Using Data from Comprehension and Production to Test Competing Theories in Sentence Impairments in Individuals with Aphasia

by

Megan Ahern

B.A., Communication Science and Disorders, University of Pittsburgh, 2017

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This thesis was presented

by

Megan Ahern

It was defended on

March 22, 2019

and approved by

Wiltrud Fassbinder, Ph.D., Adjunct Assistant Professor, Communication Science and Disorders, VA Pittsburgh Healthcare System

Bharath Chandrasekaran, Ph.D., Vice Chair for Research and Professor, Communication Science and Disorders

Erin Lundblom, Ph.D., CCC-SLP, Coordinator of Clinical Education and Assistant Professor, Communication Science and Disorders

Thesis Advisor: Michael Walsh Dickey, Ph.D., Director of PhD Program and Associate Professor, Communication Science and Disorders, VA Pittsburgh Healthcare System

Using Data from Comprehension and Production to Test Competing Theories in Sentence Impairments in Individuals with Aphasia

Megan Ahern, B.A.

University of Pittsburgh, 2017

The ability to use language is essential for the communication of ideas, wants and needs. It is through language that human beings are able to participate in meaningful exchanges from having a conversation with a friend or family member, to understanding detailed instructions an employer, to accessing important information regarding personal health and wellness. Language is necessary to share information as well as request it. It is estimated that nearly 180,000 Americans acquire aphasia each year (NIDC, 2018), causing potentially devastating deficits in their ability to use language. In order to assess and treat individuals with aphasia (IWA), it is important to have an understanding of the linguistic deficits that occur. The current study aims to investigate syntactic deficits in the linguistic production and comprehension of 99 IWA. Two competing theories (Resource Reduction Theory and Specific Impairment Theory) are evaluated based on the results and conclusions of the current study. Research questions in the current study ask how several variables (canonicity, modality, severity and length) impact or predict participant performance on comprehension tasks (SOAP; Love & Oster) and on production tasks (NAT; Thompson, Weintraub & Mesulam, 2012): canonicity, modality, severity, and length. Findings suggest that increasing linguistic complexity in the presence of linguistic deficits results in poor performance in both production and comprehension across the variables outlined in the research questions. Overall, findings of the current study are more consistent with Resource Reduction Theory than Specific Impairment Theory.

Table of Contents

List of Tables

List of Figures

Preface

I would like to thank the members of my thesis committee for helping me learn more about an area and population I am passionate about. Thank you, Dr. Dickey for serving as my thesis advisor and for encouraging my enthusiasm for linguistics and research.

Thank you, Dr. Fassbinder for the endless time and patience spent supporting my learning beyond classroom instruction, and for the countless hours spent helping me through data analysis.

Thank you, Dr. Chandrasekaran and Dr. Lundblom for supporting this project and for being a part of my thesis committee.

I would like to thank my family, especially my parents, for always believing in me and reminding me of my ability to accomplish the goals I set for myself. Thank you also for always pushing me to pursue opportunities to better myself.

1.0 Introduction

Aphasia is an acquired communication deficit that affects the four domains of language: reading, writing, speaking and listening. Aphasia is usually caused by neurological insult, such as stroke or traumatic brain injury, in the left cerebral hemisphere (Gialanella, Prometti, Vanoglio, Comini and Santoro, 2016, p. 782). Aphasia can manifest itself in a variety of ways depending on the severity and location of the brain injury that the individual has suffered. While spontaneous recovery can occur in individuals with aphasia (IWA), it is likely that many language difficulties will remain a permanent part of their life (Gialanella et al., 2016). The World Health Organization's International Classification of Functioning, Disability and Health (ICF) has influenced the way aphasia is conceptualized and defined. The ICF defines disability through multiple dimensions including "body structure and function, activities and participation, and personal and environmental context" (Simmons-Mackie and Kagan, 2007, p. 244). Difficulty with language, in either production or comprehension, can have a tremendous impact on the life of IWA because of the unique nature of a linguistically-rooted disability. Aphasia is not a disability that impacts or reduces the individual's intelligence, although this is a common misconception about the disorder. Instead, communication deficits in IWA can mean problems conveying ideas, wants and needs, which are functions of language most individuals do not think twice about. Experiencing neurological insult that disrupts these processes can be devastating for IWA. Understanding the way aphasia is impacting a patient can be challenging because of the variability of the disorder and the limited access IWA have to communication. When examining aphasia in

regard to the ICF dimension of activities and participation, one can see how aphasia has impacts on everyday life. Being unable to talk to or understand friends, family and loved ones is a social burden. This social burden can isolate IWA by reducing participation in activities of daily living, leisure activities, and interaction with communication partners. IWA may also become unemployed as a result of their disability (Simmons-Mackie and Kagan, 2007, p. 246). Similarly, these language deficits can prevent IWA from returning to work or managing responsibilities independently. It is important that professionals are able to understand the nature of the linguistic deficits caused by aphasia, as well as the vast implications it has on a patient when treating IWA in order to offer the best possible care and services.

The majority of IWA are "middle-aged or older" however individuals of any age can acquire aphasia. NIDCD (2018) estimates that "1 million people in the United States" have aphasia and that each year, and that almost "180,000 Americans acquire it each year according to the National Aphasia Association."

IWA can be classified as fluent or non-fluent. There are four major types of fluent aphasia, where Wernicke's aphasia and conduction aphasia are the most commonly assigned diagnostic terms, and transcortical sensory aphasia and anomia are less commonly assigned (Edwards, 2005). The linguistic profile and abilities of IWA is highly variable, although there are patterns of symptoms that are commonly observed for the aphasia types discussed. In a highly simplified breakdown of these classifications, an individual with Wernicke's aphasia would be expected to have relatively poor comprehension abilities and what may sound like fluent speech output, but consist of very little meaningful content. As such, some refer to Wernicke's aphasia as jargon aphasia because of the tendency of these individuals to use nonwords/neologisms or paraphasias, where one word is used instead of the target word (Hallowell, 2017 p.160). Repetitions may consist

of jargon rather than target words. In addition, Wernicke's aphasia is commonly associated with receptive deficits when listening to others and when reading (Hallowell, 2017, p. 162). An individual with conduction aphasia would be expected to have good comprehension abilities but relatively poor repetition abilities as a target becomes longer. While conduction aphasia is considered by some to be a "fluent aphasia," it is important to consider the task that an IWA was engaged in when assigning this classification. For example, being unable to repeat a complex target sentence would make an individual with conduction aphasia sound non-fluent, but the same individual may be able to participate in conversation in a way that appears "fluent" and uninterrupted, but conveys little meaning (Hallowell, 2017, p. 155). Conduction aphasia can also be associated with mild to moderate receptive deficits in auditory listening and when reading (Hallowell, 2017, p.157). An individual with transcortical sensory aphasia would be expected to have poor comprehension abilities and good repetition relative to other abilities, and fluent-sounding speech output. Transcortical sensory aphasia often presents in a similar manner to Wernicke's aphasia, but with superior repetition abilities and similarly impaired comprehension.

There are two or three major types of non-fluent aphasia, where Broca's aphasia is the most commonly assigned diagnostic term for individuals with lesions in "the pre-Rolandic area of the left cerebral cortex" and who have reduced abilities to produce speech and use grammar and vocabulary (Edwards, 2005, p.8). Auditory comprehension for individuals with Broca's aphasia decreases when the grammatical complexity of targets increases. Broca's aphasia is commonly associated with agrammatism, and some use this term to describe individuals with Broca's aphasia. Others assign the term agrammatism to a subgroup of individuals with Broca's aphasia. These individuals have deficits in production as well as comprehension of sentences with "certain syntactic structures, specifically when sentence constituents have been moved" from their original

position, (as discussed in section 2.1) (Edwards, 2005, p. 8). Speech output in individuals with Broca's aphasia can be considered non-fluent in that it is often comprised of short "telegraphic" phrases that do not include function words, like auxiliary verbs and pronouns. Instead, these simple phrases consist of content words (i.e., verbs, nouns and adjectives) that convey more semantic information (Hallowell, 2017, p. 162-3).

In a highly simplified breakdown of the other non-fluent aphasia classifications, an individual with transcortical motor aphasia would be expected to have some intact comprehension abilities, where auditory comprehension of grammatically complex targets decreases in a similar way to as is seen in Broca's aphasia. Repetition abilities are also severely impaired in transcortical motor aphasia. Compared to other deficits, transcortical motor aphasia is also characterized by relatively good repetition abilities (Hallowell, 2017, p. 163).

An individual with global aphasia would be expected to have severe comprehension deficits, severely impaired repetition abilities, and non-fluent or absent speech output. Global aphasia can, in some cases, be characterized by stereotypy, which occurs when an individual uses one word for everything they say (Hallowell, 2017, p. 165).

Mixed transcortical aphasia is another, more rare classification that can be assigned to IWA. An individual with mixed transcortical aphasia would be expected to have poor auditory comprehension abilities, severely non-fluent speech output. Mixed transcortical sensory aphasia can present similarly to global aphasia, where the distinction lies in the individual's preserved repetition abilities (Hallowell, 2017, p.166). Across aphasia classifications, individuals may have varying degrees of difficulty in their spoken production and their auditory or written comprehension of sentence (See Table 1).

Table 1. Aphasia Classifications

	Comprehension	Repetition	Speech Output	Common Lesion Site
Wernicke's Aphasia	Moderately-severely impaired	Poor; may be jargon or nonword-based	Fluent but without semantic meaning	Posterior superior temporal gyrus; BA 22
Conduction Aphasia	Mildly-moderately impaired	Moderately-severely impaired; worsens as target length increases	May be fluent or non-fluent depending on severity and speech task observed	Arcuate fasciculus, supramarginal gyrus; BA 40
Transcortical Sensory Aphasia	Moderately-severely impaired	Intact	Fluent with nonwords and/or paraphasias possible	Temporal lobe; angular gyrus (BA 39); posterior middle temporal gyrus (BA 37)
Anomia	Intact	Intact	Fluent with word-finding difficulty	Angular gyrus
Broca's Aphasia	Mildly-moderately impaired; worsens as grammatical complexity increases	Severely impaired	Severely impaired; non-fluent; telegraphic	Pre-Rolandic area of left hemisphere; inferior frontal lobe; BA 44 and BA 43
Transcortical Motor Aphasia	Mild-moderately impaired	Intact	Severely impaired; non-fluent	Anterior portions of left frontal lobe and prefrontal

Table 1. Aphasia Classifications

				areas; (BA 6, 8, 9, 10, and 46)
Global Aphasia	Severely impaired	Severely impaired	Severely impaired; non-fluent or absent; stereotypy	Perisylvian lesion in multiple areas of frontal, parietal and temporal lobes
Mixed Transcortical Aphasia	Severely impaired	Relatively intact	Severely impaired	Inferior frontal lobe; regions associated with transcortical motor and/or transcortical sensory aphasia

BA= Brodmann's Areas

Difficulty in production and comprehension of sentences appears most clearly in more complex sentence structures, which will be discussed in section 2.1. This increased difficulty with increased syntactic complexity may also be more common among individuals with non-fluent aphasia types (Edwards, 2005). Deficits in producing and comprehending sentences can have important consequences for individuals' ability to function in everyday settings. For example, failing to understand who to talk to about one's medication or a medical diagnosis could have life-threatening consequences. IWA may also have difficulty in asking or answering Wh-questions (like who- or what- questions), which are complex sentence types that involve WH-movement (see section 2.1.2 for discussion of movement). These kinds of questions are regularly used by individuals to accrue and communicate information. Such linguistic deficits can therefore

negatively affect the quality of life of IWA while putting them at risk for serious ramifications associated with misunderstanding or being unable to access language and information.

The prevalence of deficits in the comprehension and/or production of language among IWA is not well-established in the literature. However, prior work demonstrates that irrespective of fluency, IWA may have difficulty in sentence production. Relevant data comes from Caplan and collogues' 2007 study examining sentence comprehension and production performance in a large group of IWA. They compared the performance of two groups of participants: "aphasic patients with left hemisphere strokes" (n=42) and healthy controls (n=25). Patients ranged in age from "24.7 to 84.5 years (mean 60.3 years)," were English speakers with a diagnosed language impairment recruited from Boston hospitals (Caplan et al., 2007, p. 7). There were "16 female and 26 male" patients, and years of education ranged from "9 to 22 years (mean 14.7 years)" (Caplan et al., 2007, p. 7). Although this sample cannot truly be considered randomly selected, the information regarding prevalence of sentence-level deficits in the participants and controls of this study is useful when considering questions of prevalence, nature and impact of sentence-level impairments in IWA. In order to classify participants in the study, a variety of assessment tasks were used, one of which was the "oral production subtests of the Cross-Modal Morphosyntax Battery" (Caplan et al., as cited in Goodglass, Christiansen, & Gallagher, 1993), which assessed production of morphosyntactic structures. Sentences tested included different morphology types of noun phrases, such as "singular/plural, noun possession, and adjective possession," and five different verb features, such as "3rd person present tense, verb tenses, auxiliary verbs, low-content verbs, and auxiliary verbs with complements" (Caplan et al., 2007, p. 112-1). In additional subtests, the participants' production of active and passive voice was also assessed. Participants looked at a picture and listened to a verbal prompt that explained the image to provide context

related to the structures being assessed. Participants were then required to finish a lead-in prompt with one- or two-word responses or with a full sentence. Performance was scored as being correct, if responses contained the targeted grammatical morpheme/element, or incorrect. Incorrect responses were broken down by error types: substitutions (response included an incorrect grammatical element), or omissions (response did not include a grammatical element) (Caplan et al., 2007). The participants in the study were classified using the same accuracy cutoffs for the seven subtests of the morphosyntactic battery used and demonstrated to distinguish Broca's and fluent IWA in Goodglass et al., (1993). Based on the Goodglass et al. classifications, six participants in the Caplan et al., (2007) "met Goodglass et al. criteria for agrammatism" and were "classified as Broca's aphasics," and 11 participants "were classified as fluent aphasics" and also met these criteria for agrammatism (Caplan et al., 2007, p. 142). This finding suggests that both fluent and non-fluent IWA may have difficulty with producing sentences.

In a preliminary test of sentence comprehension, a Sentence-Picture Verification (SPV) task, consisting of "one hundred semantically reversible sentences" was used and included the following syntactic structures: "active, active with conjoined theme, dative, passive, truncated passive, dative passive, cleft object, conjoined, object-subject and subject-object" (Caplan et al., 2007, p. 112). Semantically reversible sentences are important in regard to the current study because IWA with agrammatism have deficits in comprehending this type of sentence (see section 2.1 below). Accuracy for these structures were reported for Broca's aphasics and fluent aphasics. Accuracies are as follows: active (Broca's= 68.6% SD=19.0, fluent= 93.1% SD 13.8), subject object (Broca's= 60.0% SD=11.0, fluent= 93.1% SD=21.4), passive (Broca's= 58.3% SD=26.4, fluent= 93.1% SD=12.8), object subject (Broca's= 61.2% SD=11.3, fluent= 79.2% SD=18.9), all sentences (Broca's= 59.5% SD=15.5, fluent= 84.5% SD=12.8) (Caplan et al., 2007, p. 146).

Furthermore, averaging across all structures tested on the SPV task, 31 out of 42 of the IWA performed outside the range of normal performance on this task. Caplan et al. based "normal performance" on the performance of a set of healthy age-matched controls. All these IWA performed within normal limits on auditory word-picture matching tests, a test of single-word comprehension. This means that 31 out of 42 IWA in this sample (around 73%) have deficits in sentence comprehension as measured by the SPV task.

This performance data is useful as it suggests that although IWA classified as fluent performed better overall, performance on the noncanonical structures (passives and object subjects) was poorer than on the canonical structures (actives, subject objects) for both fluent and non-fluent (Broca's) IWA. These findings indicate that both fluent and non-fluent IWA may have deficits in the production and comprehension of sentences. This conclusion is consistent with Edwards (2005), who reports that sentence comprehension and production deficits are also found in individuals with fluent aphasia, even though these individuals are not commonly diagnosed with having agrammatism. These findings suggest that deficits in production and comprehension of complex sentences are common in aphasia and can be found in a wide variety of IWA.

1.1 Competing Theories of Syntactic Deficits

Two theories of syntactic deficits exist in the literature and are addressed in the current study. The first suggests that IWA have difficulties producing and/or comprehending language as a result of a general reduction in the previously intact linguistic resources of an IWA prior to their injury. This is referred to as Resource Reduction Theory. The second theory suggests that IWA have difficulties producing and/or comprehending language as a result of a specific deficit that

prevents the use of linguistic operations required for certain sentence types. This is referred to as Specific Impairment Theory.

In this thesis, I will examine the performance of 99 IWA on the Subjective Objective Active Passive assessment tool (SOAP; Love & Oster, 2002) and the Northwestern Anagram Test (NAT; Thompson, Weintraub & Mesulam, 2012) to examine predictions made based on the Resource Reduction Theory and the Specific Impairment Theory, which will both be described further in section 2.2 of this paper. Data was collected as part of a larger study through the VA Pittsburgh Healthcare System (VAPHS) at 5 sites: VAPHS, the VA Northern California Health Care System (VANCHS), the VA Puget Sound Health Care System (VAPSHCS), the University of Washington, and Temple University (see section 3.0 for discussion).

2.0 Background

2.1 Syntactic Complexity in Aphasia

Agrammatism can be a major source of communication difficulties faced by IWA. Agrammatism is a characteristic seen frequently in non-fluent Broca's aphasia, and there is some disparity in the way the term is used in the literature and by clinicians, as mentioned previously. It is useful to consider the following concept from Edwards (2005): all individuals with agrammatism have Broca's aphasia, but all individuals with Broca's aphasia are not necessarily agrammatic.

Agrammatism is a characteristic speech pattern in which IWA produce grammatically incorrect sentences in which syntactic structure is reduced or not present. Agrammatism is also commonly associated with more difficulty using verbs than nouns; individuals may convey ideas using short strings of content words without function words necessary to connect them together. (Thompson, C.K., as cited in Chapey, 2008). Further, when these individuals use constructions that include verbs, it is typically done in canonical word order, which is subject, object verb (SVO) in English. More complex sentences that do not use SVO word order are referred to as noncanonical. There are treatment approaches for agrammatism which are centered around the idea that the deficit is in connecting the underlying structures that convey meaning to their correct positions in the surface-level syntax: this refers to the mapping hypothesis (Schwartz, Saffran, Fink, Myers & Martin, 1994). The mapping hypothesis is relevant to the concept of asyntactic comprehension, which refers to the inability to understand semantically reversible, noncanonical sentences such that meaning cannot be gleaned from the lexical content (Schwartz et al., 1994). While IWA can have better comprehension abilities when compared to verbal production, they

often demonstrate difficulty in both modalities when it comes to these semantically reversible sentences (Berndt, 1987; Berndt, Mitchum, & Haendiges, 1996; Caplan & Futter, 1986; Caplan, Baker, & Dehaut, 1985; Caramazza & Zurif, 1976; Grodzinsky, 1986; Saffran & Schwartz, 1988; Schwartz, Linebarger, & Saffran, 1985) as referenced in Rochon, Laird, Bose, and Scofield (2005). Semantically reversible sentences are those that contain multiple possible "actors" that could be impacted by the verb acting in the sentence (Rochon et al., 2005). For example, in the sentence "*The man is hugging the woman"* both potential actors (the man and the woman) could logically be an agent capable of performing the action (hugging). They could also both logically be the patient that is being acted upon (Rochon et al., 2005). To correctly interpret this reversible sentence, one has to be able to differentiate who is doing what by assigning thematic roles and then connecting them to the right position in the sentence. The question becomes, who is hugging and who is being hugged? If correct interpretation of the sentence would occur when "man" is assigned the role of agent and "woman" is assigned the role of patient. For individuals with aphasia, these kinds of reversible syntactic constructions can be problematic.

The six syntactic structures in the current study are: actives (A), passives (P), subject-relatives (SR), object-relatives (OR), subject Wh-questions (SWhQ) and object Wh-questions (OWhQ).

Active sentences are constructed such that the subject of the sentence is acting upon or doing something to another entity. For example, "the boy kissed the girl."

Passive sentences are constructed such that the subject of the sentence is being acted upon by another entity. For example, "the girl was kissed by the boy."

Subject relative clauses occur when "the head noun occupies the subject role in the relative clause." For example, "this is the dog [that chased the cat]" is a sentence where the head noun, (the

dog) is the subject for the relative clause, which is in brackets (Kidd, Brandt, Lieven, & Tomasello, 2007, p.861).

Object relative clauses occur when the head noun "occupies the object role in the relative clause." For example, "this is the dog [that the cat chased]" is a sentence where the head noun, (the dog) is the object for the relative clause, which is in brackets (Kidd et al., 2007).

Subject Wh-Questions are constructed such that "the entity being questioned is the grammatical subject of the verb" (Seidl, Hollich, & Jusczyk, 2003, p. 425). For example, "who kissed the girl?"

Object Wh-Questions are constructed such that "the entity being questioned is the grammatical object of the verb" (Seidl, Hollich, & Jusczyk, 2003, p. 425). For example, "who did the boy kiss?

In the current study, various canonical and noncanonical sentence structures will be compared to examine participant performance in both comprehension and production. For the purposes of this paper, the term *"word order category"* will be used to describe one way the structures can be grouped: as canonical (A, SR and SWhQ, which have SVO word order) and as noncanonical (P, OR, and OWhQ, which do not have SVO word order). The term "*sentence type"* will be used to describe another way the structures can be grouped: as relatives (SR, OR), as declaratives (A, P) or as Wh-Questions (SWhQ, OWhQ). The relative sentences on the NAT are subject-clefts and object clefts, but will be referred to as SR and OR in this paper to maintain consistency with the relative sentence types on the SOAP. The term *"movement type"* will be used to describe the final way the noncanonincal structures will be grouped: those with NP-movement (P) and those with WH-movement (OR, OWhQ). Examples of syntactic structures used on the SOAP and the NAT can be found in Appendix A and Appendix B, respectively.

When examining the theories and patterns of impairments discussed in the current study, it is helpful to group the sentence types into control-experimental pairs. The control structures (A, SR, SWhQ) have canonical word order and no trace movement. The experimental structures (P, OR, OWhQ) have noncanonical word order and trace movement of either movement type. The canonical trace-free sentence types can be thought of as control structures for the experimental counterparts of the same sentence type. The following pairs of control and experimental structures are discussed and illustrated in Table 2: actives and passives, object relatives and subject relatives, object WH-questions and subject WH-questions. Pairing structures on the NAT and the SOAP allows for comparison and facilitates questions about performance through the lens of both theories.

Table 2. Control Experimental Structures

	NP-Movement	WH-Movement
1 Clause	Actives → Passives	Subject Wh-Questions → Object Wh-Questions
2 Clauses		Subject Relatives → Object Relatives

Control structures for each pair are in blue; Experimental structures for each pair are in green; Structure pairings are represented by the arrow which points from simple structure to complex structure

2.1.1 Canonicity (Word Order Category)

As mentioned before, in English, canonical refers to SVO word order, which is found in the simplest sentence structures. Noncanonical refers to structures that deviate from the typical SVO word order and are more difficult. The active sentence, *"the boy kissed the girl"* is canonical. The passive sentence, *"the girl was kissed by the boy"* is noncanonical. Producing and comprehending sentences with increased complexity due to word order can be particularly challenging for IWA. This challenge arises as a result of the syntactic movement of sentence constituents that occurs in noncanonical sentences. Deficits or impairments in comprehension and production of noncanonical sentences can be found in many types of aphasia, as mentioned before in the discussion of aphasia classifications (Edwards, 2005).

2.1.2 Movement Type

Two types of syntactic movement exist in noncanonical sentences: NP-movement (seen in passives) and WH-movement (seen in ORs and ObWhQ) (Jacobs & Thompson, 2000). These two movement types are alike in that they both leave behind a "trace" when the moved constituent is placed in a new spot (Jacobs & Thompson, 2000). The landing spot of this constituent is how NP- and WH-movements are distinguished.

In NP-movement, a NP is moved to a new argument position, which means it still receives a thematic role. For example, the sentence, *"Joelle was kissed by Dillon"* (P) is constructed through NP movement in the underlying form *"Dillon kissed Joelle"* (A) (Shapiro, 1997, p. 14).

In WH-movement, a constituent that was previously assigned a thematic role is moved to a nonargument position, which it is not assigned a thematic role as an argument in the sentence.

For example, the object relative sentence, *"it was Joelle who Dillon kissed"* is constructed through WH-movement in the underlying form (an ObWhQ) *"who did Dillon kiss?"* (Shapiro, 1997, p. 14).

NP-movement places the NP in a new argument position within a sentence that, like its underlying form, has one clause. WH-movement places the moved constituent to a nonargument position and crosses clausal boundaries. While both movement types leave a trace behind, WH-movement structures "require more processing resources than NP-movement structures" because it involves moving sentence constituents "across clausal boundaries" leaving the moved constituent further from its original position (Jacobs & Thompson, 2000, p. 3).

2.1.3 Modality

As described in section 2.1, many IWA may have difficulty with both comprehension and production of complex, noncanonical sentences. However, there is also evidence that one modality (production) may be more difficult than the other (comprehension) (Schroder, Burchert and Stadie, 2014). In a study that investigated the cognitive processes that underlie production and comprehension, aphasic participant performance corresponded to a pattern of classical dissociation. Comprehension of OWhQ was intact in the presence of impaired production of the same OWhQ structures (Schroder et al., 2014). The researchers examined performance on ORs as well as OWhQ and found that they were not impaired in the same way across modalities. The dissociating patterns of better performance in comprehension than production was seen in five out of seven participants for both OWhQs and ORs. However, for two of the participants their performance reflected the classical dissociation described, where comprehension of OWhQs was spared, while production of OWhQs was significantly worse (Schroder et al., 2014). These

findings support the argument that the underlying syntactic processes required to build sentences are modality-specific. Impairments in either comprehension or production of noncanonincal structures cannot be attributed to a shared underlying impairment (Schroder et al., 2014). Interestingly, two participants in the study had impaired performance on syntactically more complex structures in both modalities, but impaired performance in only production on less complex structures. The results presented by Schroder et al. (2014), provide evidence that suggests both modalities can be affected by a deficit without being affected to the same degree.

Turning to language acquisition patterns in typically developing children is relevant and may be helpful in the present discussion. It is widely known that receptive language has an earlier developmental trajectory than expressive language. A child's receptive lexicon emerges earlier and is larger than their expressive vocabulary, which holds true into adulthood (Paul, Norbury & Gosse, 2018). The way in which receptive language abilities precede and surpass expressive language abilities can inform predictions in the current study by illustrating how comprehension tasks will likely be less difficult than production tasks. In addition to faster acquisition, the receptive vocabulary of a typical speaker is also larger than their expressive vocabulary. This is another factor to consider when examining how abilities become impaired in a language deficit and how this can impact performance in the modalities.

Consider how language production can be more challenging than comprehension for healthy individuals. For example, if you hear someone use a word in conversation, you may recognize it upon hearing it and receptively understand what the speaker is talking about, even if it is not a word you use and/or hear frequently. However, if you saw the word depicted in a picture without context and were asked to define or explain it, you may have difficulty. Instead of simply having to recognize the word and think about its meaning in a context, production would require

you to also select and retrieve the definition from your expressive lexicon, if you know the word to begin with. Taking this a step further, having to talk about the word (as opposed to listening to someone else and comprehending what they are saying) introduces another layer of difficulty associated with the ability to utilize correct syntax to explain the semantic meaning. Considering these ideas about typical language comprehension and production, it is unsurprising that IWA will demonstrate more difficulty with production tasks. Schroder et al. (2014) point out that IWA performed poorly on both comprehension and production when it came to more complex structures. The fact that less complex structures were only impaired in production resembles the example above where a typical speaker may be able to understand a word/concept (receptive language use) but still struggle to talk about it (expressive language use).

2.1.4 Length

Another factor that can make sentences more complex is length. Longer sentences are more complex than shorter sentences. For the purposes of the current study, length refers to the number of clauses within a sentence. In order to examine how complexity may impact participant performance, the structures being examined will be characterized by length, where longer structures are more complex than shorter structures. ORs and SRs have two clauses, and all other structures discussed (A, P SWhQ, OWhQ) have one clause.

2.2 Theories of Sentence Impairments in IWA

In order to study the patterns of sentence production and comprehension seen in IWA, two syntactic theories have been used to explain the deficit patterns observed in the literature.

2.2.1 Resource Reduction Theory

Resource Reduction Theory argues that errors in linguistic comprehension and production seen in IWA are the result of reduced resources which, when exhausted, results in the inability to perform the all necessary operations required to connect meaning to structure in complex sentences (Caplan, Waters, DeDe, Michaud and Reddy, 2007). For the purposes of the current study, resources can be defined as the cognitive reserve an individual draws upon in order to process linguistic information. IWA may have limited processing resources, which refers to relevant cognitive abilities such as: "verbal working memory, phonological store, and short-term semantic memory" (Des Roches, Vallila-Rohter, Villard, Tripodis, Caplan, & Kiran, 2016, p. 744). By this logic, one would expect participant performance to deteriorate as structures become more complex or for performance to be better on structures or operations that occur in isolation. This prediction stems from the idea that the IWA will be able to handle and process linguistic information, but only up to a certain point as a result of the reduced resources they can allocate to the task. It would not be expected for IWA to make the same kinds of errors on specific kinds of structures or in specific operations. Rather, IWA performance is more likely predicted by complexity of the structures or on their severity and the requirement to allocate more resources than they possess.

If participants are operating under the limitations of their reduced linguistic resources, disruptions in normal mental processes may result in the inability or reduced ability to comprehend

and/or produce the meaning of sentences. Four relevant arguments in support of this theory are outlined by Caplan et al., 2007:

1. Patients can process sentences with isolated operations or structures, as opposed to sentences that use combinations the same components (Caplan and Hildebrandt, 1988; Hildebrandt et al., 1987).

2. Deteriorating participant performance due to reduced resources, has a greater negative impact on more complex sentences than on less complex sentences (Caplan et al., 1985, 1996).

3. Performance is based on the nature of factors used "in factor analyses and principle components analyses, which have yielded first factors that account for the majority of the variance in syntactic comprehension tasks, on which all sentence types load" (Caplan et al., 1985, 1996 as cited in Caplan et al., 2007, p. 106). Reduced resources are considered a crucial factor.

4. In normal individuals "concurrent load, speed of stimulus presentation, perceptual degradation or masking of input" that are assumed to impact ability to complete tasks result in poor participant that mirrors what would be expected of an IWA (Miyake et al., 1994; Dick et al., 2001; Kilborne, 1991 as cited in Caplan, 2007, p. 106).

In regard to the first and second arguments, less severe participants in the current study should be expected to perform better than more severe participants on items with increased complexity (noncanonicals, longer structures). Based on these arguments and data, the current study examines the extent to which participants with mild aphasia perform better than those with moderate and severe aphasia on noncanonincal items, and potentially canonical items as well.

Factor analysis findings by Caplan and colleagues 2007 regarding comprehension of three syntactically contrasted structures are consistent with the third argument. Like previous findings that demonstrate the that the most important component when considering performance is a

process that impacts all sentence types to the same degree, Caplan et al., 2007 found the same factor influenced end-of-sentence RTs for participants as well healthy controls. Based on this information, the current study asks performance on all items demonstrate a similar pattern, regardless of participant severity. Resource Reduction Theory predicts that severity and linguistic complexity of items will predict performance.

2.2.2 Specific Impairment Theory

Specific Impairment Theory argues that IWA are missing or unable to access a specific function that is required in order to process a component or procedure for a given sentence type (Caplan et al., 2007). By this logic, one would expect IWA to have errors on structures that require a specific linguistic operation or use of a particular constituent found in difficult sentences. Further, it would need to be demonstrated that the observed errors are not the result of another, coexisting impairment. Chomsky's 1957, 1965, 1981, 1986, 1995 model of syntax (as cited in Thompson, Tait, Ballard and Fix, 1999) can elucidate the ideas behind this theory, as the model suggests that difficulty in syntactic construction arises due to the movement of a relative clause from its original position to a new one in noncanonical sentences. When this occurs, the movement leaves a "trace" behind where the head of the relative clause would be in the canonical version of the sentence. This is relevant to Grodzinsky's Trace Deletion Hypothesis (THD; 1984, 1986), which states: "traces are deleted from agrammatic aphasic subject's syntactic representations," leaving the individual unable to assign thematic meaning to the "trace-antecedent chain" (as cited in Thompson et al., 1999, p. 170). Under this syntactic theory, noncanonical sentences all involve syntactic movement and a trace-antecedent chain.

If IWA interpret sentences using a predetermined strategy for interpretation that assigns the initial NP as the agent (THD; Grodzinsky, 1984, 1986), one would expect IWA to struggle with comprehension of sentences that do not share this syntactic construction. In the current study, intact participant comprehension of active and subject-relative clause items and impaired comprehension of passive and object relative clause items on the SOAP is predicted. This will be demonstrated through better performance, in particular, higher accuracy, on actives than on passives, and on subject-relatives than on object relatives. This prediction is based on the fact that A and SR sentences are canonical structures, which means they do not contain trace movement and are therefore less complex. For example, consider the following experimental items on the SOAP, which can be found in the complete list of SOAP items in Appendix A:

4. *"The young boy with the brown hair grabs the man"* (A)

22. *"The man that grabs the little boy has brown hair"* (SR)

6. *"The boy with the brown hair is grabbed by the man"* (P)

10. *"The man that the young boy grabs has brown hair"* (OR)

The canonical sentence in item 4 (A) should elicit better performance than the noncanonical sentence in item 6 (P). The canonical sentence in item 22 (SR) should elicit better performance than the noncanonical sentence in item 10 (OR). Errors made by participants may demonstrate a pattern where selection of the incorrect mismatch picture is made instead of the target. This may occur because the syntax in P and OR items does not place agents before the verb that assigns their thematic role. This would leave a trace in the final position, creating ambiguity for an IWA where word order will not be enough to differentiate the agent (*"young boy"* in item 10) from the patient (*"the man"* in item 10). Thus, it is reasonable to expect the participants to select the mismatch condition for these items. This logic predicts the same difficulties in comprehension to arise on

production of items of the NAT, where experimental structures would be expected to elicit poorer performance than their controls.

Specific Impairment Theory does not argue that noncanonical sentence impairments are caused by overall aphasia severity. Similarities between comprehension and production of canonical and noncanonical structures by groups of mild, moderate, and severe participants would be consistent with Specific Impairment Theory. The performance of severe expressive IWA, who received a score less than 3 on the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1972), on the SOAP was the same as performance of mild expressive aphasics (BDAE score over 3) on noncanonical items. Both groups logically were able to comprehend canonical structures, but not noncanonical ones, and when presented with sentences that contained moved components, group performance dropped to chance level (Love & Oster, 2002). This demonstrates the SOAP's sensitivity to the mild expressive group's less obvious deficits in in comprehension, which the BDAE would not predict. These findings are consistent with Specific Impairment Theory because they suggest that IWA will all perform poorly on noncanonical sentences, even if they are considered mildly impaired.

2.3 Implications of the Theories

2.3.1 Task-Related Factors/Implications

2.3.1.1 Canonicity

In regard to performance on noncanonical items, Resource Reduction Theory predicts that in a comparison of canonical and noncanonical structures, participants will make more errors on noncanonical structures because they require more linguistic and/or cognitive resources to produce and comprehend them. According to this theory, errors made are not rooted in the syntax of the structures or the underlying linguistic systems IWA use to mentally process them, but in the limited resources IWA have available when using said system.

Resource Reduction theory does not require that *all* noncanonical structures will be impaired, as this would rule out the possibility that a mild participant may be able to comprehend or produce some noncanonical items. For example, a mild participant may perform better on noncanonical passives than the canonical, but more complex SRs. A scenario where performance on noncanonical items is spared in the presence of poor canonical items would be a finding that is in opposition to Specific Impairment Theory.

Specific Impairment Theory predicts that participants will perform more poorly (lower accuracy) on the noncanonical experimental structures than on their canonical control parings. That is, passives, ORs and OWhQs should elicit poorer performance than actives, SRs and SWhQs, respectively. A deficit in the specific process or operation required to produce and/or comprehend the trace movement in each of the experimental structures (P, ORs, OWhQs) should prevent participants from responding correctly on test items. While the two theories make similar predictions, Specific Impairment Theory assumes that all participants with a specific canonicity-deficit should be consistent in their errors. This would be demonstrated incorrect responses on all noncanonical structures, regardless of individual severity levels, length of structures and modality.

2.3.1.2 Movement Type

When comparing items based on their trace movement types (WH-movement vs. NP-movement) similar predictions can be made for each theory. Resource Reduction Theory would predict that when comparing items based on their movement types (WH-movement vs. NP-movement), more severe participants will perform more poorly than moderate and mild participants on all structures. However, Resource Reduction Theory also predicts that WH-movement may have a more prominent pattern of errors than NP-movement. While both movement types leave behind a trace, WH-movement occurs across the clausal boundaries of a sentence, making it more difficult and requiring more linguistic resources than NP-movement (Jacobs & Thompson, 2000). ORs, OWhQs and passives are all noncanonical, however ORs and OWhQs require WH-movement, and passives require NP-movement. In the current study, Resource Reduction Theory predicts poorer performance on items with ORs and OWhQs than on items with passives because of the increased resources needed for computation of WH-movement.

Specific Impairment Theory predicts that if participants with a specific impairment exist, they will consistently perform poorly on structures of the same type, regardless of their severity. This prediction would be supported if a pattern of errors in either NP- or WH-movement is seen. If there is a difference in performance between these two structures, Specific Impairment Theory predicts that the errors will be consistent, such that if one movement type is impaired, it will result in poor performance on all structures with that movement. This theory posits that impairment can be movement-specific. However, it does not predict which movement type will be impaired or suggest that participants may have an impairment in *only* WH- or NP- movements. So, if NP-movement is the specific impairment, passives will elicit poorer performance than ORs and

OWhQs. Conversely, if WH-movement is the specific impairment, ORs and OWhQs will elicit poorer performance than passives.

2.3.1.3 Modality

Resource Reduction Theory predicts that participants will have more difficulty overall in production tasks (NAT) than in comprehension tasks (SOAP).

Specific Impairment Theory predicts that if participants are demonstrating consistent error patterns on a structure type, they should demonstrate that error pattern on both the SOAP and the NAT. For example, if participants cannot comprehend passives, they should not be able to produce them either. A deficit in the linguistic system required to process passives (whether comprehending or producing them) will prevent success in either modality. In the current study, this linguistic deficit in processing noncanonical structures will be reflected in overall poorer performance on the NAT than on the SOAP.

While Specific Impairment Theory predicts the error patterns will be uniform, as described above, a caveat may exist where slightly different pattern of performance may also be permitted and support this theory. While the same structure type should be impaired in both modalities, the actual difference in performance (accuracy) may be parallel rather than identical. For example, OWhQs should be more impaired than SWhQs in both modalities. However, as long as the difference in performance between these two structure types is of the same magnitude, where comprehension yields better performance than production overall, the pattern may still support Specific Impairment Theory.

Based on previously observed dissociating patterns of better performance in comprehension than production (Schroder et al., 2014) the current study predicts better

performance on the comprehension tasks of the SOAP than on the production tasks of the NAT. In addition to patterns in IWA, the language acquisition patterns of typically developing children discussed in section 2.1.3 are also valuable to predictions regarding modality (Paul, Norbury, and Gosse, 2018).

2.3.1.4 Length

Resource Reduction Theory predicts that a more severe participant will have fewer resources than a mild participant and will have more difficulty on all structures. Once length is accounted for, Resource Reduction Theory also predicts that ORs will elicit the poorest performance, followed by SRs, followed by passives, followed by actives which would elicit the best performance. This prediction is based on the idea that sentences with two clauses will elicit poorer performance than those with one clause because they require more resources to process. So longer structures (ORs and SRs) will elicit poorer performance for severe participants when compared to one-clause structures of the same canonicity.

Specific Impairment Theory predicts that the noncanonical passives and ORs will elicit poorer performance than their canonical active and SR "controls" regardless of length.

2.3.2 Person-Related Factors/Implications: Severity

Overall aphasia severity may influence production and comprehension of sentences by IWA. The two competing theories discussed have different predictions regarding the influence of severity on comprehension and production. Resource Reduction Theory predicts that the most severely impaired participants will have the poorest performance in all structure types, and that

they will struggle more with noncanonical and longer structures. Thus, a severe participant may perform poorly on actives, SRs and SWhQs as well as their noncanonical experimental counterparts (passives, ORs, OWhQs). Here, the errors made are not rooted in the syntax of the structures or the underlying linguistic systems used to process them, but in the limited resources the participants have available to use that system. Specific Impairment Theory predicts that participants will perform more poorly experimental structure types when compared to control pairings. According to this theory, patterns of errors will not rely on participant severity, but on the classification of test items that were incorrect.

Aphasia severity is an important factor when considering performance of IWA, but there are other person-specific factors to consider for this population. Age, aphasia type, months post-onset, and other demographic information are examples of person-specific factors that relevant to the linguistic abilities of IWA. However, because that information is not available for all participants, additional factors will not be considered for the purpose of this thesis.

2.4 Tests of Sentence Comprehension and Production Performance

2.4.1 The SOAP

A variety of behavioral assessments that exist in the fields of neurolinguistics and aphasiology and are used to address and evaluate language impairment that results from brain damage. Many of these tools, such as the Porch Index of communicative Ability (PICA; Porch, 1967), the Minnesota Test for Differential Diagnosis of Aphasia (MTDDA; Schuell, 1965), the

Western Aphasia Battery (WAB; Kertesz, 1982), and the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1972) are used to measure a variety of aspects within language comprehension and production as cited in Love & Oster (2002, p. 503-4). While these measures have enabled researchers and clinicians to open the door to classification and providing treatment services, they generally focus on a broader view of language impairment. As Love & Oster point out, the wide range of impairments targeted in these diagnostic measures make it impossible for one to "expect them to necessarily demonstrate sensitivity to individual subsystems of language processing" and as a result, other "more focused subtests have begun to be developed by researchers" who seek to target "specific subsystems of interest" (Love & Oster, 2002, p. 504). In order to investigate the efficacy of the SOAP as a valid measure of specific comprehension abilities, 76 individuals participated in a study conducted at The Laboratory for Research on Aphasia and Stroke (LRAS) at The University of California, San Diego (UCSD) or at The Aphasia Research Center at the Boston Veteran's Administration Medical Center. Participants were placed into 1 of 3 testing groups: (1) those with neural involvement (n=26), (2) younger unimpaired controls (n=46), and (3) older (age-matched) unimpaired controls (n=4) and performance on the SOAP was compared. Data collected by Love and Oster (2002) demonstrate the SOAP's sensitivity and reliability in distinguishing comprehension abilities of broad aphasia groups.

2.4.2 The NAT

The Northwestern Anagram Test (NAT; Thompson, Weintraub & Mesulam, 2012) was created to assess syntactic production abilities and to assess accuracy of word order in the production of sentences by individuals who present with "speech production, word comprehension and/or word-finding difficulties, or reduced working memory capacity" (Weintraub et al., 2009, p.

410). Sentence production deficits are known to be an issue for IWA, but understanding the exact underpinnings of where and why the deficits occur is a challenge for researchers and clinicians who seek to provide services. Understanding these underpinnings is important for clinical purposes, and many existing assessment tools do not address certain sentence properties. Difficulty with noncanonical sentence, particularly those with reversible syntax, has been well documented in individuals with agrammatic aphasia. Some assessment tools have been created to examine sentence production deficits in aphasia, and they control for these syntax variables. Unfortunately, it is not uncommon for IWA to present with severe motor speech, word retrieval, and/or working memory deficits which can impact their performance on tasks related to sentence production (Weintraub et al., 2009). When one intends to examine syntactic impairments, it is important to be able to look at participant performance in a way that these potential coexisting deficits will interfere. The Northwestern Assessment of Verbs and Sentences (NAVS; Cho-Reyes & Thompson, 2012) is a recently developed test that contains subtests for both canonical and noncanonical sentences. The NAVS includes a sentence comprehension test (SCT) using picture-matching, a sentence production priming test (SPPT), which primes the target structures with a sample sentence that describes an image before asking the participant to create their own in response to a semantically reversed sentence. (NAVS; Cho-Reyes & Thompson, as cited in Weintraub et al., 2009). While the NAVS is beneficial in that in controls for sentence structure type, it may be too difficult for some patient populations to participate in. Thus, the NAT uses an anagram task to test production without having participants speak by using word cards that match the action pictured. IWA arrange these word cards in order to create their response (NAT; Thompson, Weintraub & Mesulam, 2012).

2.5 Research Questions

The current study aims to address questions regarding comprehension and production of sentences by IWA using data collected from the SOAP and the NAT. Performance refers to participant reaction time (RT) and accuracy.

1. Do the noncanonical experimental structures elicit poorer performance (lower accuracy) than their noncanonical control pairings?

2. Does modality (comprehension or production) predict performance in the same way on both the SOAP and the NAT?

3. Does participant severity predict performance in the same way regardless of canonicity, length and modality?

4. Do longer structures (2 clauses) elicit poorer performance (lower accuracy) than shorter structures (1 clause)?

3.0 Methods

3.1 Participants

99 IWA participated in a larger study conducted through the VA Pittsburgh Healthcare System (VAPHS), the VA Northern California Health Care System (VANCHS), and the VA Puget Sound Health Care System (VAPSHCS). Two non-VA sites also participate and contributed to the study: The University of Washington, Seattle and Temple University, Philadelphia. The participants completed approximately 20 hours of cognitive and language testing that assessed phonological, semantic, syntactic processing and short-term memory and cognitive control. This testing included the language comprehension (SOAP) and production (NAT) tasks reported in this paper. Fourteen screening and descriptive tasks were also part of the testing that occurred in the larger VA study described, and were used to determine whether participants met the following criteria for inclusion: adequate peripheral vision and hearing, no present visuospatial neglect or dementia, no history of TBI, degenerative nervous system illnesses, schizophrenia, manic-depressive disorder, schizoaffective disorder, or current substance abuse. Presence of aphasia was determined based on medical history, and below-cutoff performance on at least two subtests of the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004) and/or the listening version of the Computerized Revised Token Test (CRTT-L;, McNeil et al., 2015). Screening and descriptive tasks were also used to obtain the following information: presence and severity of aphasia. Participants were all native speakers of English. There were 31 female, 67 male participants, and one participant that did not report their gender (NR). Participants included 64 Veterans and 35 non-veterans. Participants ranged in age

from 26 to 84 years (mean 63.7 years) and had from 10 to 24 years of education (mean 15.4 years). Participants months post onset (MPO) ranged from four to 216 months (mean 84.2 months). Participant severity was measured as the average T-score on the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004). CAT T-score is valuable as it is a normalized score that represents performance on CAT, which is a standardized and comprehensive tool used to assess all modalities and areas of language in order to diagnose aphasia. Participant severity (T-score) ranged from 44 to 66 (mean 56.4) Participant race/ethnicity is broken down as follows: American Indian/Alaskan Native (2), Asian (3), Black/African-American (12), Hispanic/Latino (1), Native Hawaiian/other Pacific Islander (0), White/Caucasian (76), Other (1), Mixed (4).

Table 3. Participant Demographics

	Age (Years)	**Education (Years)**	**MPO**	**Severity (CAT T-score)**
Mean (SD)	63.7 (10.8)	15.4 (2.7)	84.2 (56.2)	56.4 (5.6)
Range	26-84	10-24	4-216	44-66

MPO= months post onset, SD=standard deviation, F=female, M=male, Y= Veterans, N= Non-veterans

3.2 Procedures

3.2.1 The SOAP

The SOAP was one of 14 descriptive tasks conducted in the first two-three sessions of testing. As outlined in Love & Oster (2002), the SOAP syntactic battery includes the following materials: "5 practice sentences (involving active and subject-relative constructions only) and 40 experimental sentences targeting reversible actions with active passive, subject-relative, and object-relative constructions (10 each)." Appendix A provides a list of the practice and experimental sentences in the SOAP. The 40 experimental sentences are pseudo-randomized so that the same syntax structure cannot be presented more than twice in a row, and all sentences are controlled for length of sentence (within two words) across the syntactic categories: subjective-relatives (SR), object-relatives (OR), actives (A), and passives (P).

A computerized adaption of the SOAP task was presented to participants on a laptop (Yoo and Dickey, 2014). Sentences are matched with three simple line drawings that involve two characters and one action and fall into three conditions. The first condition ("match condition") includes pictures that correctly depict the sentence. The second condition ("mismatch condition") include pictures that involve the same characters and actions but with reversed thematic roles (such that the patient is performing the action). The third condition is a foil picture and depicts an unrelated action and characters. See Figure 1 for an example of experimental pictures presented for the sentence, *The woman in the bathing suit dries the child*.

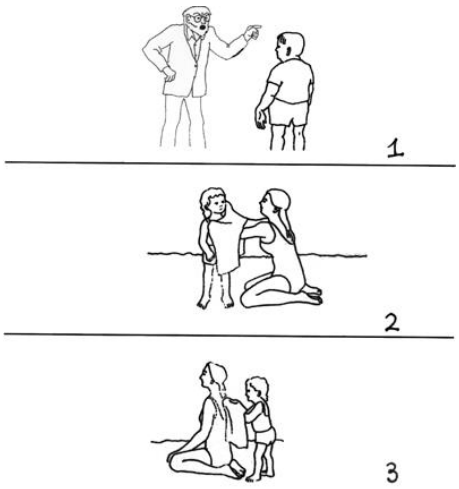

Figure 1. Sample of experimental pictures from the SOAP

Procedures

Participants listened to the following instructions, presented auditorily on a laptop:

"*You will see three pictures on the screen while listening to a sentence. Match the sentence you hear to one of the three pictures on the screen. Press one of the buttons (1, 2, or 3) to indicate your answer. When you are ready to move to the next item, please press the "2" button. Do you have any questions? Let's try a few for practice"* (Yoo and Dickey, 2014).

The five practice items were then administered. The sentences presented for practice items can be found in Appendix A. Once the participants completed practice items, they listened to the following instructions before beginning the experimental items:

"*Now you will listen to more sentences with more pictures. Press one of the buttons (1, 2, or 3) to match the sentence to a picture."*

For each trial, three pictures (one for each condition described) were presented vertically on the computer screen (see Figure 1) while the participants listened to the sentence. They pressed a button on a response box to indicate their choice. The inter-trial interval was four seconds and

there was no response deadline. Participants pressed a button on the response box when they were ready for the next item (Yoo and Dickey, 2014).

3.2.2 The NAT

The NAT (Thompson, Weintraub & Mesulam, 2012) includes a Long Version and a Short Version, which both measure sentence production of canonical and noncanonical sentences. The Long Version includes the same sentence types as the NAVS (NAVS; Thompson, 2011) and consists of two practice items and 30 experimental test items: five actives (A), five passives (P), five subject extracted Wh-questions (SWhQ), five object extracted Wh-questions (OWhQ), five subject-cleft structures (SR), and five object-cleft structures (OR). Appendix B provides a list of the practice and experimental sentences in the NAT. (NAT; Thompson, Weintraub & Mesulam, 2012)

The images in the Picture Stimulus Book are a subset of those used in the NAVS (Thompson, 2011) to allow for comparison of performance on the NAVS and NAT, which can be used to look at performance when production is done orally versus manually in anagram tasks. Simple black and white line drawings are presented to the participant. The corresponding word cards are placed below the picture card in a random order so that the participant may move them to create their response. See Figure 2 for an example of the picture stimulus and word cards presented for the target sentence, *"The dog is biting the cat"* (NAT; Thompson, Weintraub & Mesulam, 2012).

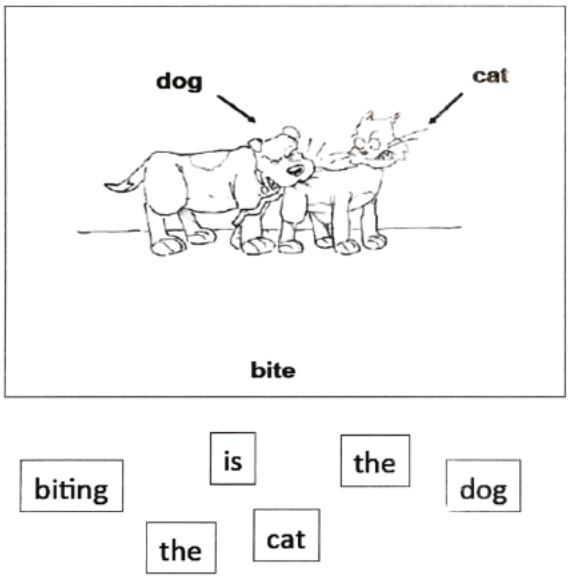

Figure 2. Sample of an experimental picture and word card from the NAT

Instructions are read to the participant and pictures are presented one at a time, with corresponding Word Cards, using the Picture Stimulus Book. The examiner administers the two practice items first. For practice items, participants the examiner should place the Word Cards in the correct order under the picture if the participant does not respond within 30 seconds or responds incorrectly. On all items (practice and experimental), the following instructions/cues are read aloud to the participants:

"*This picture shows a _____"* (point to and name entity on the left side of the picture) "*and a_____"* (point to and name the entity on the right side of the picture) (use participant's left and right for orientation). "*The action is _____"* (point to action word and name the action). Present corresponding word cards in the work space beneath the picture in a random array. *"Use these words"* (point to scattered words) "*to make a sentence to go with the picture. Start with these words"* (slide the word(s) underlined on the response form for each target item, in correct order, to the top left of the work space). *"Be sure to use all of the words to make your sentence."*

On experimental items, the examiner should not provide corrections to incorrect or incomplete responses. Instead, responses are recorded, and the next item is presented after the 30-second response deadline. Credit is only given for responses in which all words are in the correct order of the target sentence and no partial credit is given for pairs of words that occur in the correct order.

While the NAT tests language production, it requires participants to rely on reading ability in order to manipulate word cards to construct a sentence. The NAT administration procedure described above helped ensure that participants could read and identify the content words (nouns and verb) in the task. This means that any problems an individual may have had in reading would not interfere with their completing the NAT. Also, in the current study, participants' reading ability was assessed through the CAT, which includes subtests that test participant ability to use complex words, function words and non-words (CAT; Swinburn, Porter, & Howard, 2004). Although not carried out for this study, future analyses could examine whether CAT reading comprehension T-scores affected participants' performance on the NAT. (See section 5.5.)

3.3 Data Analysis

Data from the SOAP and the NAT were used to address the research questions outlined in section 2.5. Data from both tasks were combined to investigate questions 1-3. Question 1 examined the effect of word order type (do noncanonical items elicit poorer performance than canonical items). Question 2 examined the effect of modality (does production elicit poorer performance than comprehension?). Question 3 examined the effect of severity (do severe participants perform more poorly than mild participants?). Data from the NAT were used to

investigate question 4, which examined the effect of length (do sentences with two clauses elicit poorer performance than sentences with one clause?). In order to isolate length, question 4 was addressed using data from the NAT alone, rather than combined data from both tasks. The NAT contains Wh-questions (SWhQs and OWhQs), which were compared to relatives (SRs and ORs) so that noncanonical items shared the same movement type (WH-movement). Wh-questions on the NAT contain one clause and relatives on the NAT contain two clauses.

The SOAP data are item-level data, where the response for each trial (item) is available. The NAT data is different from the SOAP data because it is not item-level data. The response for each participant's response to each item on the NAT is not available. Instead, the number of items correct in each of the six sentence categories (A, P, SR, OR, SWhQ, OWhQ) is reported for each participant. This leaves out information (for example, which Active items specifically did the participant get correct or incorrect). Raw data was transformed in order to combine data sets from the NAT and the SOAP, and was reported as mean accuracy within both modalities. Planned analysis included investigation of sentence type in addition to the discussed variables; however, sentence type was removed from data analysis due to a nonconvergence issue: the data for sentence type conflates movement with length, and movement type was not under investigation in this study. Subsequently, two separate, repeated measures ANOVAs were used to analyze the variance in accuracy data from the NAT and the SOAP and address canonicity, modality, word order category, severity and length . A three-way ANOVA used combined data from both the SOAP and the NAT to investigate the following factors: word order category (noncanonical and canonical), severity (mild, moderate, severe), and modality (production and comprehension). This three-way ANOVA will be referred to as ANOVA 1. A two-way ANOVA used data from the NAT to

investigate two factors: length (one clause vs. two clauses) and word order category (noncanonical and canonical). This two-way ANOVA will be referred to as ANOVA 2.

Significant main effects were further investigated with post-hoc analysis to determine which groups of data were significantly different. Pairwise comparisons for interaction effects were made using t-tests with pooled standard deviations using the False Discovery Rate (FDR) method. Power was reported as eta squared (η^2) and was interpreted using Cohen's 1988 guidelines, where a value of 0.01= small, 0.06= medium, and 0.14= large.

4.0 Results

4.1 Results from ANOVA 1

Data from three participants was missing in the SOAP due to non-compliance. Data from three participants were missing in the NAT due to experimenter error and technical difficulties. The combined data from the SOAP and the NAT analyzed in ANOVA 1 was collected from 93 participants. Results from ANOVA 1 showed a significant effect of canonicity, ($F(1, 90)=218.78$, $p<0.00$, $\eta 2=0.21$), where canonical sentences elicited higher accuracy than noncanonical sentences (canonical: M=0.76, SD= 0.32; noncanonical: M=0.54, SD=0.36). Results from ANOVA 1 showed a significant main effect of modality ($F(1, 90)=80.40$, $p<0.00$, $\eta 2=0.20$), where comprehension elicited higher accuracy than production (comprehension: M=0.76, SD=24 production: M= 0.54, SD=0.43).

ANOVA 1 yielded an interaction effect, where canonicity interacted with modality ($F(1, 90)=13.28$, $p<0.00$, $\eta^2=0.01$). Canonical items showed a smaller modality difference than noncanonical items. Within the SOAP, the difference between noncanonical and canonical items is smaller than the difference between noncanonical and canonical items on the NAT. Pairwise post-hoc comparison of canonicity and modality using t-tests with pooled standard deviation and the FDR method showed that participants performed more poorly on noncanonical sentences on both tasks, but average noncanonical sentence performance on the NAT (*M=0.41, SD=0.4)* was 0.26 lower than average noncanonical sentence performance on the SOAP (*M=0.67, SD=0.25).* Pairwise comparisons investigating canonicity and modality are outlined in Table 4. The interaction between canonicity and modality is displayed in Figure 3.

Table 4. Post-Hoc Analysis of Canonicity and Modality

	NAT Canonical	SOAP Canonical	NAT Noncanonical
SOAP Canonical	*p<0.00*		
NAT Noncanonical	*p<0.00*	*p<0.00*	
SOAP Noncanonical	p=0.80	*p<0.00*	*p<0.00*

p-values yielded by pairwise t-tests. Significant results are italicized

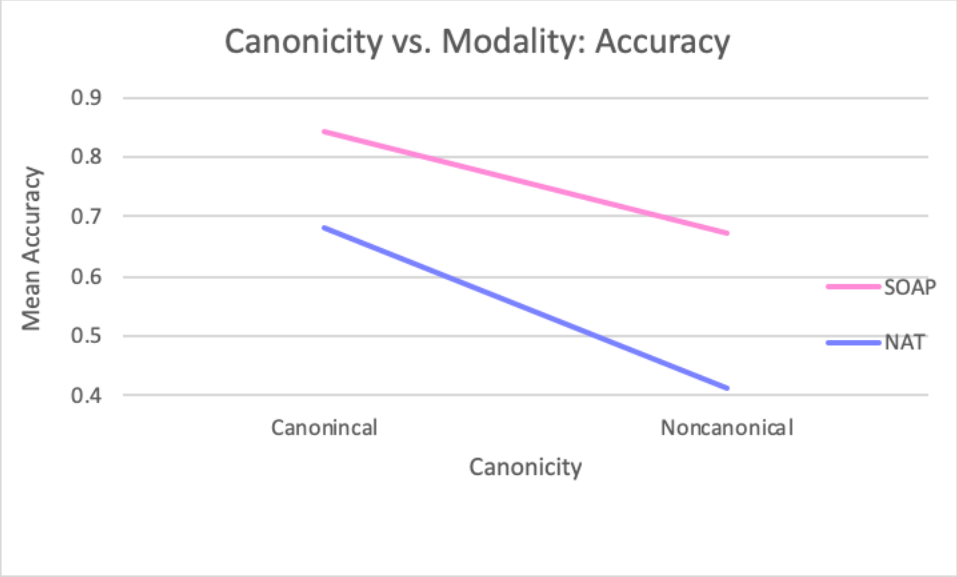

Figure 3. Canonicity vs. Modality: Accuracy

The ANOVA yielded a main effect of severity, ($F(2, 90)=36.72$, $p<0.00$, $\eta^2= 0.30$), where severe participants (*M=0.46, SD=0.36)* performed more poorly than moderate participants (*M=0.68*, *SD=0.34*), who performed more poorly than mild participants (*M=0.80*, *SD=0.29)*. The ANOVA also yielded an interaction effect where severity interacted with modality ($F(2, 90)=11.47$, $p<0.00$, $\eta^2=0.07$). Pairwise post-hoc comparison of severity and modality using t-tests with pooled standard deviation and applying the FDR method showed that severe participants performed more poorly than moderate participants, who performed more poorly than mild

participants on both tests. Pairwise comparisons investigating severity and modality are outlined in Table 5. The interaction between severity and modality is displayed in Figure 4.

Table 5. Post-Hoc Analysis of Severity and Modality

	Severe NAT	Moderate NAT	Mild NAT	Severe SOAP	Moderate SOAP
Moderate NAT	*p<0.00*				
Mild NAT	*p<0.00*	*p<0.00*			
Severe SOAP	*p<0.00*	p=0.18	p=0.0		
Moderate SOAP	*p<0.00*	*p<0.00*	p=0.73	*p<0.00*	
Mild SOAP	*p<0.00*	*p<0.00*	p=0.14	*p<0.00*	*p=0.03*

p-values yielded by pairwise t-tests. Significant results are italicized

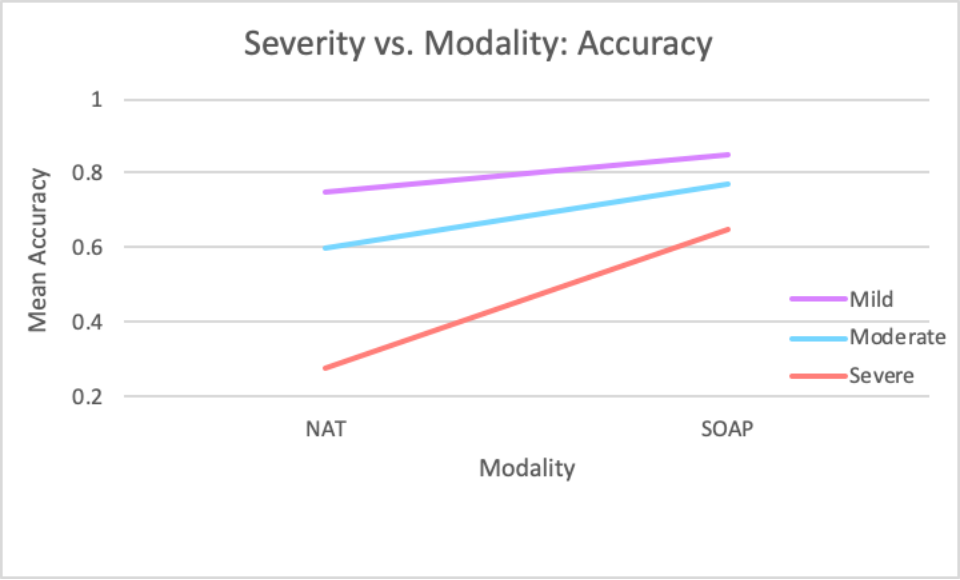

Figure 4. Severity vs. Modality: Accuracy

4.2 Results from ANOVA 2

To analyze the effects of length on participant performance, a two-way ANOVA was performed using data from the NAT (ANOVA 2). Length was a two-level factor (one clause vs two clauses). Canonicity was a two-level factor (canonical vs noncanonical). Only data from the NAT was used for this analysis because it allowed for the isolated comparison of length using Wh-questions and relatives in the same modality (production) while controlling for movement type (WH-movement). Data from three participants were missing in the NAT due to experimenter error and technical difficulties. NAT data analyzed by ANOVA 2 was collected from 96 participants. Analysis showed a significant main effect for length ($F(1, 97)=46.00$, $p<0.00$, $\eta^2=0.04$) and a significant main effect for canonicity ($F(1, 97)=84.65$, $p<0.00$, $\eta^2=0.12$), the latter being consistent with results from ANOVA 1. Analysis showed a significant interaction effect between canonicity and length ($F(1, 97)=9.18$, $p<0.00$, $\eta^2=0.009$). Pairwise post-hoc comparison of length and canonicity using t-tests with pooled standard deviation and applying the FDR method showed significant differences in all comparisons except for longer canonicals and shorter canonicals ($p=0.11$). Pairwise comparisons investigating length and canonicity are outlined in Table 6. The interaction between length and canonicity is displayed in Figure 5.

Table 6. Length vs. Canonicity

	Short Canonical	Short Noncanonical
Short Noncanonical	*p<0.00*	
Long Canonical	p=0.11	*p=0.03*
Long Noncanonical	*p<0.00*	*p<0.00*

p-values yielded by pairwise t-tests. Significant results are italicized

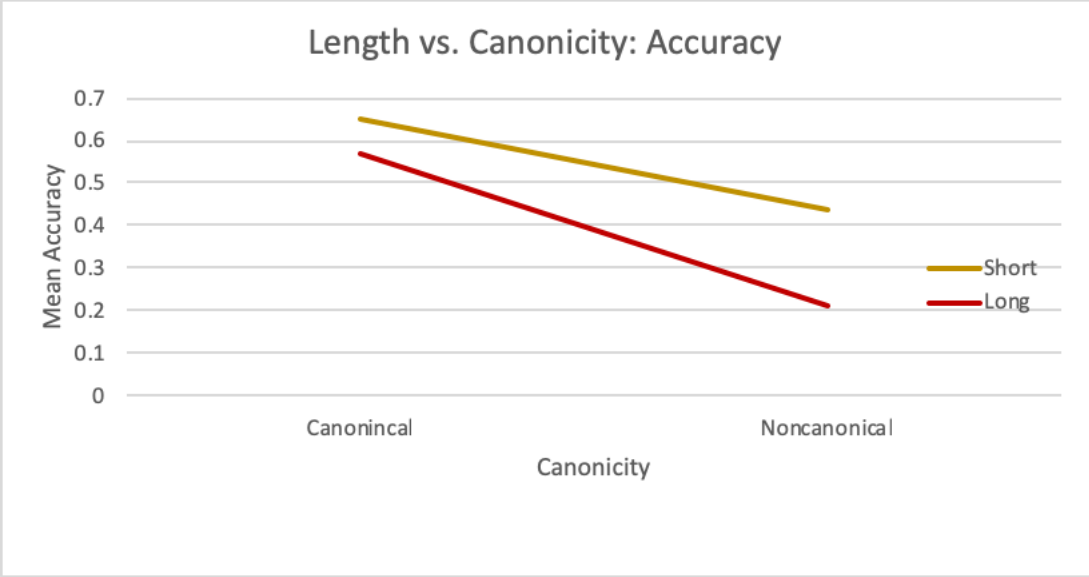

Figure 5. Length vs. Canonicity: Accuracy

5.0 Discussion

5.1 Canonicity

Canonicity had a significant effect on participant performance, where noncanonical items elicited poorer performance than canonical items on both tasks. The main effect of canonicity had a large effect size (η^2=0.21), suggesting that canonicity had a strong impact on participant performance. This finding is consistent with Resource Reduction Theory and Specific Impairment Theory. The main effect of canonicity is consistent with Resource Reduction Theory because this effect suggests that noncanonical structures of increased complexity are more difficult because they require more resources than less complex, canonical structures. The main effect of canonicity is also consistent with Specific Impairment Theory because the effect suggests that one syntax type (reversible sentence structure with movement) is problematic for participants with a specific impairment .

5.2 Modality

Modality had a significant main effect on participant performance, where participants performed more poorly overall on items of the NAT than on items of the SOAP. The main effect of modality had a large effect size (η^2=0.20), suggesting that modality had strong effect on participant performance. These findings are consistent with the argument that production requires more linguistic resources than comprehension (Schroder et al., 2014) and with Resource Reduction

Theory. The main effect of modality is inconsistent with Specific Impairment Theory. Impairment Theory claims that poor performance can be attributed to a deficit in a central process that occurs in specific structures across both modalities. A broken syntactic process, rather than increased difficulty in one modality or another, would be reflected by equally impaired performance in both modalities.

Further evidence in favor of Resource Reduction Theory comes from the interaction between canonicity and modality, where noncanonical items on the NAT elicited poorer performance than noncanonical items on the SOAP. This suggests that, while noncanonical syntax resulted in poorer performance overall, requiring participants to *produce* these structures was more difficult than requiring them to *comprehend* them. This is consistent with Resource Reduction Theory because it suggests that the size of canonicity effects depend on modality. This is inconsistent with Specific Impairment Theory, which predicts equally poor performance in both production and comprehension. It is important to be cautious of interpretations regarding the interaction of canonicity and modality because of the low power (η^2=0.01) of this interaction. Cohen (1988) defined an η^2 of 0.01 as small, 0.06 as medium and 0.14 as large. Given these guidelines, the power of this interaction effect can be considered small. This may be because the effect is small for most participants in the sample or because the pattern for holds true for some but not all participants in the sample. Determining which of these two possibilities is correct would require analysis of individual PWA data. Regardless of which of these possibilities is correct, this significant but small-power interaction effect between canonicity and modality may not be generalizable to all IWA.

5.3 Severity

Severity had a significant main effect on participant performance, which followed a predicted pattern: severe participants performed more poorly than moderate participants, who performed more poorly than mild participants overall on both tasks (the SOAP and the NAT). The main effect of severity had a large effect size (η^2=0.30), offering support for the strong impact severity had on participant performance. This finding suggests that poor performance is determined by participant characteristics, in particular severity, in addition to the possible inability to process specific syntactic structures, in particular, noncanonical structures with WH- or NP-movement. The main effect of severity is consistent with Resource Reduction Theory, as this finding demonstrates that participants who have more severe deficits performed more poorly than participants with mild deficits on all structures. The main effect of severity is inconsistent with Specific Impairment Theory, which does not predict a main effect of severity. Specific Impairment Theory suggests that specifically impaired syntactic structures will elicit poor performance by all participants who have this deficit, regardless of their individual severity levels.

The interaction between severity and modality demonstrates that the difference between performance on the NAT and the SOAP was greatest for severe participants. The interaction effect between severity and modality had a medium effect size (η^2=0.07). This interaction effect is consistent with Resource Reduction Theory, as it demonstrates the way increasing the amount of cognitive and linguistic demands on participants with severe language deficits results in significantly poorer performance on harder tasks. Mild and moderate participants also performed more poorly on the NAT; however, the effect of increased difficulty is most easily observed when looking at the performance of severe participants on the NAT. Although the interaction of severity

and modality is not addressed by Specific Impairment Theory, this pattern is unexpected because it suggests that factors other than canonicity play an important role in participant performance. One may consider the effects of severity as missed generalizations by the Specific Impairment Theory, where sentence structure is considered the primary predictor of performance.

5.4 Length

Data from the NAT was used in ANOVA 2, as it allowed isolated comparison of length without the influence of movement type (WH- vs. NP-movement) and modality (comprehension vs. production). Wh-questions (SWhQs, OWhQs), found only on the NAT, and relatives (SRs, ORs) were used to examine length in production. Wh-questions on the NAT contain one clause. Relatives on the NAT contain 2 clauses. The noncanonical structures (OWhQs and ORs) share the same movement type (WH-movement).

Length had a significant main effect on participant performance on the NAT, where participants performed more poorly on longer structures (two clauses) than shorter structures (one clause). The main effect of length is consistent with Resource Reduction Theory, as this finding suggests that increased structure length results in increased difficulty, which would require greater linguistic resources for participants. This finding is inconsistent with Specific Impairment Theory as it suggests that length of a structure, rather than a specific syntactic impairment, predicts performance.

Pairwise post-hoc analysis revealed an interaction effect between length and canonicity, where performance on longer, noncanonical items (ORs) were significantly poorer than shorter, noncanonical items (OWhQs). Longer structures elicited poorer performance than shorter

structures, however the degree to which longer structures were impaired was greater for long noncanonical structures than short noncanonical structures. This interaction is consistent with Specific Impairment Theory, as it suggests that participants are able to produce longer, canonical items to a similar degree as shorter canonical items. The increased difficulty associated with length occurs in noncanonical structures, where a specific syntactic movement is required. This interaction is consistent with Resource Reduction Theory. The degree to which longer structures are impaired is larger in the noncanonical condition than in the canonical condition. This suggests that, while length adds complexity and longer structures may be more difficult to produce, the difficulty is greater when a specific syntactic process is involved. Noncanonical structures containing syntactic movement may require more linguistic resources, thus making them more difficult for participants to produce, especially when they contain an additional clause. It is important to be cautious of interpretations regarding the interaction of length and canonicity because of the low power (η^2=0.009) of this interaction, which has an effect size that approaches small (Cohen, 1988). This may be attributed to high variability in the sample, which reflects the fact that this pattern of length and canonicity in performance may hold true for some but not all participants included. Thus, this significant interaction effect between length and canonicity may not be generalizable to all IWA.

5.5 Limitations and Future Research

Statistical methods discussed in section 3.3 were selected in order to investigate the current study's research questions as thoroughly and effectively as possible within constraints of retrospective investigation and time-constraints. Limitations of the current study include the

limited number of observations for items available in both the SOAP and the NAT. The SOAP consists of 40 experimental items, half of which are declaratives (A, P) and half of which are relatives (SR, OR). The NAT consists of 30 experimental items and also includes Wh-questions, which further reduces the number of items in each sentence type (five items each). As such, the interpretations made were limited by the relatively small number of trials in investigated structures, especially for Wh-questions, which were only observed in production.

In addition, structures observed in production rely on the ability to read the words presented on the NAT word cards. Relying on reading could have increased difficulty for some participants. One way to examine whether reading impairments affected participants' ability to complete the NAT would be to perform additional analyses using the reading-comprehension subtest from the CAT. Although not carried out for this study, future analyses could examine whether CAT reading comprehension T-scores affected participants' performance on the NAT, with people with lower T-scores having worse performance on the NAT. Furthermore, although reading-based measures were collected for participants in the current study, in the form of the CAT Reading Comprehension subtests, these subtests do not test reading in the same way the NAT does. For example, on the CAT, oral reading is measured; however, on the NAT participants are not required to read aloud. Future research should include additional assessments to measure participant reading ability as it is utilized by participants completing the NAT.

A notable limitation of the current study is the retrospective design used to comparing performance on the two tasks (the SOAP and NAT) that were not directly matched for sentence stimuli and structures. As noted above, the SOAP does not include Wh-questions. Similarly, the specific active, passive, OR and SR sentences differed across the tasks, which can be seen in the

stimuli found in Appendix A and Appendix B. Future prospective studies may account for limitations such as unmatched stimuli by using sentences that are matched across modalities.

While the results of the analyses in the current study were significant, the variability in performance in the sample was also very high. This is illustrated by considering mean performance in relation to standard deviations reported. The following results had small effect sizes: the interaction effect between canonicity and modality (η^2=0.01), and the main effect of length (η^2=0.04), (Cohen, 1988). The interaction between length and canonicity approached a small effect size (η^2=0.009), (Cohen, 1988). The possible causes for variability and small effect sizes, as mentioned in Section 5.2, may be addressed by future research where individual participant performance is analyzed. Further, while the potential variability in the current study can limit the extent to which generalizable conclusions can be made, the variability is logical when one considers the sample. As outlined in section 1.0, aphasia is a disorder that can impact the four domains of language in a variety of ways. IWA do not fit perfectly into one linguistic profile, and results of the current study reflect this.

Participant demographics outlined in section 3.1 may also be considered limitations of the current study. Effects observed in the current study may not be representative of IWA because of the high proportion of males (67) to females (31), high proportion of Veterans (64) to non-veterans (35) , and unequal representation of race/ethnicity in the sample. Although the range of months post onset (four to 216 months) was wide, potentially making this sample more representative of IWA as a group, the impact of this variability has not yet been examined.

Limitations of the current study also include the limited number of analyses that were performed on the data. Future research addressing individual analysis of participant performance would add valuable insight regarding questions of specific impairments. For example, examining

each participant's performance would allow researchers to identify patterns of specifically impaired noncanonical structures as well as inform questions regarding movement type (NP- vs. WH-movement). Individual analysis in future research may also help investigate the variability associated with the range of months post onset.

Finally, future research investigating syntactic deficits in IWA should include a larger number of trials in both comprehension and production, as well as in various sentence types like those discussed in the current study. It is possible that in addition to assessing linguistic abilities of IWA, performance of healthy controls without aphasia on the same measures may also provide useful information regarding Resource Reduction Theory and Specific Impairment Theory.

6.0 Conclusions

Overall, patterns of performance demonstrated the impact of linguistic complexity on participant ability to accurately comprehend and produce language in the presence of person-related linguistic deficits. The variables that contributed to linguistic complexity in the current study include the following task-related variables: canonicity, modality, and length. Severity contributed to person-related linguistic deficits and was also a variable that impacted participant performance.

In regard to how the results of the current study compare to those found by Caplan et al., (2007) several points are worth noting. Caplan et al. investigated group performance and reported that overall, the results were consistent with the conclusion that participant performance is determined by the processing demands required to complete linguistic tasks (Caplan et al., 2007: 145). This is more consistent with the predictions of Resource Reduction Theory. Caplan et al. also investigated individual performance and reported that if results were interpreted according to the most strict rules of Specific Impairment Theory, where all canonical structures are intact and all experimental structures are impaired, only one participant's performance is reflective of a specific deficit (Caplan et al., 2007: 125-6). If results were interpreted according to more lenient Specific Impairment criteria, Caplan et al. report one other participant who demonstrated a deficit in passives and object-extracted structures. These individual results may be attributed to the presence of a deficit impacting structures with traces, which is consistent with Specific Impairment Theory.

The current study did not include individual analysis, and as such the results cannot be directly compared with individual findings reported in Caplan et al., (2007). Group analysis in the

current study yielded results generally consistent with Resource Reduction Theory, where factors that added complexity (increased severity and length, production and noncanonical syntax) resulted in poorer participant accuracy. These findings, although not as detailed as those presented by Caplan et al., are consistent with their findings, and the idea that increased processing demand yields poorer linguistic performance.

6.1 Canonicity

In regard to canonicity, noncanonical syntax elicited poorer performance than canonical syntax. This suggests that reversible sentences are more difficult to produce and comprehend for IWA. As mentioned in the discussion, this observation is consistent with both Resource Reduction Theory and Specific Impairment Theory. It is possible that noncanonical syntax requires an individual to draw upon more resources as a result of the linguistic processes (movement) present in passives, ORs and OWhQs. It is also possible that noncanonical syntax is more difficult for IWA regardless of available linguistic resources, as a result of a nonfunctioning component of their language system required to process noncanonical syntax. Distinguishing which of the two theories is best supported or refuted by the observed effects of canonicity may be possible with further, individual analyses that look closely at each participant's performance on the different sentence types assessed.

6.2 Modality

In regard to modality, production elicited poorer performance than comprehension. This suggests that it is more difficult for IWA to produce language than to comprehend it. This observation is consistent with Resource Reduction Theory and suggests that IWA struggle with expressive language more than receptive language. As such, noncanonical syntax is not the only factor that increases complexity of a language task and poor performance cannot be attributed solely to a specific syntactic deficit as Specific Impairment Theory would suggest. Further, the interaction between canonicity and modality illustrates that noncanonical syntax is not the only factor involved in poor performance: participants performed more poorly on noncanonical items in production than in comprehension. This is important, as it may inform the abilities of IWA who have a relatively preserved ability to comprehend reversible syntax despite difficulty when faced with the more linguistically-taxing modality of production.

6.3 Severity

In regard to severity, participants performed as predicted overall, where decreased performance occurred in the presence of increased severity. In addition, the interaction between severity and modality observed suggests that severe participants had more difficulty in production tasks than moderate or mild participants. As discussed, these patterns are consistent with Resource Reduction Theory. The observed patterns of severity pose a question as to how the effect of severity should be interpreted according to Specific Impairment Theory. Specific Impairment Theory focuses on a universally "broken" language structure or process in IWA as the cause for

poor performance on language tasks. As such, severity effects observed in the current study do not offer support for Specific Impairment Theory. Additionally, the effects of severity observed demonstrates that IWA do not share one pattern of language deficits; linguistic abilities differ across individuals and language tasks. It is important to consider the impact of severity on linguistic abilities when assessing and treating IWA as an understanding of reduced available resources should inform clinical decision making.

6.4 Length

In regard to length, participants performed more poorly on longer structures than on shorter structures overall. This main effect is consistent with Resource Reduction Theory and inconsistent with Specific Impairment Theory because it suggests that increased sentence length contributes to difficulty, as opposed to canonicity alone. However, the interaction effect observed between length and canonicity offered support for Specific Impairment Theory and qualified the degree to which the main effect of length supports Resource Reduction Theory. While longer structures did elicit poorer performance overall, increased length had a stronger negative impact on items with noncanonical syntax. For certain participants, this may be interpreted as evidence for a specific impairment that is makes it more difficult to produce long structures if they have noncanonical syntax. Given that length was examined in only one modality (production on the NAT) and using only two sentence types (relatives vs. Wh-questions), further investigation is required to draw conclusions about the way length effects can be generalized to IWA in general. Investigating length in both modalities, with a larger number of items and sentence types may provide beneficial information to support ideas discussed in the current study.

6.5 Competing Theories

The current study produced results that can contribute to evaluations made regarding Resource Reduction Theory and Specific Impairment Theory. Through the investigation of how canonicity, modality, severity and length impacted participant performance, the current study has produced results that are overall more consistent with predictions of Resource Reduction Theory than Specific Impairment Theory. Overall, noncanonical syntax, production, increased severity, and increased length elicited poorer performance. These findings support Resource Reduction Theory because they suggest that the more difficult conditions resulted in poorer performance than less difficult conditions (i.e. noncanonical syntax vs. canonical syntax) which illustrates that increased complexity, rather than simply a specific impaired sentence type, is likely to increase difficulty for participants. It is also important to note that Specific Impairment Theory fails to address factors other than canonicity in a way that can be used to predict performance to the degree Resource Reduction Theory does. As such, missed generalizations occur when one considers and predicts performance through the lens of Specific Impairment Theory.

Appendix A SOAP Stimuli

A.1 Practice Sentences

1. The old man with the beard leads the little boy.
2. The woman in the bathing suit dries the child.
3. The teacher that scolds the student has black hair.
4. The nurse on the 5th floor examines the doctor.
5. The policeman that threatens the man has black hair.

A.2 Experimental Sentences

1. (P) The doctor with blonde hair is questioned by the soldier.
2. (SR) The man that pushes the boy is wearing a red shirt.
3. (A) The thin cowboy with the blue pants captures the Indian.
4. (A) The young boy with the brown hair grabs the man.
5. (OR) The doctor that the bedridden patient accuses has black hair.
6. (P) The boy with the brown hair is grabbed by the man.
7. (A) The little boy with the big book yells at the soldier.
8. (SR) The cowboy that captures the Indian has on blue pants.
9. (OR) The boy that the girl chases is wearing a green shirt.
10. (OR) The man that the young boy grabs has brown hair.

11. (SR) The man that records the young woman has brown hair.

12. (P) The boy in the green shirt is chased by the girl.

13. (A) The girl with blonde hair photographs the nurse.

14. (P) The man in the red shirt is pushed by the boy.

15. (A) The little boy in the blue shirt instructs the teacher.

16. (SR) The doctor that accuses the bedridden patient has black hair.

17. (OR) The teacher that the little boy instructs has on a blue shirt.

18. (SR) The girl that photographs the nurse with the camera is blonde.

19. (P) The cowboy with blue pants is captured by the Indian.

20. (OR) The soldier that the little boy yells at has black hair.

21. (A) The young doctor with blonde hair questions the soldier.

22. (SR) The man that grabs the little boy has brown hair.

23. (A) The man in the red shirt pushes the little boy.

24. (A) The man with brown hair records the young woman.

25. (SR) The soldier that yells at the small boy has black hair.

26. (A) The bedridden patient with the black hair accuses the doctor.

27. (P) The boy with black hair is yelled at by the soldier.

28. (SR) The soldier in uniform that questions the doctor has blonde hair.

29. (SR) The teacher that instructs the boy is wearing a blue shirt.

30. (OR) The girl that the nurse photographs with the camera is blonde.

31. (OR) The man that the young woman records has brown hair.

32. (P) The boy in the blue shirt is instructed by the teacher.

33. (OR) The cowboy that the Indian captures has on blue pants.

34. (OR) The man that the boy pushes is wearing a red shirt.

35. (P) The patient with black hair is accused by the doctor.

36. (OR) The soldier in uniform that the doctor questions has blonde hair.

37. (A) The girl in the green shirt chases the small boy.

38. (P) The man with brown hair is recorded by the woman.

39. (P) The girl is photographed by the nurse with blonde hair.

40. (SR) The boy chases the girl is wearing a green shirt.

Appendix B NAT Stimuli

B.1 Practice Sentences

1. (A) The girl is tickling the boy.
2. (SWhQ) Who is carrying the bride?

B.2 Experimental Sentences

1. (A) The boy is pulling the girl.
2. (P) The cat is chased by the dog.
3. (OC) It is the girl who the boy is pulling.
4. (SWhQ) Who is chasing the cat?
5. (SC) It is the man who is saving the woman.
6. (OWhQ) Who is the dog watching?
7. (P) The man is kissed by the woman.
8. (A) The man is saving the woman.
9. (SC) It is the man who the woman is kissing.
10. (SWhQ) Who is saving the woman?
11. (SC) It is the boy who is pulling the girl.
12. (OWhQ) Who is the boy pulling?
13. (A) The dog is watching the cat.

14. (SWhQ) Who is pulling the girl?

15. (OC) It is the cat with a dog is watching.

16. (P) The cat is watched by the dog.

17. (OWhQ) Who is the man saving?

18. (SC) It is the dog who is watching the cat.

19. (OC) It is the cat who the dog is chasing.

20. (P) The girl is pulled by the boy.

21. (A) The woman is kissing the man.

22. (OWhQ) Who is the dog chasing?

23. (A) The dog is chasing the cat.

24. (SWhQ) Who is watching the cat?

25. (SC) It is the dog who is chasing the cat.

26. (OWhQ) Who is the woman kissing?

27. (OC) It is the woman who the man is saving.

28. (SWhQ) Who is kissing the man?

29. (SC) It is the woman who is kissing the man.

30. (P) The woman is saved by the man.

Bibliography

Berndt, R. S. (1987). Symptom co-occurrence and dissociation in the interpretation of agrammatism. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language* (pp. 221–223*).* Hove, UK: Lawrence Erlbaum Associates Ltd.

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. Cognition. https://doi.org/10.1016/0010-0277(95)00682-6

Caplan, D., & Futter, C. (1986). Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language*, *27*(1), 117–134. https://doi.org/10.1016/0093-934X(86)90008-8

Caplan, D., Waters, G. S., & Hildebrandt, N. (1997). Determinants of sentence comprehension in aphasic patients in sentence-picture matching tasks. *Journal of Speech, Language, and Hearing Research, 40(3),* 542-55. Retrieved from http://pitt.idm.oclc.org/login?url=https://search.proquest.com/docview/232347696?accountid=14709

Caplan, D., & Hildebrandt, N. (1988). Disorders of syntactic comprehension. Cambridge, MA: MIT Press (Bradford Books).

Caplan, D., Hildebrandt, N., & Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain,* 119(3), 933–949. https://doi.org/10.1093/brain/119.3.933

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition, 21(2),* 117–175. https://doi.org/10.1016/0010-0277(85)90048-4

Caplan, D., Waters, G., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain and Language*, *101*(2), 103–150. https://doi.org/10.1016/J.BANDL.2006.06.225

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language, 3(4),* 572–582. https://doi.org/10.1016/0093-934X(76)90048-1

Chapey, R. (2008). *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th Ed.). Baltimore, MA: Lippincott Williams & Wilkins

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). Aphasiology, 26(10), 1250–1277. https://doi.org/10.1080/02687038.2012.693584

Chomsky, N. (1957). *Syntactic structure.* Berlin: Mouton, The Hague.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, Mass: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language.* New York: Praeger.

Chomsky, N. (1995). *The minimalist program.* Cambridge, Mass: MIT Press.

Chomsky, N. (1995.) *Bare phrase structure.* In G. Webelhuth (Ed.), *Government and binding theory and the minimalist program.* Cambridge, MA: Blackwell. Pp. 383–439.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Retrieved from http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Des Roches, C.,A., Vallila-Rohter, S., Villard, S., Tripodis, Y., Caplan, D., & Kiran, S. (2016). Evaluating treatment and generalization patterns of two theoretically motivated sentence comprehension therapies. *American Journal of Speech-Language Pathology 25(4),* 474-757. doi:http://dx.doi.org/10.1044/2016_AJSLP-15-0134

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review, 108*(4), 759-788. doi:10.1037/0033-295X.108.4.759.

Edwards, S. (2005). *Fluent aphasia*. New York, NY: Cambridge University Press.

Gialanella, B., Prometti, P., Vanoglio, F., Comini, L., & Santoro, R. (2016). Aphasia and activities of daily living in stroke patients. *European Journal of Physical and Rehabilitation Medicine.* Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27098299

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders.* Philadelphia: Lea and Febiger.

Goodglass, H., Christiansen, J. A., & Gallagher, R. E. (1993). Comparison of morphology and syntax in free narrative and structured tests: fluent vs. non-fluent aphasics. *Cortex, 29(3),* 377–407. Retrieved from: http://www.sciencedirect.com/science/article/pii/S001094521380250X

Grodzinsky, Y. 1984. The syntactic characterization of agrammatism. *Cognition, 16 (2),* 99–120. Retrieved from: https://doi.org/10.1016/0010-0277(84)90001-5

Grodzinsky, Y. 1986. Language deficits and the theory of syntax. *Brain and Language, 27(1),* 135–159. Retrieved from: https://doi.org/10.1016/0093-934X(86)90009-X

Hallowell, B. (2017). Aphasia and other acquired neurogenic language disorders: A guide for clinical excellence. San Diego, CA: Plural Publishing, Inc.

Jacobs, B. J., & Thompson, C. K. (2000). Cross-Modal Generalization Effects of Training Noncanonical Sentence Comprehension and Production in Agrammatic Aphasia. *Journal of Speech Language and Hearing Research, 43(1).* doi:10.1044/jslhr.4301.05

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes, 22(6),* 860-897. doi:10.1080/01690960601155284

Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: implications for aphasia. *Brain and Language, 41*, 275–288. doi:10.1016/0093-934X(91)90156-U

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

Love, T., & Oster, E. (2002). On the Categorization of Aphasic Typologies: The SOAP (A Test of Syntactic Complexity). *Journal of Psycholinguistic Research,31(5),* 503-529. doi:10.1023/A:1021208903394

McNeil, M. R., Pratt, S. R., Szuminsky, N., Sung, J. E., Fossett, T. R. D., Fassbinder, W., & Lim, K. Y. (2015). Reliability and validity of the computerized Revised Token Test: comparison of reading and listening versions in persons with and without aphasia. *Journal of Speech, Language, and Hearing Research: JSLHR,* 58(2), 311–324. https://doi.org/10.1044/2015_JSLHR-L-13-0030

Miyake, A., Carpenter, P. C., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: making normal adults perform like aphasic patients. *Cognitive Neuropsychology,* 11, 671–717. doi: 10.1080/02643299408251989

Paul, R., Norbury, C., & Gosse, C. (2018). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating.* St. Louis, MO: Elsevier.

Porch, B. E., (1967). Porch index of communicative ability: Theory and development. Chicago, IL: Riverside Publishing Company

Rochon, E., Laird, L., Bose, A., & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation,15(1),* 1-36. doi:10.1080/09602010343000327

Saffran, E. M., & Schwartz, M. F. (1988). Agrammatic comprehension it's not: Alternatives and implications. *Aphasiology, 2,* 389–394. doi: 10.1080/02687038808248943

Schuell, H. (1965). The Minnesota Test for Differential Diagnosis of Aphasia. Minneapolis: University of Minnesota Press.

Schröder, A., Burchert, F., & Stadie, N. (2014). Training-induced improvement of noncanonical sentence production does not generalize to comprehension: Evidence for modality-specific processes. Cognitive Neuropsychology, 32(3-4), 195-220. doi:10.1080/02643294.2014.968535

Schwartz, M. F., Linebarger, M. C., & Saffran, E. M. (1985). The status of the syntactic deficit theory of agrammatism. In M. L. Kean (Ed.), *Agrammatism* (pp. 83–104). New York: Academic Press.

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology, 8(1),* 19-54. doi:10.1080/02687039408248639

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early Understanding of Subject and Object Wh-Questions. *Infancy, 4(3),* 423-436. doi:10.1207/s15327078in0403_06

Shapiro, L. P. (1997). Tutorial: An Introduction to syntax. *Journal of Speech Language and Hearing Research, 40(2),* 254. doi:10.1044/jslhr.4002.254

Simmons-Mackie, N., & Kagan, A. (2007). Application of the ICF in aphasia. *Seminars in Speech and Language.* doi: 10.1055/s-2007-986521

Swinburn, K., Porter, G., & Howard, D. (2004). CAT: comprehensive aphasia test. Psychology Press.

Thompson, C. K., Tait, M. E., Ballard, K. J., & Fix, S. C. (1999). Agrammatic Aphasic Subjects Comprehension of Subject and Object Extracted Wh Questions. *Brain and Language,67(3),* 169-187. doi:10.1006/brln.1999.2052

Thompson, C.K., Weintraub, S., & Mesulam, M.M. (2012). Northwestern Anagram Test (NAT).

U.S. Department of Health & Human Services. (2018, June 15). *Aphasia*. Retrieved from https://www.nidcd.nih.gov/health/aphasia

Weintraub, S., Mesulam, M., Wieneke, C., Rademaker, A., Rogalski, E. J., & Thompson, C. K. (2009). The Northwestern Anagram Test: Measuring sentence production in primary progressive aphasia. *American Journal of Alzheimers Disease & Other Dementiasr,24(5),* 408-416. doi:10.1177/1533317509343104

Yoo, H. & Dickey, M.W. (2017). Aging effects and working memory in garden-path recovery. *Clinical Archives of Communication Disorders, 2(2),* 91-102. doi:10.21849/cacd.2017.00122