IRR) was conducted on randomly generated transcripts for 20% of all of the transcriptions. Two trained graduate students who completed the transcriptions were in agreement for 99.54% across tasks and a third lab member was brought in to resolve the small number of discrepancies. Individual task-level IRR was: BW: 98.99%; RU: 100%; CR: 99.78%; CIND: 99.9% PUPP: 99.05%; SS: 99.29%; and LETT: 99.87%. Additionally, 20% of the CLAN transcripts were coded separately by two trained lab members and an IRR of 92.07% was achieved. The two lab members met to discuss the small volume of discrepancies for CLAN coding and came to 100% agreement.

Lexical Diversity Metrics

The metrics used in this study consisted of NDW, TTR, and MATTR. These calculations were all based on lemmas, meaning that words such as “eat”, “eats”, and “ate” were all counted as one word in order to minimize interference from different grammatical elements such as past tense (Fergadiotis, 2013). In addition, the scope of this study was focused on different words as opposed to derivations of the same word. NDW was chosen as a parameter in part because it has been used in research to “estimate the diversity of conversational vocabulary” and is not as affected by sample size as TTR due to its very nature of essentially being a word count (Watkins et al., 1995). TTR was included because although it is limited due to its reliance on discourse sample size, it is one of the most widely used measures of LD and as such can be easily compared to other samples in the existing literature. In addition, similar to NDW for TTR, TTR is a basis that is built upon in order to create different analysis procedures such as MATTR. MATTR was chosen as a parameter because it is a measurement that employs a customizable moving analysis window that one can manipulate to capture the number of words and is not

bound exclusively by discourse sample size like TTR. While it was originally anticipated that this study would include *VOCD* as a measurement parameter for LD, the data collected for our participants (made up entirely with individuals who present with non-fluent aphasia subtypes) were unable to provide the necessary 50 words (or tokens) necessary to complete the analysis for a majority of the discourse tasks.

Descriptive Statistics

Descriptive statistics including mean, standard deviation, and minimum-maximum ranges were used to describe the results for each of the different discourse tasks. The tables below (Tables 2-4) and in Appendix D (Tables 5-11) will assist in illustrating the trends within the group data between tasks, as well as demonstrate the changes in TNW, NDW, TTR, and MATTR for each client on each individual discourse task.

Number of Different Words (NDW)

Based on the data table shown below (Table 2), there is not a consistent upward trend between pre-, post-, and follow up data for the total number of different words across tasks for the group averages. Throughout the data for NDW there are large standard deviations indicating that there is a large amount of variability between the samples and the tasks and makes it difficult to find clear trends within the data.

For the *Cinderella* and *Cat Rescue* tasks, there is a slight upward trend. This slight upward trend is reflected in the individual data (see Tables 7 and 8 in Appendix D), specifically for Clients 1 and 2. Client 1 reflected this trend in both the *Cinderella* and *Cat Rescue* tasks, while Client 2 demonstrated this slight increase in only the *Cat Rescue* tasks due to only providing a sample for the *Cinderella* tasks during the post- collection time point.

Table 2

Number of Different Words per Discourse Task

	N	Mean (#)	SD	Range
BROKEN WINDOW (SEQENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	25	7.67	18 - 42
Post-MIAP	8	24.5	9.4	14 – 38
Follow up	5	26.4	7.23	19 – 38
REFUSED UMBRELLA (SEQUENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	34	15.18	16 - 60
Post-MIAP	6	30.33	13.94	13 - 47
Follow up	5	36.8	14.67	18 - 58
CAT RESCUE (PICTURE DESCRIPTION)				
Pre-MIAP	6	33.17	14.36	20 - 55
Post-MIAP	6	37.17	11.97	25 - 55
Follow up	5	39.8	14.2	24 - 61
CINDERELLA (FAMILIAR NARRATIVE)				
Pre-MIAP	6	104.67	43.03	40 - 157
Post-MIAP	7	110.71	61.56	28 - 214
Follow up	4	123.5	70.32	63 - 225
PUPPY LOVE (UNFAMILIAR NARRATIVE)				
Pre-MIAP	8	50	23.96	21 - 82
Post-MIAP	8	55.5	21.67	16 - 82
Follow up	4	50	33.43	16 - 90
STROKE STORY (PERSONAL NARRATIVE)				
Pre-MIAP	7	95.57	65.47	22 - 167
Post-MIAP	8	61.5	48.52	11 - 162
Follow up	6	66.17	65.99	12 - 186
LETTER (PROCEDURAL)				
Pre-MIAP	7	44.43	24.52	5 - 71
Post-MIAP	8	26.13	18.87	5 - 53
Follow up	6	30.33	19.4	11 -56

Descriptive statistics of number of different words (NDW) of each discourse task administered during MIAP 2019

Conversely, both the *Stroke Story* and *Letter* tasks show a considerable drop between the pre- and post- group means. This trend is also reflected in the individual data for *Stroke Story* (see Table 10 in Appendix D) by Client 1, who went from 68 to 50 NDW, Client 5 who went

from 167 to 67 NDW, and Client 20 who went from 158 to 57 NDW from pre- to post-treatment. The individual data from the *Letter* task (see Table 11 in Appendix D) reflected this decline between NDW also in Client 1 (71 to 38 NDW), Client 5 (54 to 25 NDW), and Client 20 (45 to 15 NDW).

Type-Token Ratio (TTR)

Table 3 (shown below) reflects the group data for TTR and while clear trends are not apparent between the means, the standard deviations are much smaller and the means are clustered together and generally fall between 5.0 and 6.7 with not much variability between pre-, post-, and the follow up data. The *Cinderella* task shows the smallest means, while the *Letter* task proved to have the highest means. The individual data (see Tables 8 and 10 respectively in Appendix D) show the TNW for each task and *Cinderella* had the largest number of overall words while *Letter* had the least.

Table 3

Type-Token Ratio per Discourse Task

	N	Mean (#)	SD	Range
BROKEN WINDOW (SEQENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	.671	.11	.56 - .91
Post-MIAP	8	.63	.07	.52 – .7
Follow up	5	.64	.11	.53 – .76
REFUSED UMBRELLA (SEQUENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	.63	.08	.52 - .76
Post-MIAP	6	.57	.09	.47 - .68
Follow up	5	.58	.05	.54 - .67
CAT RESCUE (PICTURE DESCRIPTION)				
Pre-MIAP	6	.61	.12	.41 - .76
Post-MIAP	6	.53	.1	.42 - .65
Follow up	5	.57	.09	.46 - .71
CINDERELLA (FAMILIAR NARRATIVE)				
Pre-MIAP	6	.46	.08	.37 - .58
Post-MIAP	7	.44	.04	.35 - .5
Follow up	4	.47	.1	.34 - .94
PUPPY LOVE (UNFAMILIAR NARRATIVE)				
Pre-MIAP	8	.56	.09	.39 - .68
Post-MIAP	8	.52	.1	.37 - .65
Follow up	4	.56	.1	.48 - .7
STROKE STORY (PERSONAL NARRATIVE)				
Pre-MIAP	7	.52	.1	.39 - .64
Post-MIAP	8	.59	.17	.36 - .94
Follow up	6	.56	.19	.32 - .87
LETTER (PROCEDURAL)				
Pre-MIAP	7	.66	.2	.47 - 1
Post-MIAP	8	.7	.2	.49 - 1
Follow up	6	.67	.18	.48 -.92

Descriptive statistics of type-taken ratio (TTR) of each discourse task administered during MIAP 2019

Moving Average Type-Token Ratio (MATTR)

Table 4 (shown below) reports the group means, standard deviations, and ranges for the MATTR parameter. All the discourse tasks and all three data collection points yielded at least a

.85 MATTR ratio. As all the data points are clustered together (between .85 and .92) this demonstrated minimal variation in MATTR due to the elicitation method of the discourse tasks.

Table 4

Moving Average Type-Token Ratio per Discourse Task

	N	Mean (#)	SD	Range
BROKEN WINDOW (SEQENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	.89	.06	.81 - .99
Post-MIAP	8	.85	.06	.76 – .92
Follow up	5	.86	.03	.83 – .9
REFUSED UMBRELLA (SEQUENTIAL PICTURE DESCRIPTION)				
Pre-MIAP	8	.92	.03	.86 - .96
Post-MIAP	6	.88	.07	.79 - .96
Follow up	5	.9	.05	.84 - .97
CAT RESCUE (PICTURE DESCRIPTION)				
Pre-MIAP	6	.87	.06	.79 - .93
Post-MIAP	6	.89	.07	.76 - .95
Follow up	5	.9	.03	.85 - .93
CINDERELLA (FAMILIAR NARRATIVE)				
Pre-MIAP	6	.92	.01	.91 - .93
Post-MIAP	7	.89	.03	.85 - .93
Follow up	4	.91	.03	.89 - .94
PUPPY LOVE (UNFAMILIAR NARRATIVE)				
Pre-MIAP	8	.89	.05	.8 - .94
Post-MIAP	8	.9	.04	.82 - .94
Follow up	4	.86	.07	.76 - .91
STROKE STORY (PERSONAL NARRATIVE)				
Pre-MIAP	7	.89	.06	.8 - .95
Post-MIAP	8	.88	.09	.67 - .96
Follow up	6	.86	.1	.69 - .98
LETTER (PROCEDURAL)				
Pre-MIAP	7	.92	.02	.89 - 95
Post-MIAP	8	.86	.09	.68 - .93
Follow up	6	.9	.09	.73 - .97

Descriptive statistics of moving average type-token ratio (MATTR) of each discourse task administered during MIAP 2019

Chapter IV: Discussion

Results of the preliminary analysis of the data suggest that there were no consistent trends to be found within the group data, partially due to the large individual variations between participants and the tasks. This is consistent with other discourse research where interpretation is difficult due to large variability between participants in terms of severity and subtype, as well as other factors and a lack of collaboration between research and clinical application (Bryant et al., 2017). It is also difficult at this stage of the analysis to point to a singular explanation for why some of these changes occurred. For example, in *Stroke* Story, there was a visible trend with a decrease in NDW between pre- and post- treatment. One can speculate that this may be because of the time between pre- and post- testing being only a week and individuals can sometimes be disinclined to repeat a whole story to the same individual if they believe the listener probably remembers it. This decrease may also be attributed to other outside factors, such as the participants wanting to finish quickly, fatigue, confusion regarding what the clinician is asking, etc. While it is difficult to interpret the data with its large variability between tasks and individuals, some of the inconsistent trends found in the group data such as a slight upward trend in NDW across all three data collection points is supported by slight upwards trends found within the individual data.

The patterns found in the group data for TTR, namely *Cinderella* having the smallest means across all three data collection points and *Letter* having the largest means, make sense when you take into account the TNW for each of these tasks. The story retell of *Cinderella* by far had the most TNW across all the discourse tasks with an average of 290.69 TNW while the *Letter* task averaged at 63.41 TNW across all three data collection points. Because TTR is directly impacted by discourse sample size as it is essentially the NDW divided by the TNW, it

follows that with as the discourse sample size increases, the more likely the sample will include the same words therefore decreasing the TTR (Carroll, 1964). The opposite is also true, where if there are fewer words in a discourse sample, there is a higher likelihood of them being more diverse so these smaller discourse samples appear richer when compared to longer ones.

The slight trends that were illustrated by the data for the narrative discourse samples (*Cinderella, Puppy Love,* and *Stroke Story*) included *Cinderella* having the smallest mean with regard to TTR and showing a slight upward trend for NDW across all three data points, and the *Stroke Story* data illustrating a considerate drop between pre- and post- treatment in regard to TNW. Based on our preliminary analysis, *Puppy Love* showed no trends that would differentiate it from the other elicitation tasks. While Stark (2019) found that narrative discourse may be best at evaluating both vocabulary depth and content richness when compared to procedural and picture description tasks, our data did not reveal any conclusive evidence to back up this finding. This may be in part due to the wide variability between individuals (especially for NDW and TNW) which impacts TTR and to a lesser extent MATTR. In addition, the fact that the number of participant discourse samples varies between tasks due to client refusal or other factors makes it difficult to compare between tasks.

Fergadiotis and Wright (2011) found that sequential pictures yielded more lexically diverse samples than single pictures. Our data was inconclusive and unable to support or refute this finding. The data for all three elicitation tasks representing expositional narratives (*Broken Window, Refused Umbrella,* and *Cat Rescue*) were clustered together and while *Cat Rescue* showed a slight upward trend in NDW between data collection points, in and of itself, NDW does not reflect the diversity of a sample. The means of the TTR for all tasks showed that *Broken Window* had a slightly higher mean, but once again, factoring in the difference in discourse

sample sizes, *Broken Window* generally had the smallest TNW. Focusing on MATTR values for these tasks shows the results range between .87 and .89 which does not denote any major variation or lend itself to any kind of trend within the data.

To reflect on the data from a more clinical perspective, what is clinically meaningful in a discourse sample must be addressed. Following the A-FROM principles, meaningful change is reflective of the individual client and their goals and needs, so while an increase of two words may not ultimately be statistically significant, if those two words are the client's loved one's name or their favorite food, it can still be meaningful to the client. In addition, the severity of the client's aphasia needs to be factored into interpretation of discourse samples. If a client increases their NDW from three to five, that is proportionally more meaningful than if the client increased their NDW from 17 to 22. This is why honoring the A-FROM and WHO-ICF guidelines is critical because what is meaningful for one client may not be for another and there are large individual variations in what is meaningful.

Study Limitations

One of the major limitations of this study is the small number of participants that comprise the data collected. This limitation is prevalent in communication sciences and disorders and makes it difficult to generalize results to greater populations (Haynes and Johnson, 2009). It is important to note however, that the typical number of participants in ICAP programs is approximately six individuals within a cohort, according to Rose et al., (2013), and therefore, our data reflecting eight participants is reflective of a typical ICAP program.

If this study had access to a larger number of participants, additional findings may have been possible, especially if the participants were a more diverse group in terms of aphasia subtype and severity. As all the data comes from participants who present with a non-fluent

aphasia subtype, this may skew results and is not inclusive of fluent subtypes. In addition, two participants were not included in the data due to the severity of their aphasia and need for high clinician support causing the discourse samples to be unusable in the data analysis.

A second limitation includes the difficulty of retaining participants to come back for the one-month therapy probe. A number of studies, including Babbitt et al., (2015), acknowledge that one of their limitations in ICAP studies is not including a follow-up therapy probe due to poor attendance by previous participants. Of the eight original participants, four to six of them returned and participated in all the discourse tasks for the probe. This means that for some of the tasks (such as *Puppy Love* and *Cinderella)*, probe data was only able to be calculated on half of the original participants, which also skew the descriptive statistics to not be reflective of the group as a whole.

Another limitation of this study was how to analyze and include those participants who presented with more severe aphasias who required more clinician support and interjections meaning the transcripts of these discourse samples were not viable for data analysis. Including a range and variety of severities and subtypes of aphasia was not within the scope of this study but is critical for future research. In order to analyze the data for measures of LD, the number of tokens in each discourse sample needed to be taken into account as discourse sample size directly impacts TTR and MATTR requires a minimum token window of 5 (a token window of 10 was used in this study) in order to analyze the data (Covington & McFall, 2010; Fraser et al., 2014). For those who present with severe aphasias, they may not manage to produce enough tokens to be analyzed with the chosen measures in this study.

It is important to find ways to include participants who present with more severe non-fluent aphasias into discourse analyses and using only verbal tokens does not encompass the full

range of functional communication these individuals have. A main aspect of aphasia treatment is a focus on functional communication, which means that verbal output may not be the sole focus. Instead, gestures, writing, sign language, vocalizations, facial expressions, and more may be utilized in order to communicate in a functional manner (Armstrong & Ferguson, 2010). Currently this study does not account for these different forms of communication and as such does not allow us to represent a considerable sample of the aphasia population.

A final limitation of this study is that LD was not a focus of treatment during MIAP. As such, none of the variables could be manipulated to influence LD and any changes found in LD will have to be attributed to the ICAP model as a whole and not tied to one specific aspect of the service delivery model. While these findings could provide evidence for the overall effectiveness of the ICAP model, it provides no further information on which component was most beneficial for changing LD. However, one could look at it the other way and see it as a strength. For example, if participants in an ICAP program made improvements on untreated targets, such as LD, this could be indicative of the validity of the comprehensiveness that makes up an ICAP program. If it is possible to prove that positive changes can be made on untreated targets, this means that gains outside of the trained targets can be made within an ICAP service delivery model.

Future Directions

The first future direction is to increase the sample size beyond eight viable participants and diversify the aphasia population. As previously discussed, having a population comprised only of individuals presenting with non-fluent aphasia subtypes limits the types of analyses one can perform due to not having enough words or tokens to analyze. If it is not possible to increase the size of the cohort, future research can include data from previous years in order to increase

the overall quantity of data available to be analyzed. Furthermore, finding a way to include and analyze samples from more severe non-fluent aphasia subtypes will diversify the results of the study and encompass this population as a whole. In this way, the data may be less likely to have a Type II error and become more generalizable to different populations.

Another future direction is to look into different data analysis methods, such as the Word Information Measurement (WIM) in order to analyze for different factors such as the influence of aphasia severity and subtype. While MATTR is a good tool to identify the presence of aphasia and mild cognitive impairment, WIM was found to be more sensitive at detecting the variations in both severity and aphasia subtype for PWA (Cunningham & Haley, 2020). Because many existing studies on LD diversity focus strictly on how to measure it (Cunningham & Haley, 2020; Fergadiotis et al., 2013) to the field needs to push for more studies on the application of LD in the assessment and treatment domains.

In addition to exploring other data analysis methods, comparing ICAP programs of different lengths may provide further insight into what is the optimal length of the ICAP program. According to Rose et al. (2013), ICAPs generally run from 12 to 33 days with the ISU's MIAP program being an exception as it lasts 5 days in total. This brings up the question of what length of time allows for recognizable change for our patients and if it would be beneficial to extend the MIAP program in order to provide the patients with enough time to demonstrate significant change.

A final future direction would be to include qualitative data such as patient interviews, as well as possibly interviews with family/caregivers in order to encompass the WHO-ICF and A-FROM principles to see if they report a difference in LD or in overall functional communication as a result of MIAP participation. While gaining two new words may not be statistically or

clinically significant, if those words include a loved one's name or a favorite activity, one can imagine that it would be meaningful for the client and as such holds significance in its own right. Furthermore, it would be important to research and acknowledge the possible presence of fatigue or burnout that the clients may be experiencing. An article by Barker-Collo et al. (2007), described that 39-72% of stroke survivors presented with persistent and often significant fatigue which frequently had a negative impact of daily life and rehabilitation for these individuals. As ICAP is a very intensive program (typically 48 to 150 hours of service; Rose et al., 2013), burnout is a very real possibility for these clients and may be a barrier for further growth. Adding a qualitative parameter and measuring the potential influence of fatigue to this line of research would better encompass the ICAP model as a whole as one of the principles for the model is treating the client as a complete individual and not simply their symptoms or diagnosis.

Conclusion

In conclusion, the results from this retrospective within-subjects cohort-based study add to the literature base for the ICAP service delivery model, elicitation methods for discourse tasks, and the application of LD in the assessment of PWA. Primary outcomes were inconclusive due to a high variability between discourse tasks and individual participants, but some trends were noted. These trends include (1) a slight increase in NDW for both the *Cinderella* (narrative discourse) and *Cat Rescue* (expositional narrative) from pre-treatment, post- treatment, and the follow up probe; (2) a considerable drop in NDW for *Stroke Story* (narrative discourse) and *Letter* (procedural discourse) between pre- and post- treatment; and (3) with MATTR as a measurement of LD, there was very little variability between tasks and data collection points. Because the results were inconclusive, it is recommended that future research investigate different measures of LD (such as WIM) in order to determine if other measures are more

sensitive to the changes in elicitation protocol, severity of aphasia, and aphasia subtype. LD is a strong parameter of interest in discourse tasks, especially when abiding by the principles of A-FROM by providing clients with the opportunities to express their sense of self-identity and be able to engage in a variety of social situations. Further research is necessary to continue investigating the application of LD in discourse analysis with different elicitation methods across a wider range of PWA within a variety of treatment contexts.

References

American Speech-Language-Hearing Association. (n.d.). *Aphasia*. American Speech-Language-Hearing Association. https://www.asha.org/Practice-Portal/Clinical-Topics/Aphasia/.

Andersen, E. S., Brizuela, M., Dupuy, B., & Gonnerman, L. (1999). Cross-linguistic evidence for the early acquisition of discourse markers as register variables. *Journal of Pragmatics*, *31*(10), 1339–1351. https://doi.org/10.1016/s0378-2166(98)00108-8

Andreetta, S., & Marini, A. (2014). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, *29*(6), 705–723. https://doi.org/10.1080/02687038.2014.979394

Armstrong, E. (2005). Expressing opinions and feelings in aphasia: Linguistic options. *Aphasiology*, *19*(3-5), 285–295. https://doi.org/10.1080/02687030444000750

Armstrong, E., & Ferguson, A. (2010). Language, meaning, context, and functional communication. *Aphasiology*, *24*(4), 480–496. https://doi.org/10.1080/02687030902775157

Armstrong, E., & Ulatowska, H. (2007). Making stories: Evaluative language and the aphasia experience. *Aphasiology*, *21*(6-8), 763–774. https://doi.org/10.1080/02687030701192364

Babbitt, E. M., Worrall, L., & Cherney, L. R. (2015). Structure, processes, and retrospective outcomes from an intensive comprehensive aphasia program. *American Journal of Speech-Language Pathology*, *24*(4). https://doi.org/10.1044/2015_ajslp-14-0164

Babbitt, E. M., Worrall, L., & Cherney, L. R. (2016). Who benefits from an intensive comprehensive aphasia program? *Topics in Language Disorders*, *36*(2), 168–184. https://doi.org/10.1097/tld.0000000000000089

Barker-Collo, S., Feigin, V. L., & Dudley, M. (2007). Post-stroke fatigue—where is the evidence to guide practice? *The New Zealand Medical Journal*, *120*(1264). http://www.nzma.org.nz/journal/120-1264/2780/.

Boyle, M., & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology*, *4*(4), 94–98. https://doi.org/10.1044/1058-0360.0404.94

Brown, D. W. (2008). *Curricular approaches to linguistic diversity: Code-switching, register-shifting and academic language* (dissertation).

Bryant, L., Spencer, E., & Ferguson, A. (2016). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology*, *31*(10), 1105–1126. https://doi.org/10.1080/02687038.2016.1239013

Carroll, J. B. (1964). *Language and thought*. Prentice-Hall.

Chapey, R., Duchan, J. F., Elman, R. J., Garcia, L. J., Kagan, A., Lyon, J. G., & Simmons Mackie, N. (2000). Life participation approach to aphasia: A statement of values for the future. *The ASHA Leader*, *5*(3), 4–6. https://doi.org/10.1044/leader.ftr.05032000.4

Covington, M. A., & McFall, J. D. (2010). Cutting the gordian knot: The moving-average type–token ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94–100. https://doi.org/10.1080/09296171003643098

Crosson, B., Rodriguez, A. D., Copland, D., Fridriksson, J., Krishnamurthy, L. C., Meinzer, M., … Leff, A. P. (2019). Neuroplasticity and aphasia treatments: new approaches for an old problem. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(10), 1147–1155. https://doi.org/10.1136/jnnp-2018-319649

Cunningham, K. T., & Haley, K. L. (2020). Measuring lexical diversity for discourse analysis in aphasia: Moving-average type–token ratio and word information measure. *Journal of Speech, Language, and Hearing Research*, *63*(3), 710–721. https://doi.org/10.1044/2019_jslhr-19-00226

Dagenais, L. M., Boliek, C. A., Woodhouse, L., & Paslawski, T. M. (2014). Comparing service delivery models on communication outcomes in adults with aphasia secondary to stroke. *International Journal of Medicine and Allied Health Sciences*, *4*(2), 321–334.

Dalemans, R. J. P., De Witte, L. P., Beurskens, A. J. H. M., Van Den Heuvel, W. J. A., & Wade, D. T. (2010). An investigation into the social participation of stroke survivors with aphasia. *Disability and Rehabilitation*, *32*(20), 1678–1685. https://doi.org/10.3109/09638281003649938

Dalemans, R. J., de Witte, L., Wade, D., & van den Heuvel, W. (2010). Social participation through the eyes of people with aphasia. *International Journal of Language & Communication Disorders*, *45*(5), 537–550. https://doi.org/10.3109/13682820903223633

Dalton, S. G., & Richardson, J. D. (2015). Core-lexicon and main-concept production during picture-sequence description in adults without brain damage and adults with aphasia. *American Journal of Speech-Language Pathology*, *24*(4). https://doi.org/10.1044/2015_ajslp-14-0161

Dalton, S. G., Hubbard, H. I., & Richardson, J. D. (2019). Moving toward non-transcription based discourse analysis in stable and progressive aphasia. *Seminars in Speech and Language*, *41*(01), 032–044. https://doi.org/10.1055/s-0039-3400990

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social participation for older people with aphasia: The impact of communication disability on friendships. *Topics in Stroke Rehabilitation*, *15*(4), 325–340. https://doi.org/10.1310/tsr1504-325

DeDe, G., Hoover, E., & Maas, E. (2019). Two to tango or the more the merrier? A randomized controlled trial of the effects of group size in aphasia conversation treatment on standardized tests. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1437–1451. https://doi.org/10.1044/2019_jslhr-l-18-0404

Dignam, J. K., Rodriguez, A. D., & Copland, D. A. (2016). Evidence for intensive aphasia therapy: Consideration of theories from neuroscience and cognitive psychology. *PM&R*, *8*(3), 254–267. https://doi.org/10.1016/j.pmrj.2015.06.010

Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *American Journal of Speech-Language Pathology*, *4*(4), 130–134. https://doi.org/10.1044/1058-0360.0404.130

Eaton, S. E. (n.d.). Language register and why it matters (or: Why you can't write an academic paper in gangsta slang). https://drsaraheaton.wordpress.com/2012/05/22/language-register-and-why-it-matters-or-why-you-cant-write-an-academic-paper-in-gangsta-slang/.

Elman, R. J. (2016). Aphasia centers and the life participation approach to aphasia. *Topics in Language Disorders*, *36*(2), 154–167. https://doi.org/10.1097/tld.0000000000000087

Elman, R. J., & Bernstein-Ellis, E. (1999). The efficacy of group communication treatment in adults with chronic aphasia. *Journal of Speech, Language, and Hearing Research*, *42*(2), 411–419. https://doi.org/10.1044/jslhr.4202.411

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*(11), 1414–1430. https://doi.org/10.1080/02687038.2011.603898

Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology*, *25*(10), 1261–1278. https://doi.org/10.1080/02687038.2011.606974

Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research*, *58*(3), 840–852. https://doi.org/10.1044/2015_jslhr-l-14-0280

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology*, *22*(2). https://doi.org/10.1044/1058-0360(2013/12-0083)

Fraser, K. C., Hirst, G., Graham, N. L., Meltzer, J. A., Black, S. E., & Rochon, E. (2014). Comparison of different feature sets for identification of variants in progressive aphasia. In P. Resnik, R. Resnik & M. Mitchell (Eds.), Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical

Reality (pp. 17–26). Association for Computational Linguistics. https://doi.org/10.3115/v1/W14-3203

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B. (2017). Discourse characteristics in aphasia beyond the western aphasia battery cutoff. *American Journal of Speech-Language Pathology*, *26*(3), 762–768. https://doi.org/10.1044/2016_ajslp-16-0071

Gordon, J. K. (2008). Measuring the lexical semantics of picture description in aphasia. *Aphasiology*, *22*(7-8), 839–852. https://doi.org/10.1080/02687030701820063

Griffin-Musick, J. R., Off, C. A., Milman, L., Kincheloe, H., & Kozlowski, A. (2020). The impact of a university-based intensive comprehensive aphasia program (ICAP) on psychosocial well-being in stroke survivors with aphasia. *Aphasiology*, 1–27. https://doi.org/10.1080/02687038.2020.1814949

Hall, N., Boisvert, M., & Steele, R. (2013). Telepractice in the assessment and treatment of individuals with aphasia: A systematic review. *International Journal of Telerehabilitation*, *5*(1). https://doi.org/10.5195/ijt.2013.6119

Harris Wright, H., & Capilouto, G. J. (2012). Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. *Aphasiology*, *26*(5), 656–672. https://doi.org/10.1080/02687038.2012.676855

Harris Wright, H., Silverman, S., & Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology*, *17*(5), 443–452. https://doi.org/10.1080/02687030344000166

Helm-Estabrooks, N., Fitzpatrick, P. M., & Barresi, B. (1982). Visual action therapy for global aphasia. *Journal of Speech and Hearing Disorders*, *47*(4), 385–389. https://doi.org/10.1044/jshd.4704.385

Herdan, G. (1960). Type-token mathematics. The Hague: Mouton.

Hess, C. W., Sefton, K. M., & Landry, R. G. (1986). Sample size and type-token ratios for oral language of preschool children. *Journal of Speech, Language, and Hearing Research*, *29*(1), 129–134. https://doi.org/10.1044/jshr.2901.129

Hilari, K., Needle, J. J., & Harrison, K. L. (2012). What are the important factors in health-related quality of life for people with aphasia? A systematic review. *Archives of Physical Medicine and Rehabilitation*, *93*(1). https://doi.org/10.1016/j.apmr.2011.05.028

Kagan, A., & Simmons-Mackie, N. (2007). Beginning with the end: Outcome-driven assessment and intervention with life participation in mind. *Topics in Language Disorders*, *27*(4), 309–317. https://doi.org/10.1097/01.tld.0000299885.39488.bf

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., … Sharp, S. (2008). Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*, *22*(3), 258–280. https://doi.org/10.1080/02687030701282595

Kim, H., Kintz, S., Zelnosky, K., & Wright, H. H. (2018). Measuring word retrieval in narrative discourse: core lexicon in aphasia. *International Journal of Language & Communication Disorders*, *54*(1), 62–78. https://doi.org/10.1111/1460-6984.12432

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, *51*(1). https://doi.org/10.1044/1092-4388(2008/018)

Leahy, M. M. (2004). Therapy talk. *Language, Speech, and Hearing Services in Schools*, *35*(1), 70–81. https://doi.org/10.1044/0161-1461(2004/008)

Lee, H., Lee, Y., Choi, H., & Pyun, S.-B. (2015). Community integration and quality of life in aphasia after stroke. *Yonsei Medical Journal*, *56*(6), 1694. https://doi.org/10.3349/ymj.2015.56.6.1694

Linnik, A., Bastiaanse, R., & Höhle, B. (2015). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology*, *30*(7), 765–800. https://doi.org/10.1080/02687038.2015.1113489

MacWhinney, B. (2000). The CHILDES project: Tools for analysing talk (3rd ed.). Mahwah, NJ: Erlbaum.

MacWhinney, B., Fromm, D., Forbes, M. & Holland, A. (2011). AphasiaBank: Methods for studying discourse. Aphasiology, 25,1286-1307.

Marini, A., Andreetta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, *25*(11), 1372–1392. https://doi.org/10.1080/02687038.2011.584690

Matos, M. A., Jesus, L. M., & Cruice, M. (2014). Consequences of stroke and aphasia according to the ICF domains: Views of Portuguese people with aphasia, family members and professionals. *Aphasiology*, *28*(7), 771–796. https://doi.org/10.1080/02687038.2014.906561

McCarthy, PM. Doctoral dissertation. 2005. An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity. Available from Proquest Dissertations and Theses. (UMI No. 3199485)

McCarthy PM, Jarvis S. Voc-D: A theoretical and empirical evaluation. Language Testing. 2007; 24(4):459–488.10.1177/0265532207080767

McKee G, Malvern D, Richards B. Measuring vocabulary diversity using dedicated software. Literary and Linguistic Computing. 2000; 15(3):323–337.

Mohr, B. (2017). Neuroplasticity and functional recovery after intensive language therapy in chronic post stroke aphasia: Which factors are relevant? *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00332

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research*, *36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338

Northcott, S., Moss, B., Harrison, K., & Hilari, K. (2016). A systematic review of the impact of stroke on social support and social networks: associated factors and patterns of change. *Clinical Rehabilitation*, *30*(8), 811–831. https://doi.org/10.1177/0269215515602136

Olness, G. S., & Ulatowska, H. K. (2011). Personal narratives in aphasia: Coherence in the context of use. *Aphasiology*, *25*(11), 1393–1413. https://doi.org/10.1080/02687038.2011.599365

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language & Communication Disorders*, *52*(6), 689–732. https://doi.org/10.1111/1460-6984.12318

Raymer, A. M., Ciampitti, M., Holliway, B., Singletary, F., Blonder, L. X., Ketterson, T., … Gonzalez Rothi, L. J. (2007). Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia. *Neuropsychological Rehabilitation*, *17*(2), 244–270. https://doi.org/10.1080/09602010600814661

Ribeiro Lima, R., Rose, M. L., Lima, H. do, Guarinello, A. C., Santos, R. S., & Massi, G. A. (2020). Socio-demographic factors associated with quality of life after a multicomponent aphasia group therapy in people with sub-acute and chronic post-stroke aphasia. *Aphasiology*, 1–16. https://doi.org/10.1080/02687038.2020.1727710

Rose, M. L., Cherney, L. R., & Worrall, L. E. (2013). Intensive comprehensive aphasia programs: An international survey of practice. *Topics in Stroke Rehabilitation*, *20*(5), 379–387. https://doi.org/10.1310/tsr2005-379

Simmons-Mackie, N., & Elman, R. J. (2010). Negotiation of identity in group therapy for aphasia: the Aphasia Café. *International Journal of Language & Communication Disorders*, 100917004613033. https://doi.org/10.3109/13682822.2010.507616

Stiles-Shields, C., Plevinsky, J. M., Psihogios, A. M., & Holmbeck, G. N. (2020). Considerations and future directions for conducting clinical research with pediatric populations during the COVID-19 pandemic. *Journal of Pediatric Psychology*, *45*(7), 720–724. https://doi.org/10.1093/jpepsy/jsaa055

Stark, B. C. (2019). A comparison of three discourse elicitation methods in aphasia and age-matched adults: implications for language assessment and outcome. *American Journal of Speech-Language Pathology*, *28*, 1067–1083. https://doi.org/10.31219/osf.io/vbu2d

Strong, K. A., & Shadden, B. B. (2020). The power of story in identity renegotiation: Clinical approaches to supporting persons living with aphasia. *Perspectives of the ASHA Special Interest Groups*, *5*(2), 371–383. https://doi.org/10.1044/2019_persp-19-00145

Templin, M. C. (1957). *Certain language skills in children: their development and interrelationships*. Greenwood Press.

Ulatowska, H., Streit Olness, G., Wertz, R., Samson, A., Keebler, M., & Goins, K. (2003). Relationship between discourse and western aphasia battery performance in african americans with aphasia. *Aphasiology*, *17*(5), 511–521. https://doi.org/10.1080/0268703034400102

Watkins, R. V., Kelly, D. J., Harbers H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research 38*, 1349-1355.

Wells, S. Y., Morland, L. A., Wilhite, E. R., Grubbs, K. M., Rauch, S. A. M., Acierno, R., & McLean, C. P. (2020). Delivering prolonged exposure therapy via videoconferencing during the COVID-19 pandemic: An overview of the research and special considerations for providers. *Journal of Traumatic Stress*, *33*(4), 380–390. https://doi.org/10.1002/jts.22573

Whitworth, A., Claessen, M., Leitão, S., & Webster, J. (2015). Beyond narrative: Is there an implicit structure to the way in which adults organise their discourse? *Clinical Linguistics & Phonetics*, *29*(6), 455–481. https://doi.org/10.3109/02699206.2015.1020450

Wilmskoetter, J., Fridriksson, J., Gleichgerrcht, E., Stark, B. C., Delgaizo, J., Hickok, G., … Bonilha, L. (2019). Neuroanatomical structures supporting lexical diversity, sophistication, and phonological word features during discourse. *NeuroImage: Clinical*, *24*, 101961. https://doi.org/10.1016/j.nicl.2019.101961

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2010). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. https://doi.org/10.1080/02687038.2010.508530

Appendix A - Measurements of Lexical Diversity

MATTR. The MATTR is similar to the TTR in that it measures the number of unique words, or **types**, divided by the number of words produced in the sample, or **tokens**. The key difference between the two is that the MATTR employs a customizable moving analysis window that one can manipulate to capture the number of words the individual is looking for. This window goes sequentially through the discourse sample one word at a time until the token ratio has been reached in all the windows that can be obtained in the sample and the mean token ratio of all of these sequential windows is what makes up the MATTR value (Cunningham, 2020). Depending on the area of interest, the size of the analysis window changes, for example if calculating the overall vocabulary is the goal then the window should be large. Limitations of this method include not being able to compare different windows coupled with the variety of language fluency found between aphasia types which can make it difficult to set a standardized window for the data set, though it is agreed that the minimum window is set at five words.

MTLD. The MTLD is similar to the MATTR in that it uses a sequential analysis but instead of windows, the scores indicate "the average number of words in a row for which a certain TTR is maintained" (Fergadiotis, 2013). The algorithm creates a factor count that increases by 1 every time the TTR drops below a value that has been predetermined. Once the entire sample has been analyzed, the number of total words is divided by the total factor count. After this value is found the entire text of the sample is reversed and the calculation is run again; the final MTLD score is a reflection of the mean of the scores calculated from both the forward and reversed scores.

VOCD (d). VOC*D* was developed as an alternative method to address the ongoing difficulties surrounding the varying sample sizes found in discourse samples. This measurement

employs random samplings throughout the sample and produces an empirical curve representative of the TTR against the tokens. After this step, the empirical curve is compared to a theoretical curve to create a value. Since D is calculated using random sampling and so varies, the data is often run through 3 times to establish a higher level of consistency. The average of these calculations is taken and considered to be the final D output. This value tends to range between 10 and 100 with greater LD being represented by higher values (McCarthy, 2010). Conceptually, VOC*D* represents how fast TTR decreases within a sample; for example, if the sample is comprised of the types that are used repeatedly, TTR would decrease more quickly (Fergadiotis, 2013). While this method addresses the difficult sample size variable found in discourse sampling, a sample of more than 50 words, which may present difficulties if the sample is obtained from an individual with a non-fluent aphasia subtype and may not be the best measure of LD for this population.

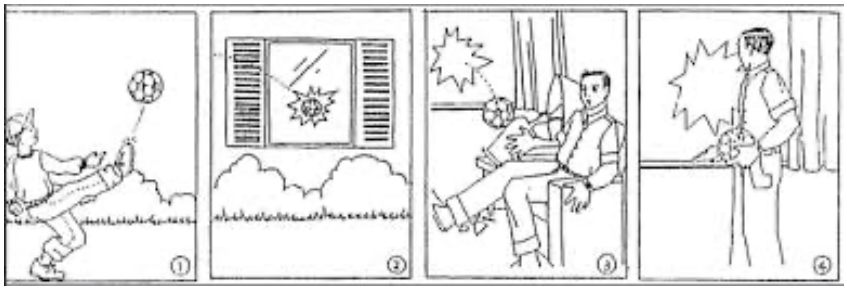
HD-D. HD-*D* is a method that combines *D* with hypergeometric distribution (HD) and the main assumption is that if the discourse sample contains a high number of tokens for a specific word, then the probability of drawing a sample containing at least one token of that word is high (Fergadiotis, 2013). The main difference between *D* and HD-*D* is that *D* is calculated based upon random sampling throughout a sample and introduces error through curve fitting, while HD-*D* is "directly estimated based on probabilities of word occurrence" (Fergadiotis, 2013). Another difference is that HD-*D* does not require a minimum of 50 words or tokens, making it more effective for analyzing those samples collected from individuals with non-fluent aphasias.

Appendix B - Discourse Tasks

Broken Window

Sequenced Picture Description (1)

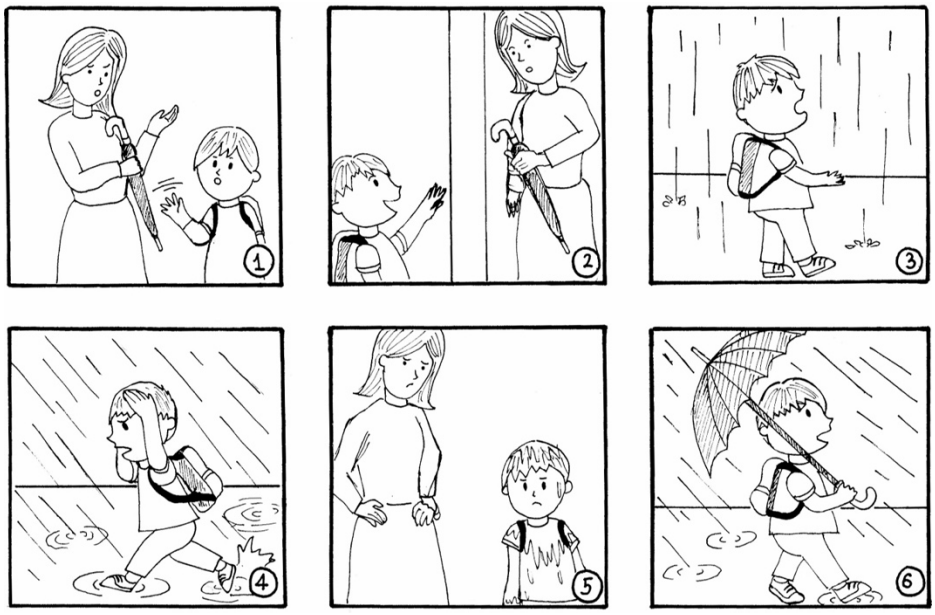
Figure B1



Refused Umbrella

Sequenced Picture Description (2)

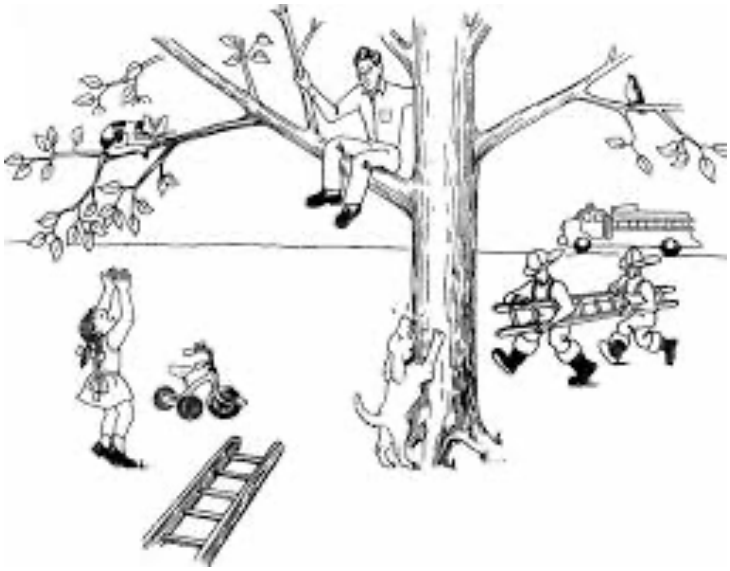
Figure B2



Cat Rescue

Picture Description

Figure B3



Cinderella Story

Story Retell (after looking at pictures)

Puppy Love

Dynamic Retell – Watch the video and retell the story looking at pictures provided

https://www.youtube.com/watch?v=dlNO2trC-mk

Personal Narrative

Narrative

Tell me about your stroke and your recovery process

Procedural Task

How to write and send a letter

Appendix C - Weekly MIAP Schedule

Schedule of Meridian Intensive Aphasia Program (MIAP)

Time	Monday	Tuesday	Wednesday	Thursday	Friday
9:00		Group	Group	Group	Group
10:00		Individual	Individual	Individual	Individual
11:00		Physical Therapy*	Group	Physical Therapy	Group
12:00		Lunch	Lunch	Lunch	Lunch
1:00	Group	Individual	Individual	Individual	Home Exercise Training
2:00	Individual	Group	Group	Group	Ice Cream Social
3:00	Home	Home	Home	Home	Home

*Physical therapy consisted of sessions led by physical therapists who completed small group sessions with student-participant dyads targeting balance and strength exercises.

Appendix D - Individual Data per Discourse Task

Table D1

Number of Different Words for Broken Window Sequential Picture Description

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	31	18	0.64	0.811	24	15	0.682	0.831	35	22	0.733	0.733
Client 2	30	18	0.62	0.84	35	17	0.515	0.758	45	27	0.643	0.643
Client 5	45	25	0.56	0.878	62	38	0.644	0.888	72	38	0.543	0.543
Client 7	35	23	0.62	0.846	24	14	0.636	0.838	29	19	0.76	0.76
Client 8	23	21	0.91	0.993	35	21	0.656	0.865	--	--	--	--
Client 17	45	26	0.58	0.914	50	24	0.522	0.795	50	26	0.53	0.53
Client 19	58	42	0.72	0.955	51	32	0.681	0.913	--	--	--	--
Client 20	38	27	0.71	0.907	57	35	0.7	0.922	--	--	--	--
Mean	38.1	25	.67	.89	42.25	24.5	0.62	.85	46.2	26.4	0.6418	0.8624

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Broken Window discourse task for MIAP 2019

Table D2

Number of Different Words for Refused Umbrella Sequential Picture Description

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	53	26	0.52	0.953	35	18	0.465	0,834	66	34	0.542	0.874
Client 2	42	25	0.64	0.953	76	33	0.465	0.834	59	32	0.542	0.874
Client 5	59	42	0.71	0.958	--	--	--	--	104	58	0.592	0.963
Client 7	31	18	0.64	0.863	20	13	0.684	0.88	31	18	0.667	0.839
Client 8	23	16	0.76	0.9	50	26	0.542	0.792	--	--	--	--
Client 17	121	60	0.51	0.912	--	--	--	--	73	42	0.575	0.936
Client 19	75	44	0.62	0.948	87	45	0.608	0.926	--	--	--	--
Client 20	66	41	0.64	0.904	75	47	0.653	0.957	--	--	--	--
Mean	58.75	34	0.63	0.92	57.17	30.33	0.57	0.88	66.6	36.8	0.58	0.9

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Refused Umbrella discourse task for MIAP 2019

Table D3

Number of Different Words for Cat Rescue Picture Description

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	108	40	0.41	0.929	128	55	0.47	0.934	148	61	0.459	0.902
Client 2	31	21	0.68	0.8	72	28	0.418	0.76	71	35	0.547	0.907
Client 5	--	--	--	--	94	47	0.573	0.905	92	46	0.568	0.904
Client 7	39	20	0.57	0.785	63	25	0.455	0.822	40	24	0.706	0.848
Client 8	33	22	0.76	0.915	--	--	--	--	--	--	--	--
Client 17	95	55	0.66	0.915	--	--	--	--	58	33	0.579	0.933
Client 19	79	41	0.56	0.897	52	29	0.63	0.924	--	--	--	--
Client 20	--	--	--	--	61	39	0.65	0.947	--	--	--	--
Mean	64.17	33.17	0.61	0.87	78.33	37.17	0.53	0.88	81.8	39.8	0.572	0.9

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Cat Rescue discourse task for MIAP 2019

Table D4

Number of Different Words for Cinderella Familiar Narrative

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	77	40	0.58	0.912	162	65	0.445	0.875	--	--	--	--
Client 2	--	--	--	--	68	28	0.438	0.856	--	--	--	--
Client 5	289	128	0.47	0.923	400	162	0.422	0.92	702	225	0.34	0.936
Client 7	--	--	--	--	--	--	--	--	128	63	0.578	0.881
Client 8	162	73	0.5	0.916	270	112	0.455	0.851	268	107	0.448	0.893
Client 17	230	98	0.48	0.908	189	89	0.497	0.91	200	99	0.532	0.921
Client 19	465	157	0.37	0.932	649	214	0.353	0.926	--	--	--	--
Client 20	368	132	0.38	0.913	239	105	0.461	0.911	--	--	--	--
Mean	265.17	104.66	0.46	0.92	282.4	110.71	0.44	0.89	324.5	123.5	0.47	0.91

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Cinderella discourse task for MIAP 2019

Table D5

Number of Different Words for Puppy Love Unfamiliar Narrative

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	117	45	0.39	0.89	100	43	0.48	0.90	--	--	--	--
Client 2	62	33	0.56	0.85	126	44	0.37	0.82	--	--	--	--
Client 5	179	82	0.5	0.91	155	76	0.51	0.9	197	90	0.48	0.91
Client 7	36	21	0.68	0.91	26	16	0.64	0.91	25	16	0.7	0.87
Client 8	40	22	0.58	0.8	82	49	0.65	0.87	61	30	0.52	0.76
Client 17	132	61	0.53	0.93	119	65	0.56	0.92	121	64	0.55	0.9
Client 19	154	80	0.58	0.89	227	82	0.41	0.92	--	--	--	--
Client 20	86	56	0.68	0.94	127	69	0.58	0.94	--	--	--	--
Mean	100.75	50	0.56	0.89	120.25	55.5	0.52	0.9	101	50	0.56	0.86

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Puppy Love discourse task for MIAP 2019

Table D6

Number of Different Words for Stroke Story Personal Narrative

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	130	68	0.53	0.91	115	50	0.57	0.88	--	--	--	--
Client 2	--	--	--	--	25	11	0.46	0.67	23	12	0.52	0.69
Client 5	429	167	0.45	0.95	124	67	0.57	0.91	41	33	0.87	0.98
Client 7	39	22	0.58	0.8	20	15	0.94	0.96	87	49	0.58	0.84
Client 8	73	40	0.64	0.81	175	92	0.58	0.85	31	20	0.67	0.83
Client 17	75	47	0.64	0.92	62	38	0.67	0.91	251	97	0.44	0.9
Client 19	433	167	0.41	0.92	488	162	0.36	0.91	--	--	--	--
Client 20	403	158	0.39	0.94	102	57	0.56	0.91	598	186	0.317	0.91
Mean	226	81.22	0.54	0.87	138.9	55.45	0.61	0.86	171.8	53.11	0.57	0.84

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Stroke Story discourse task for MIAP 2019

Table D7

Number of Different Words for Letter Procedural Task

	PRE MIAP				POST MIAP				FOLLOW UP			
	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR	TNW	NDW	TTR	MATTR
Client 1	158	71	0.47	0.95	70	38	0.59	0.93	--	--	--	--
Client 2	--	--	--	--	8	6	1	--	16	14	0.88	0.97
Client 5	111	54	0.56	0.91	48	25	0.65	0.89	92	56	0.62	0.93
Client 7	5	5	1	--	5	5	1	--	29	18	0.62	0.73
Client 8	26	21	0.88	0.94	23	17	0.77	0.86	12	11	0.92	0.97
Client 17	81	44	0.57	0.92	112	53	0.49	0.88	77	32	0.49	0.89
Client 19	145	71	0.52	0.89	106	50	0.53	0.91	--	--	--	--
Client 20	73	45	0.62	0.9	28	15	0.58	0.68	102	51	0.53	0.95
Mean	85.57	44.43	0.66	0.92	50	26.13	0.7	0.86	54.67	30.33	0.67	0.9

Total number of words, number of different words, type-token ratio and moving average type-token ratio for Letter discourse task for MIAP 2019