GENERALIZATION ACROSS VERB TYPES AFTER VERB NETWORK STRENGTHENING TREATMENT (VNeST): A TREATMENT STUDY

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF ARTS

---

by
Elizabeth Brophy
December 2016

Thesis Approvals:

Dr. Edwin Maas, Thesis Advisor, Department of Communication Sciences & Disorders
Dr. Gayle DeDe, Department of Communication Sciences & Disorders
Dr. Jamie Reilly, Department of Communication Sciences & Disorders

ABSTRACT

Research in communication disorders suggests that training linguistically complex forms will generalize to untrained, simpler forms with similar structural properties (see Thompson, 2007 for review). The present study investigated generalization patterns from transitive verbs to two classes of intransitive verbs following administration of Verb Network Strengthening Treatment (VNeST; Edmonds, Nadeau & Kiran, 2009). Based on the Argument Structure Complexity Hypothesis (ASCH; Thompson, 2003), it was predicted that greater generalization would occur to unergatives because unergatives bear a structural relationship to transitive verbs and unaccusatives do not. Results at post-treatment supported the hypothesis with both intransitive verb types showing generalization and slightly higher effect sizes observed for unergative than for unaccusative verbs. At maintenance, this pattern was not maintained due to improvements in production of unaccusative verbs. Results support the findings of Edmonds et al. (2009) that administration of VNeST results in gains on measures of untrained, semantically related verbs as well as standardized measures of lexical retrieval and connected speech. These results also suggest that training transitive verbs results in slight generalization to untrained intransitive verbs; however, it is inconclusive whether unergative and unaccusatives intransitives show differential improvement.

Dedicated to my cat

ACKNOWLEDGMENTS

This thesis would not have been possible without the support of many at Temple University, including Dr. Edwin Maas and Dr. Nadine Martin for their academic contributions to the development of this study, guidance during treatment, and review of the data; Dr. Jamie Reilly for stepping in as a committee member; Bridget Parkes for her contributions as a treating clinician and unfailing moral support; and Megan Montgomery for her hard work and constant optimism. I would also particularly like to thank Dr. Lisa Edmonds at Teachers College, Columbia University, for her advice and generous contribution of VNeST materials, and to Dr. Tara McAllister Byun at New York University for contributing additional probe materials. Most importantly, I would like to extend all of my gratitude to Dr. Gayle DeDe at Temple University for her guidance, contributions, and support through this process and without whom this thesis certainly would have fallen apart.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

A primary interest of clinicians working in the field of speech-language pathology is maximizing the impact and efficiency of language-based therapy. To this end, speech and language targets should be selected that will result in the most overall change in the language system and generalization to untrained contexts. Research using single subject designs has suggested that greater gains result from training linguistically complex stimuli compared to simpler forms. The present study investigated whether clinical treatment involving transitive verbs generalized to two classes of syntactically simpler intransitive verbs.

In domains including phonology, semantics, and syntax, researchers have noted that training complex language structures results in generalization to untrained, related forms with simpler structures (Gierut, 2001; Kiran, 2007; Kiran & Thompson, 2003; Maas et al., 2002; Thompson, 2007; Thompson & Shapiro, 2007; Thompson, Shapiro, Kiran, & Sobecks, 2003). This phenomenon, dubbed the Complexity Account of Treatment Efficacy (CATE; Thompson et al., 2003), is believed to emerge when there is a super- and subset relationship between the target item and untrained form. Gierut (2001) observed that training phonologically marked targets generalizes to untrained targets with unmarked features in the remediation of phonological disorders among children. Similarly, in treatment of sentence production deficits among adults with aphasia, Thompson and Shapiro (2007) found that targeting syntactic structures that trigger phrasal movement, such as passive sentences, resulted in generalization to related

canonical forms, such as active sentences, while there was no generalization in the opposite direction. Finally, in treatment of naming deficits in aphasia, Kiran and Thompson (2003) found that the training of atypical noun exemplars generalized to typical noun exemplars within a single semantic category. In each of these areas, the complex training targets encompass features of the simpler form, resulting in implicit generalization without direct treatment.

One domain that has received comparably little attention in research on the complexity effect is verb treatment. Verbs play a central role in both the meaning and structure of sentences, and individuals with language disorders, particularly agrammatic aphasia, may have disproportionate difficulty in producing verbs compared to nouns (Berndt, Mitchum, Haendiges, & Sandson, 1997; Conti-Ramsden & Jones, 1997; Zingeser & Berndt, 1990). As a result, verbs have become the primary focus of treatment in a number of protocols, such as Cueing Verbs Treatment (CVT; Loverso, Prescott, & Selinger, 1988), Sentence Production Program for Aphasia (SPPA; Helm-Estabrooks, Nicholas, & Helm, 2000), and Verb Network Strengthening Treatment (VNeST; Edmonds, Nadeau, & Kiran, 2009). While CVT and SPPA involve eliciting verbs in a hierarchical progression from simpler to more complex forms, VNeST is unique in that treatment begins with transitive verbs within a simple sentence frame.

Verb Network Strengthening Treatment

VNeST, developed by Edmonds and colleagues in 2009, is theoretically motivated by research that shows that activation of a target verb (e.g. *chop*) results in bidirectional priming with its associated network of thematic roles, including agents

(*lumberjack*), patients (*wood*), and instruments (*saw*) (Ferretti, McRae, & Hatherell, 2001; McRae, Hare, Elman, & Ferretti, 2005). Additionally, individual verbs can be associated with a variety of event schemas (e.g. *lumberjack-chop-wood, chef-chop-onions*), resulting in widespread semantic network activation for both the target verb as well as semantically related verbs and nouns. VNeST aims to use this heightened activation to promote generalization during clinical treatment by targeting production of a transitive verb and its associated arguments in the context of a basic subject-verb-object sentence structure. This production is then elaborated in a series of steps designed to heighten activation of both the target verb and semantically related arguments and verbs, thereby promoting effortful semantic processing across grammatical categories, strengthening distinctions between potential semantic competitors, and developing associations with personally salient memories (Edmonds, 2016).

VNeST has been investigated in a series of single-subject designs in English and Korean (Edmonds & Babb, 2011; Edmonds et al., 2009; Furnas & Edmonds, 2014; Kwag, Sun, Kim, & Cheon, 2014) as well as a group analysis of outcomes for 11 participants (Edmonds, Mammino, & Ojeda, 2014). In the original 2009 investigation, Edmonds and colleagues administered VNeST for two-hour long sessions twice per week for six weeks with four participants: two with moderate transcortical motor aphasia and two with moderate conduction aphasia. The authors found that for all four participants, administration of VNeST using 10 transitive verbs resulted in generalization to a set of 10 untargeted, semantically related verbs as well as improvements in noun retrieval as measured using the Boston Naming Test (BNT; Kaplan, Goodglass & Weintraub, 1983), sentence production on the Northwestern Verb Production Battery (NVPB; Thompson,

2002), and in overall aphasia severity as measured using the Aphasia Quotient of the Western Aphasia Battery (WAB; Kertesz, 1982). Additionally, three of four participants showed improvement in verb naming on the NVPB and improvements in measures of lexical retrieval during connected speech. A subsequent literature review by Edmonds (2016) found that in six studies encompassing 22 total participants, the majority demonstrated generalized improvements on standardized measures across several domains: 86% of those tested showed improvement in noun naming, 58% improved in verb naming, and 59% showed increases in complete utterances produced during a measure of connected speech. Additionally, 11 of 11 participants' caregivers who were administered the Communicative Effectiveness Index (CETI; Lomas et al., 1989) reported improvements, suggesting that improved communication as a result of VNeST may transfer to functional settings at home.

Several participants have also shown syntactic generalization as a result of this treatment. Edmonds (2016) noted that across the current literature, nearly 75% of tested participants showed generalized improvement in sentence production, including production of verbs with untargeted argument structures (e.g. intransitive and ditransitive verbs). Edmonds theorized that because VNeST targets verb production in the context of repeated elicitation of canonical subject-verb-object utterances, it may also strengthen basic sentence syntax for related verbs bearing the same canonical structure. However, no research has yet systematically investigated generalization properties of transitive verbs based on their syntactic or structural features during clinical treatment.

Verb Structure

In addition to encoding semantic properties of actions, verbs also specify obligatory participants of those actions in the form of argument structure. Intransitive verbs, such as *sleep,* assign a single obligatory argument to the subject position, as in the sentence *John sleeps*. Transitive verbs, such as *kick*, assign obligatory arguments to both the subject and object position, as in *John kicks the ball*. Behavioral research has suggested those verbs that assign a greater number of optional or obligatory arguments, such as transitives and ditransitives, are associated with greater difficulty in production than intransitives (Kim & Thompson, 2000; Shapiro & Levine, 1990; Shapiro, Zurif, & Grimshaw, 1997; Thompson, Lange, Schneider, & Shapiro, 1997). Additionally, individuals with agrammatic aphasia may have specific retrieval deficits for verbs containing more than one argument. Thompson et al. (1997) found that individuals with agrammatic aphasia produced intransitive verbs with greater accuracy than transitive and ditransitive verbs in confrontation naming and single word elicitation tasks, even when no arguments were produced, while Kim and Thompson (2000) found the same hierarchy within naming and categorization tasks. Additional research by Shapiro and colleagues found that both healthy adults and individuals with nonfluent aphasia demonstrated slowed reaction times during cross-modal lexical decision tasks using transitive and ditransitive verbs (Shapiro & Levine, 1990). These results suggest that verb retrieval may be influenced by the number of obligatory and optional arguments assigned by a verb. Those verbs that assign a greater number of obligatory and optional arguments are more difficult to produce than those that assign only one, and therefore may be considered more complex.

However, even verbs with the same number of obligatory arguments may have structural differences that impact their access and production. *Unaccusative* and *unergative* verbs are two classes of intransitive verbs that differ based on their underlying syntactic structure, as illustrated in the following sentences adapted from Friedmann, Taranto, Shapiro, and Swinney (2008):

1. Unergative: The $\text{bird}_{\text{AGENT}}$ chirped.
2. Unaccusative: The $\text{leaf}_{\text{i THEME}}$ fell t_{i}.

While both sentences contain an intransitive verb with a single argument occupying the surface (s-structure) subject position, they differ in both the underlying location of the argument and the thematic role that is assigned in their deep (d-structure) representations. *Unergative* verbs (such as *chirp*) assign an agent thematic role to the external (subject) position and do not assign a verb-internal object. *Unaccusative* verbs (such as *fall*) assign a theme to the internal (object) position in d-structure, and the theme is then raised to the subject position in s-structure, leaving a trace (*t*) behind. Such verbs can be differentiated using the nominal modifier test: if the nominal form of the verb can be used to modify a noun, it is an unaccusative verb (c.f. *The fallen leaf* versus **The chirped bird*).

According to McAllister, Bachrach, Waters, Michaud, and Caplan (2009), this movement operation can be considered analogous to the movement of the internal object to subject position in the generation of passive sentences from the canonical active form. In addition to their structural differences, these two verb types may also convey semantic

differences. Unergative verbs typically encode arguments performing intentional actions (verbs such as *swim* and *bark*), while unaccusatives typically encode arguments undergoing events or changing state (such as *melt* and *sink*; Black & Chiat, 2008). Several researchers have suggested that these factors may contribute to the relative complexity of unaccusative verb production.

Unaccusative Verb Production in Aphasia

Like other non-canonical or movement-induced syntactic structures, such as passive sentences, the unaccusative construction may be particularly difficult for individuals with aphasia to comprehend and produce (Kegl, 1995). Thompson (2003) examined patterns of verb production and found that individuals with agrammatic aphasia were less likely to produce unaccusative intransitive verbs in naming and narrative elicitation tasks compared to unergative verbs. Based on this observation, Thompson proposed the Argument Structure Complexity Hypothesis (ASCH): verb complexity is determined by the combination of the number, type, and movement of arguments contained within a verb's lexical entry, and verbs with greater argument structure complexity are more difficult for individuals with aphasia to produce. Later studies observed that individuals with agrammatic aphasia were indeed less likely to correctly name unaccusative verbs as well as produce them in connected speech compared to participants with anomic or Wernicke's aphasia (Bastiaanse & van Zonneveld, 2005) and compared to healthy adults (McAllister et al., 2009). Additional behavioral and imaging studies on healthy adults also lend support to the ASCH. Barbieri et al. (2011) found that unaccusative verb production was both slower and less accurate during sentence

completion tasks with both healthy and nonfluent aphasic adults. More recently, an fMRI investigation by Meltzer-Asscher, Mack, Barbieri, and Thompson (2015) found that during lexical decision tasks, presentation of unaccusative verbs resulted in longer response times compared to transitive verbs as well as increased activation in the left inferior frontal gyrus, an area associated with non-canonical sentence processing.

To summarize, the structural complexity of verbs appears to be determined by a combination of several factors. Verbs that assign more than one obligatory arguments are more complex than those that assign only one; additionally, verbs that assign arguments that undergo a movement transformation from d-structure to s-structure are more complex than those that do not. These factors would support a hierarchy of verb complexity in which transitives are more complex than intransitives, and in which unaccusative intransitives are more complex than unergative intransitives, as predicted by Thompson and colleagues (Thompson, 2003; Thompson & Shapiro, 1997). According to the principles of CATE, a complexity effect may emerge in which the training of transitive verbs generalizes to intransitive verbs; however, this account would also predict that greater generalization would occur between transitives and unergative intransitives because they bear a structural relationship and unaccusatives do not. To date, no study has examined the syntactic generalization properties of transitive to unergative and unaccusative verbs during clinical intervention.

Objectives

The purpose of the present study was to investigate the implicit generalization properties of transitive to two types of intransitive verbs following of VNeST. Because

VNeST targets verb and argument structure production in the context of a basic sentence frame, it is ideal for examining the syntactic generalization properties of transitive verbs to other forms. The present study sought to replicate the findings from Edmonds et al. (2009) and subsequent research that VNeST results in semantic generalization while also exploring the syntactic generalization properties of transitive to intransitive verbs using controlled sets of probes.

Research Questions

1. (Replication) Does training a set of transitive verbs using VNeST generalize to a set of untrained, semantically related transitive verbs and agent/patient pairs as well as generalized improvements on measures of lexical retrieval, connected speech, and quality of life?

2. Does training a set of transitive verbs using VNeST generalize to untrained, semantically unrelated intransitive verbs, including agent-assigning (unergative) intransitive verbs and non-agent-assigning (unaccusative) intransitive verbs?

Predictions

For Research Question #1, it was predicted that the present study would replicate the findings of Edmonds et al. (2009) and subsequent investigations by finding generalization to both untrained verbs as well as measures of lexical retrieval, connected speech, and quality of life. For Research Question #2, it was predicted that training transitive verbs would result in generalization to untrained intransitive verbs during a

weekly probe elicitation task because transitive verbs are more complex than intransitive verbs and therefore may exhibit a generalization pattern in line with CATE (Thompson et al., 2003). In accordance with the ASCH (Thompson, 2003), it was also predicted that greater generalization would occur to unergative than to unaccusative intransitive verbs because unergatives bear a structural similarity to transitives and unaccusatives do not.

CHAPTER 2

METHOD

A single subject, ABA, repeated probe experimental design with an unrelated, untreated control condition was used. The study consisted of four phases: (a) baseline testing, (b) treatment of trained items with weekly probes, (c) post-treatment testing one week after final treatment, and (d) maintenance at one month post final treatment. Weekly probes consisted of a sentence elicitation task using pictures of trained transitive verbs to monitor treatment effects and untrained verb targets to monitor semantic and syntactic generalization. A single-word adjective retrieval task was administered weekly for experimental control to ensure that treatment effects were not due to generalized semantic improvements.

Participant

Several inclusionary and exclusionary criteria were set, including 1) aphasia due to single left hemisphere stroke, 2) native English speaker, 3) impaired lexical access for nouns and verbs, 4) no prior history of speech, language, or learning disabilities, and 5) adequate vision and hearing to participate in treatment. A participant meeting these criteria was recruited from the Temple University Speech-Language-Hearing Center (TUSLHC).

XH was a right-handed, 48-year-old male with chronic aphasia resulting from a left middle cerebral artery (MCA) infarct approximately 28 months prior to the present study. According to radiology reports, XH sustained multifocal occlusions to the M2

branch of the left distribution of the MCA with partial recanalization and subsequent re-occlusion, resulting in right-sided weakness and aphasia. He remained hospitalized for two weeks following his CVA. Upon discharge, XH received outpatient speech therapy at a health center and TUSLHC for two years, ending one and half months prior to the beginning of the present study. At TUSLHC, XH received a modified and abridged version of VNeST for ten weeks at a low weekly dosage (approximately 15 minutes per week). Prior to his stroke, XH reported having completed seven years of formal education and previously having worked as a recidivism counselor through a religious institution. He reported no known history of visual impairment or speech, language, or learning disabilities.

Tests Administered

XH passed a hearing screening at 25 dB HL and a screening for cognitive deficits and visual neglect using the Cognitive Screen subtest of the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2005). XH was also administered portions of the Apraxia Battery for Adults - Second Edition (ABA-2; Dabul, 2000) and was found to demonstrate a mild apraxia of speech.

XH underwent comprehensive testing during the baseline phase in order to characterize the type and severity of his aphasia. Further assessment included evaluation of lexical retrieval for nouns and verbs, connected speech, and functional measures of disability. All tests were administered by a licensed speech-language pathologist or by the author, a graduate student in speech-language pathology. Pre- and post-treatment scores for all measures can be found in Table 1.

Table 1. Pre- and post-treatment testing.

Language testing	Pre	Post
Comprehensive Aphasia Test[a]		
Mean modality score	52	52
Comprehension of spoken language	60	62
Comprehension of written language	50	46
Repetition	55	53
Naming	54	54
Spoken picture description	54	51
Reading	46	46
Writing	47	49
Written picture description	53	53
Disability quotient	56	56
Boston Naming Test[b]		
Standard Form (n=60)	28/60	35/60
Northwestern Assessment of Verbs and Sentences[c]		
Verb Naming Test (n=18)	72.2%	83.3%
1-place verbs (n=4)	75.0%	100.0%
2-place verbs (n=7)	85.7%	71.45%
3-place verbs (n=7)	57.1%	85.7%
Argument Structure Production Test (n=16)	68.8%	81.35%
1-place verb sentences (n=4)	100.0%	100.0%
2-place verb sentences (n=5)	70.0%	80.0%
3-place verb sentences (n=7)	57.1%	71.4%
Spoken narrative sample		
Percent correct information units (%CIUs)	50.6%	66.2%
Verb type-token ratio[d]	50.0	50.0
Quality of life measures		
Aphasia Communication Outcome Measure		
Adaptive ACOM	52.9	48.04
Stroke and Aphasia Quality of Life Scale		
Overall	4.6	4.5
Physical	4.9	5.0
Communication	3.5	3.6
Psychosocial	4.8	4.6

Notes. All reported scores are standard scores unless otherwise indicated. [a]T-scores. [b]Raw scores. [c]Percent correct using modified administration procedures described in text. [d]Raw ratio.

Standardized Measures of Aphasia

The Comprehensive Aphasia Test (CAT) was administered to determine the nature and severity of XH's aphasia. Because the CAT does not provide an overall severity score, a mean modality score of 52 was calculated to evaluate pre- to post-treatment changes across subtests (following the procedures of Winans-Mitrik et al., 2014). XH's performance on the CAT was characterized by largely intact auditory comprehension with impairments in written comprehension, naming, repetition, reading, and writing, consistent with moderate fluent conduction aphasia. XH demonstrated inefficient and impaired naming characterized by perseverations, self-corrections, and circumlocutions as well as occasional semantic paraphasias (e.g. *jazz* for saxophone). During naming tasks, XH was frequently observed to attempt self-cueing by writing the first letter or multiple letters of a word. However, he was typically unable to produce the word he had written. According to the interactive model of lexical retrieval, XH's deficits in naming are suggestive of a deficit in encoding phonological output and his errors are typical for individuals with conduction aphasia (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). XH also demonstrated deficits in reading, with disproportionately impaired reading of low-frequency, low-imageability words and no-response errors for all function and non-words, suggestive of deficits to both the lexical-semantic and grapheme-to-phoneme correspondence routes in the dual route model of reading (Coltheart, 2005). Similar deficits were observed in writing, with preserved ability to copy letters but impaired writing of words in picture naming and dictation tasks. In tests of sentence processing, XH demonstrated intact comprehension of canonical and non-canonical sentences but mild deficits in comprehension of reversible

sentences, suggestive of an impairment in mapping syntactic information onto underlying semantic representations (Linebarger, Saffran & Schwartz, 1983).

The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) and the Northwestern Assessment of Verbs and Sentences (NAVS; Thompson, 2011) were administered in order to further evaluate XH's single word lexical retrieval of nouns and verbs as well as verb argument structure production. The results of the BNT confirmed a lexical retrieval deficit for confrontation naming of objects characterized by frequent self-corrections, semantic paraphasias, and circumlocutions. On the Verb Naming subtest of the NAVS, XH demonstrated impaired single-word naming of one-, two-, and three-place verbs, substituting semantically light verbs such as *giving water* for *pour* and circumlocutions such as *He taking steps - no, he taking baby steps* for *crawl*.[1] The Argument Structure Production subtest was used to assess XH's ability to produce one-, two-, and three-place verbs and their associated obligatory and optional arguments to describe pictures. While the standard administration requires the examiner to supply the verb and noun targets, administration of this subtest was modified by removing them in order to better assess XH's independent lexical retrieval. Additionally, any practice, training, or generalization targets used in the present study were not administered during pre-testing. While XH reached ceiling in production of argument structure for one-place verbs, he was observed to produce atypical substitutions of two- and three-place verbs, such as producing the three-place response *The woman is giving the man giggles* for the two-place target *The girl is tickling the boy*.

[1] All one-place verbs on the NAVS are unergative verbs, therefore the test does not provide norm-referenced information about the production or comprehension of unaccusative verbs.
[2] While Edmonds et al., 2009 probed all 10 training items weekly as well as their matched 10 untrained,

Connected Speech

A one-minute connected speech sample was obtained using a picture scene. The speech sample was analyzed using percent correct information units (%CIUs; Nicholas & Brookshire, 1993) and verb type-token ratio in order to establish baseline measures of accuracy and efficiency in speech production and diversity of verb production in connected speech. XH was observed to produce inefficient speech characterized by multiple revisions of sentences, neologisms, and semantic paraphasias. Additionally, the majority of verbs were semantically light (e.g. *taking*, *doing*).

Quality of Life Measures

The Stroke and Aphasia Quality of Life Scale (SAQOL-39; Hilari, Byng, Lamping, & Smith, 2003) and Aphasia Communication Outcome Measure (ACOM; Hula et al., 2015) were administered to assess XH's affective responses to stroke-based deficits in physical, social, cognitive, and emotional domains. On both assessments, XH reported relatively little impairment in physical or cognitive function since his stroke. XH reported slight reduction in his overall social and communicative function since his stroke, noting particular difficulty in reading and writing without support.

Stimuli Development

Training Targets

Following the procedures of Edmonds and colleagues, a set of ten transitive verbs was selected for use in biweekly treatment during the study's treatment phase using materials supplied by Edmonds (personal communication, May 7, 2016). Because these

ten verbs were developed for use in research on semantic generalization, each item was also associated with a matched verb (e.g. *bake* matched with *fry*) that had been normed for semantic relatedness on groups of younger and older healthy adults (Edmonds et al., 2009). In addition to their use in treatment, five of these verbs were randomly selected as training probes. The remaining five verbs' associated pairs were used as semantic generalization probes.[2]

Probes

A set of 27 probes was developed to monitor treatment effects and generalization to untrained verbs using a sentence elicitation task. Probes were grouped into four categories: (a) *training probes,* consisting of five randomly selected trained transitive verbs to monitor the effects of treatment on trained verbs; (b) *semantic generalization probes*, consisting of five transitive verbs that were semantically matched to the remaining trained verbs to assess semantic generalization to untrained transitive verbs; (c) *unergative probes*, consisting of nine untrained unergative verbs to assess syntactic generalization to untrained verbs; and (d) *unaccusative probes*, consisting of eight untrained unaccusative verbs to assess syntactic generalization to untrained verbs. The unequal number of probes in each set was due primarily to lexical variable constraints; for example, unaccusative verbs tend to be less imageable than unergatives and therefore could not be adequately depicted in picture stimuli or matched with the training verbs (see Table 2). Consistent with previous studies (e.g. McAllister, Bachrach, Waters, Michaud, & Caplan, 2009), a mix of alternating (e.g. *break,* that can be used in a

[2] While Edmonds et al., 2009 probed all 10 training items weekly as well as their matched 10 untrained, semantically related probes, the present study used sets of five each in order to keep weekly probes from exceeding 30 minutes in length.

transitive construction) and non-alternating unaccusatives (e.g. *rise*, that cannot be used in a transitive construction) were chosen for use in treatment. In order to limit the potential confounds of semantic generalization, the unergative and unaccusative probes were judged to be semantically unrelated to the target items via consensus between the author and two healthy adult volunteers.

Using data obtained from Balota et al. (2007) and Wilson (1988), all probes were matched to the ten training items by six phonological and lexical variables: frequency ($t(35) = .8005$), number of syllables ($t(35) = .2861$), phonological neighborhood density ($t(35) = .6405$), concreteness ($t(22) = .3402$), imageability ($t(25) = .8784$), and age of acquisition ($t(7) = .6093$). These variables were selected to ensure that phonological and semantic features did not influence patterns of generalization from treated targets to probes. Not all data were available for each item, which accounts for differences in the degrees of freedom. See Table 2 for details.

Table 2. Training and probe sets compared by phonological and lexical variables.

Verb	Freq[a]	NSyll[b]	PhonND[c]	Conc[d]	Img[e]	AOA[f]
Training targets						
bend	2.89	1.00	18.00	-	460.00	-
boil	2.48	2.00	17.00	467.00	533.00	-
brush	2.86	1.00	4.00	589.00	570.00	214.00
chop	2.84	1.00	16.00	555.00	575.00	317.00
fly	3.64	1.00	13.00	525.00	582.00	-
measure	2.73	2.00	0.00	366.00	379.00	344.00
pull	3.87	1.00	21.00	360.00	446.00	-
scrub	2.50	1.00	2.00	-	-	-
sew	2.45	1.00	37.00	-	478.00	-
stir	2.48	1.00	6.00	-	-	-
Average (all training targets, n=10)	2.87	1.20	13.40	77.00	502.88	291.67
SD	0.42	11.10	68.60	97.00	73.59	0.50
Probes - training						
boil	2.48	2.00	21.00	467.00	533.00	-
pull	3.87	1.00	6.00	360.00	446.00	-
scrub	2.50	1.00	17.00	-	-	-
sew	2.45	1.00	2.00	-	478.00	-
stir	2.48	1.00	37.00	-	-	-
Average (n=5)	2.76	1.20	16.60	413.50	485.67	n/a
SD	0.62	0.45	13.79	75.66	44.00	n/a
Probes – semantic[g]						
drive	3.89	1.00	7.00	-	-	-
fold	2.64	1.00	25.00	-	-	-
shave	2.85	1.00	17.00	-	-	-
slice	2.64	1.00	6.00	443.00	507.00	292.00
weigh	2.56	1.00	24.00	363.00	411.00	-
Average (n=5)	2.92	1.00	15.80	403.00	459.00	292.00
SD	0.56	0.00	9.04	56.57	67.88	n/a
Probes - unergative						
cry	3.53	1.00	7.00	436.00	478.00	-
dance	3.88	1.00	5.00	502.00	510.00	-
dive	2.82	1.00	14.00	494.00	586.00	322.00
howl	2.03	1.00	24.00	434.00	536.00	269.00
jump	3.55	1.00	8.00	449.00	506.00	222.00
pray	3.27	1.00	14.00	372.00	444.00	-
shiver	1.90	2.00	6.00	455.00	578.00	308.00
sleep	4.07	1.00	12.00	484.00	530.00	-
wink	2.26	1.00	13.00	515.00	501.00	-

Average (n=9)	3.03	1.11	11.44	460.11	518.78	280.25
SD	0.81	0.33	5.88	44.25	45.08	44.84
Probes - unaccusative						
bounce	2.70	1.00	3.00	-	-	-
choke	2.52	1.00	16.00	-	-	-
crash	3.17	1.00	11.00	-	-	-
fall	3.78	1.00	35.00	409.00	547.00	-
float	2.58	1.00	12.00	451.00	525.00	-
melt	2.57	1.00	13.00	413.00	461.00	-
rise	3.15	1.00	37.00	368.00	451.00	-
sink	2.94	1.00	18.00	590.00	599.00	219.00
Average (n=8)	2.93	1.00	18.13	446.20	516.60	219.00
SD	0.43	0.00	11.89	85.59	61.60	n/a
Average (all probes, n=27)	2.93	1.07	15.19	444.72	506.68	272.00
SD	0.61	0.27	9.92	60.13	51.42	43.62

Notes. All phonological and lexical variable data obtained from Balota et al. (2007) and Wilson (1988). Missing data is marked with a hyphen (-). [a]Frequency. [b]Number of syllables. [c]Phonological neighborhood density. [d]Concreteness. [e]Imageability. [f]Age of acquisition. [g]Semantic relatedness ratings for paired targets and probes (e.g. *fly/drive*) can be found in Edmonds el al. (2009).

Each verb and corresponding subject and object, if applicable, was depicted in a black-and-white line drawing measuring approximately 4" x 6" and centered on 8 ½" x 11" white paper. Images were obtained from multiple sources, including original development with norming on a population of healthy adults by Edmonds and colleagues as well from prior research on unaccusative and unergative verb production by McAllister and colleagues (2009). The remaining unergative and unaccusative images were independently developed by the author and were informally normed on three undergraduate volunteers.

Control Task

A single word adjective retrieval task was administered using materials supplied by Edmonds and colleagues in order to maintain experimental control and ensure that

semantic generalization between verbs was not due to overall improvements in semantic processing. This task required the participant to complete a sentence with a synonym for a provided adjective; for example, "Someone who is sick is also said to be…" (target *ill*). The adjective targets in this task were matched to the original set of training verbs on frequency, imageability, and familiarity; additionally, the task was normed on a population of healthy adults. See Edmonds, Nadeau, and Kiran (2009) for details.

Procedures

Treatment was administered twice per week for up to two-hour sessions (for a maximum of four hours per week) for a duration of six weeks. Probes were administered in the first half-hour of the first session each week by a trained undergraduate student in Communication Sciences and Disorders. Eight of twelve (66.7%) treatments were administered by the author; the remaining four treatments (33.3%) were administered by a second trained graduate student clinician with comparable experience in VNeST when the author was not available. Fidelity checks were administered by a trained undergraduate student for both clinicians in order to ensure adherence to the treatment protocol. See Fidelity for details.

Treatment Procedures

Treatment procedures followed the VNeST protocol as described in Edmonds et al. (2009) as well as the tutorial described in Edmonds (2014) with slight modifications. Briefly, the original treatment protocol involved a series of elaborations to a target verb, which was initially presented on an index card flanked by cards depicting

WHO and WHAT (e.g. WHO *bake* WHAT). The participant was then asked to generate potential agent/patient pairs corresponding with the verb (e.g. *My wife/bake/cookies* or *The chef/bake/wedding cake*). Minimal semantic cues (e.g. "Can you think of someone who *bakes* for their job?") or maximal cues (four written options consisting of one target and three foils) were provided by the clinician as necessary. No phonological cues were provided. Once five pairs were generated, the participant read each set aloud to facilitate production of a complete subject-verb-object utterance. The participant then chose one set to elaborate using cards depicting wh-questions WHERE, WHEN, and WHY (e.g. *The chef/bake/wedding cake/in the kitchen/before the wedding/for the bride)*. The treatment materials were removed and the participant then performed a semantic judgment task. The clinician read 12 sentences containing the target verb, varying by semantic plausibility (e.g. *The chef bakes the mailbox*) and the participant made a yes/no decision whether the sentence was semantically correct. The participant was then asked to name the target verb independently, and the initial step (generating agent/patient pairs) was repeated with no cues provided. Each target was concluded when the participant generated three schemas independently or could no longer generate agents and patients without clinician support.

Slight modifications to the original protocol were implemented. While Edmonds et al. (2009) required generation of three agent/patient pairs for each target verb, five pairs were generated in the current administration to increase the difficulty of the task as XH had prior experience with VNeST. XH was encouraged to avoid using the non-specific proper names (e.g. *John, Sally*) as agents; instead, he was prompted to think of personally salient relationships, such as a brother who cooked for the target verb *chop*.

All written items were read aloud via choral reading with the clinician including the full subject, verb, and object sentence frames (as in Edmonds & Babb, 2011) due to XH's impairments in oral reading. XH was permitted to write his own responses to facilitate word retrieval, and was frequently observed to write the first letter of a word as a form of self-cueing. The wh-question responses were both written and then read aloud in order to provide additional oral reading practice. Finally, side conversations about personally relevant topics addressed during treatment were permitted to maintain participant interest. For example, XH formerly worked as a line cook and therefore frequently engaged in conversation during treatment on the targeted verbs *boil* and *chop*.

Probe Procedures

Verb and adjective probes were administered once a week for thirty minutes at the beginning of the first session each week by an independent administrator, an undergraduate student in Communication Sciences and Disorders, who received training in administration procedures by the author as well as a licensed speech-language pathologist familiar with VNeST. The order of presentation of the verb probe and adjective probe tasks was alternated each week.

Verb probes were elicited using a black-and-white illustration of each probe verb (described under Stimuli). These were presented in a pseudo-randomized order with no verbs of the same subtype presented sequentially to reduce syntactic priming effects between targets of the same structure. For each image, XH was instructed: "Make a sentence about this picture and include this (the subject), what he/she/it is doing (the action), and this (the object)." A maximum of two prompts were provided during each

probe administration in order to elicit more specific responses. For example, if XH produced *The bartender is making a drink* instead of *The bartender is stirring the drink*, the examiner would prompt "Can you think of a more specific word for what he's doing?" Prompts for specificity were provided to elicit specific agents and patients for illustrations depicting transitive probes per the procedures of Edmonds et al. (2009), but were not targeted for intransitive probes; because the intransitive probes were developed to assess syntactic, rather than semantic, generalization, the intransitive probes were scored as correct when they contained any semantically appropriate noun in the subject position, but specific schemas were not targeted. For example, *she, the lady,* or *the woman* were all accepted for the target sentence *The woman is crying,* while *He is crying* was not.

Adjective probes were elicited using a single-word adjective retrieval task developed by Edmonds et al. (2009) that had been normed on a population of healthy adults. XH was instructed to complete a sentence with another word with the same meaning (e.g. *Someone who is sick is also said to be…* for the target *ill*). Responses were scored as correct when they matched the target word. No substitutions were accepted.

Scoring

Responses to the probe task were scored as correct when they contained a correct subject, verb, and object, if applicable. Word order and morphology were not considered in scoring probes. For the transitive probes, responses had to include specific agents and patients per the scoring criteria outlined in Edmonds et al. (2009). For the intransitive probes, responses had to include a semantically appropriate subject but specific nouns

were not targeted. A list of acceptable substitutes for each content word was developed to account for dialect variation and perceptual misinterpretations of picture stimuli. Verb substitutions were only accepted when the response bore a close semantic relationship to the target and shared the same argument structure; for example, *cough* was accepted for *choke* because they are semantically related and are both unaccusative verbs.

Reliability

Following the reliability procedures outlined in Edmonds et al. (2009), a trained student in Communication Sciences and Disorders independently transcribed and scored 33% of the sentence elicitation and control probes to determine inter-rater reliability. Using these procedures, point-by-point reliability was found to be 91.3% for verb probes and 93.9% for control probes.

Fidelity

Following the procedures described in Edmonds et al. (2009), a trained student in Communication Sciences and Disorders observed 33% of the treatment sessions live and compared the administered treatment to a checklist of the protocol to determine treatment fidelity for both clinicians. According to this measure, the treatment protocol was followed with 100% fidelity for the author and 95.8% fidelity for the second clinician. Deviations from the protocol occurred when all steps for a target verb could not be completed due to time constraints.

CHAPTER 3

RESULTS

Data were analyzed using visual inspection and calculation of effect sizes from pre-treatment baselines to post-treatment using the procedures described below.

Effect Size Calculation

Effect sizes were calculated by subtracting the post-treatment data point from the baseline average and dividing by the standard deviation of the baseline sample per the procedures described by Beeson and Robey (2006):

$$d_1 = \frac{\bar{x}_{A_2} - \bar{x}_{A_1}}{S_{A_1}}$$

In the formula, $\bar{x}_{A1}$ represents the average of the baseline phase, $\bar{x}_{A2}$ represents the average of the post-treatment phase, and S_{A1} represents the standard deviation of the baseline phase. Following the recommendations of Beeson and Robey (2006) for single-subject aphasia treatments, criteria of 2.6, 3.9, and 5.8 were set for small, medium, and large effect sizes, respectively.

Overall Probe Results

Overall results of the weekly sentence elicitation and control probes for XH are presented in Figures 1 and 2. Effect sizes for each probe type at post-treatment and maintenance are presented in Figure 3.

XH demonstrated an overall increase in performance on the weekly probe measure of correct production of subject, verb, and (if applicable) object of all trained and untrained sentence probes (n=27) from an average of 39.5% accuracy across three baseline trials ($\bar{x}$=10.67 items) to a post-treatment score of 59.3% accuracy (n=16 items). Effect size collapsed across all training and generalization probes from pre- to post-treatment was found to be 9.24, an indicator of large positive gains in performance. Treatment effects for all trained and untrained items were largely maintained after one month at 55.5% accuracy (n=15 items).

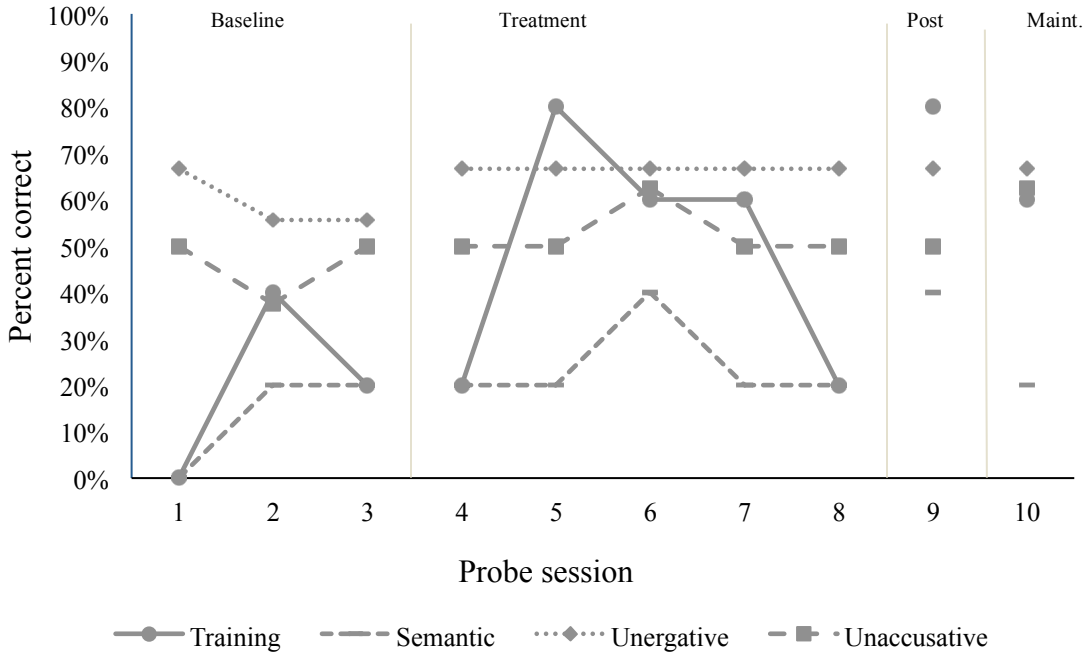
Research Question #1

Trained and Untrained Semantically Related Probes

XH's production of trained agent, verb, and patient sentences as measured using training probes (n=5) improved from a baseline average of 20.0% accuracy across three trials ($\bar{x}$=1 item) to a post-treatment accuracy of 80.0% (n=4 items), with a total effect size of 3.00, an indicator of small improvement. At maintenance, trained verbs were produced with 60% accuracy (n=3 items) with effect size of 2.00. See Figures 1 and 3.

Using the data from untrained, semantically related probes (n=5), XH improved from a baseline average performance of 13.4% accuracy ($\bar{x}$=.67 items) to a post-treatment score of 40.0% accuracy (n=2 items), with a total effect size of 2.31. At maintenance, untrained, semantically related probes were produced with 20.0% accuracy (n=1 item) with effect size of .58. See Figures 1 and 3.

Figure 1. Number of correct responses on weekly sentence elicitation probes.



No improvement was observed using the adjective control task with effect size at post-treatment and maintenance calculated to be 0, indicating that semantic generalization to targets other than verbs and nouns did not occur during treatment. See Figures 2 and 3.

Figure 2. Percent correct production of single words in adjective retrieval control.

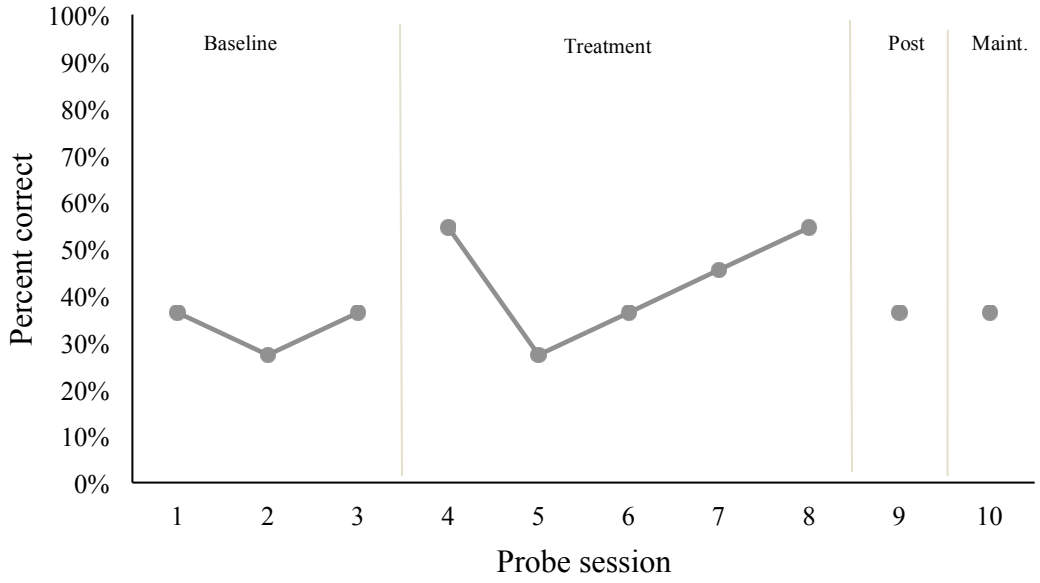
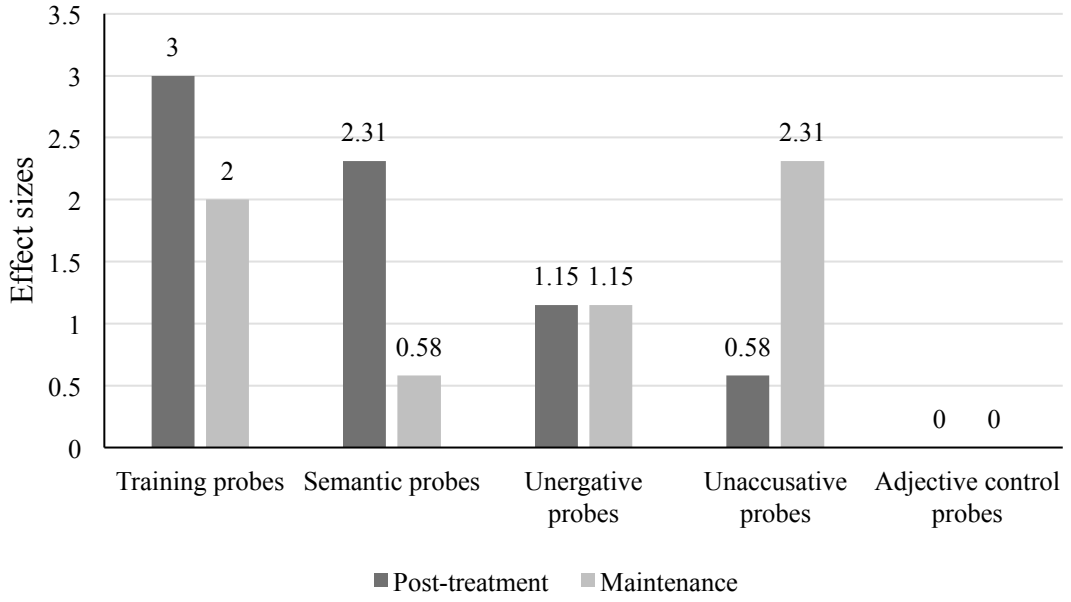


Figure 3. Effect sizes for each probe type at post-treatment and maintenance.



Pre- and Post-Treatment Testing

XH's performance on pre- and post-testing was compared. The results for tests administered in pre- and post-treatment are presented in Table 1 (see Chapter 2). There was no change in score on the CAT as measured by mean modality score, a composite score calculated by averaging t-scores from the eight language subtests of the total battery (following the procedures of Winans-Mitrik et al., 2014). On the Boston Naming Test, XH's score increased from 28/60 to 35/60, an increase of 11.6%. On the Northwestern Assessment of Verbs and Sentences, XH's single word naming of untrained and unprobed verbs improved from 72.2% to 83.3% (n=18), and argument structure production of 1-, 2- and 3-place untrained and unprobed verbs improved from 68.8% to 81.3% (n=16). Single verb naming of untrained 2-place verbs decreased slightly from 85.7% to 71.45% (n=7); however, argument structure production for 2-place verbs increased from 70.0% to 80.0% (n=5).

Measures of %CIUs in a picture description task (Nicholas & Brookshire, 1993) increased from 50.6% to 66.2%. This increase was largely due to more efficient speech production from pre- to post-treatment, with fewer revisions and reduced use of filler words than in pre-testing. Measures of verb type-token ratio remained stable at 50.0 due to the limited number of verbs produced in connected speech across both trials (n=4).

XH did not demonstrate change in measures of overall quality of life as measured by the SAQOL-39g, the Adaptive ACOM, or the Disability Quotient of the CAT, with pre- and post-testing scores on the SAQOL-39g and Disability Quotient of the CAT remaining stable while pre- and post-testing scores on the Adaptive ACOM fell within the standard error of measurement on both administrations.

Research Question #2

Unergative and Unaccusative Probes

Results for the unergative and unaccusative verb probes are depicted in Figure 1. Unergative probes improved from a baseline of 59.3% accuracy ($\overline{x}$=5.33 items) to a post-treatment accuracy of 66.7% (n=6 items) with a below-small effect size of 1.16; these effects were maintained at one month with a score of 66.7% accuracy (n=6 items) and effect size of 1.16. Finally, XH's production of unaccusative probes was analyzed. His performance increased slightly from a baseline average of 45.9% accuracy ($\overline{x}$=3.67 items) to 50.0% accuracy (n=4 items) at post-treatment with an effect size of .57, indicative of below-small effects. At maintenance, his score increased to 62.5% accuracy (n=5), with effect size of 2.31. See Figures 1 and 3.

CHAPTER 4

DISCUSSION

The purpose of the present study was to investigate the syntactic generalization properties of Verb Network Strengthening Treatment (VNeST), a verb-based treatment for naming deficits, and to replicate previous findings by Edmonds and colleagues on its efficacy in the treatment of verb and noun naming deficits for individuals with aphasia. The study addressed the following research questions: 1) Does training a set of verbs generalize to a set of untrained, semantically related transitive verbs and agent/patient pairs as well as generalized improvements on measures of lexical retrieval, connected speech, and quality of life? and 2) Does training a set of transitive verbs generalize to untrained, semantically unrelated unergative and unaccusative intransitive verbs?

To address Research Question 1, the present study provides support for previous findings from Edmonds and colleagues that VNeST results in generalization to untrained, semantically related transitive verbs and agent/patient pairs as well as generalized improvements in noun and verb retrieval on pre- and post-treatment measures including confrontation naming of nouns and verbs, sentence production, and connected speech (Edmonds, 2016; Edmonds & Babb, 2011; Edmonds, Mammino, & Ojeda, 2014; Edmonds, Nadeau, & Kiran, 2009). The original investigation by Edmonds et al. (2009) defined generalization during sentence probes as a 40% increase in probe production from the highest baseline at post-treatment. Because the semantically related probe set in the present study was 50% smaller than the 2009 investigation (n=5 compared to n=10), a 40% generalization criterion was not reached; XH's production of untrained,

semantically related verbs during probe elicitation tasks increased by only 26.6% at post-treatment. However, the generalized improvements in naming on standardized measures in the present study are comparable to those found in the original 2009 VNeST investigation. The present results show an 11.7% increase in noun retrieval on the BNT and 11.1% increase in verb naming on the NAVS, compared to respective gains of 13.2% and 8% in the original investigation using comparable measures (Edmonds et al., 2009). The present study also exceeds the average of 5.1% increase in noun naming and 8.6% increase in verb naming found in the group treatment conducted by Edmonds et al. (2014). While XH's 15.6% gain in production of correct information units (%CIUs) in connected speech fell just outside the range of 16.9% to 30.1% improvement reported by Edmonds and colleagues in 2009, the subsequent group investigation by Edmonds et al. (2014) did not find consistent increases in %CIUs among their participants. The authors attributed this variability to the relative complexity of connected speech tasks and in measures to evaluate discourse.

However, unlike the results found in previous investigations, the present study did not find generalized improvements to an overall measure of aphasia severity. All participants in Edmonds et al. (2009) and seven of 10 participants (including two of two participants with conduction aphasia) in Edmonds et al. (2014) showed clinically significant improvements on the Aphasia Quotient of the WAB. While XH showed slight increases and decreases in scores across several of the language subtests of the CAT, the mean modality score remained stable at pre- and post-treatment (see Chapter 2). It is possible that XH's lack of improvement on the CAT may be due to differences in psychometric properties between the CAT and the WAB, as the CAT is controlled for

psycholinguistic variables that may influence performance. Additionally, no changes were observed in measures of quality of life, with XH's scores on the SAQOL-39g and Adaptive ACOM remaining within the standard error of measurement at both administrations. One possibility is that XH's self-reported responses on these measures did not accurately reflect the communication difficulties he encountered in daily life, and that his high scores on these measures may have been due to his reticence to discuss his functional disability at both administrations.

The present findings therefore corroborate the claim that VNeST results in generalized improvement in lexical retrieval deficits on a variety of measures due to widespread activation of diverse semantic networks related to verbs and nouns targeted during treatment, and support the use of VNeST in the treatment of naming deficits in moderate conduction aphasia. These results also highlight the variability in performance on standardized measures among participants in single-subject designs.

To address Research Question 2, the present study replicates findings from previous studies that VNeST results in syntactic generalization to untrained intransitive verbs on standardized measures and additionally shows novel findings that may clarify to what degree different types of intransitive verbs may improve using sentence elicitation probes. XH's production of intransitive verbs in both probes and on the NAVS was higher at baseline than production of transitive verbs. This is not surprising given that intransitive verbs are known to be easier to produce because they require fewer arguments to be generated (Kim & Thompson, 2000; Shapiro & Levine, 1990; Shapiro, Zurif, & Grimshaw, 1997; Thompson, Lange, Schneider, & Shapiro, 1997). Nonetheless, treatment using transitive verbs in VNeST resulted in improved production of untrained

intransitive verbs on both probe measures and the Verb Naming subtest of the NAVS. The results from the present study at post-treatment and maintenance support the interpretation that treatment of two-place verbs will result in some degree of generalization to one-place verbs, and therefore adds to the body of literature supporting the complexity account of treatment efficacy (CATE; Thompson, Shapiro, Kiran, & Sobecks, 2003).

It was predicted that XH would demonstrate greater improvement to unergative than to unaccusative verbs in line with the Argument Structure Complexity Hypothesis (ASCH; Thompson, 2003). Because the ASCH proposes that unaccusative verbs have a non-canonical d-structure representation, it was hypothesized that they would be less susceptible to generalization from the canonical argument structure of transitive verbs. Therefore, a larger effect size was predicted for unergative than for unaccusative verbs. This pattern was observed at the post-treatment timepoint but not at maintenance. An informal error analysis revealed XH had resolved a perceptual misinterpretation of one of the unaccusative probes at maintenance, resulting in a higher score and therefore larger effect size than at post-treatment (see Limitations for more details). Additionally, both unergative and unaccusative verb subtypes showed effect sizes that were below the criteria for a small effect per the recommendations of Beeson and Robey (2006). These findings are therefore inconclusive; the present study cannot provide definitive support for the interpretation that the non-canonical d-structure representation of unaccusative verbs renders them less susceptible to generalization from the canonical argument structure of transitive verbs.

Limitations

While positive gains were observed in the present study across all probes at post-treatment and maintenance, a number of variables may have impacted the breadth of the present results and resulted in smaller effect sizes. XH had received prior VNeST clinical treatment and therefore may have had higher baseline scores for several of the training items and semantic generalization probes, which may have depressed his overall effects on the probe measure. Additionally, XH only received six weeks of treatment while subsequent VNeST studies conducted treatment for up to 15 weeks (c.f. Edmonds & Babb, 2011).

Limitations in training target selection may have also affected the breadth of the results. Because the targets for the present study were randomly selected from Edmonds' prior research, several targeted verbs such as *chop* and *boil* shared some semantic overlap (e.g. subjects like *chef* or wh-question responses such as *in the kitchen*). This may have resulted in less widespread semantic network activation than would have occurred with more structured verb selection, and may have similarly depressed overall effect sizes on the training and semantic generalization probes.

Similarly, weaknesses in probe selection may have affected XH's performance on the probe elicitation task each week. In particular, unaccusative verbs tend to be less imageable than other verbs, making them difficult to elicit via picture description. While the majority of the weekly sentence elicitation probe images were normed on a population of healthy adults, several of the unergative and unaccusative probe images were independently developed by the author and were only informally normed on two healthy volunteers. This may have depressed XH's overall scores on the weekly probe

measure for intransitive verbs as several of his responses reflected consistent misperceptions of the picture stimuli.

Finally, the effect sizes for all probe types should be interpreted with caution. At post-treatment, only training probes showed small gains using the benchmarks from Beeson and Robey (2006) selected for the present study. None of the remaining probes, including semantic, unergative, and unaccusative probes, demonstrated large enough effect sizes to qualify as small improvement. This may have been the result of several factors. XH's performance on production of intransitive verbs was higher than transitive verbs at baseline, which may have limited his gains at post-treatment. Additionally, the small number of probe items in each set (n=5 training probes, n=5 semantic probes, n=9 unergative probes, and n=8 unaccusative probes) makes the gains in post-treatment and at maintenance difficult to interpret within each category, as even small variations resulted in considerable change to effect size; for example, training probes decreased in accuracy by only one item from post-treatment to maintenance, and the corresponding effect dropped from small to below-small. Because the verb naming measure used in the present study did not test for naming or production of unaccusative verbs, pre- and post-treatment scores using a standardized measure could not be compared to verify the results of the weekly sentence probes for unergatives and unaccusatives. However, XH demonstrated an overall 25% increase in naming of untrained, intransitive verbs on the NAVS, corroborating the probe findings that generalization did occur from transitive to intransitive verbs.

Clinical Implications

The results of the present study offer a number of implications for clinicians working with individuals with acquired aphasia. Based on the breadth of generalization observed on pre- and post-treatment measures, these results support treating lexical retrieval deficits using protocols such as VNeST that target the sentence level rather than single word naming or in a hierarchical progression from simpler to more complex forms. While connected speech and confrontation naming of nouns were not explicit targets of treatment, XH made gains in these areas likely due to the widespread activation of distributed semantic networks in VNeST. Additionally, novel results from the present study support utilizing the principles from the complexity account of treatment efficacy (Thompson et al., 2003): the more complex linguistic form should be targeted. Transitive and ditransitive verbs should be selected as targets in the treatment of sentence production deficits, and clinicians can expect the use of such targets to result in some systematic generalization to intransitive verbs. Based on the results of the present study, future investigations should seek to determine the mechanism by which implicit syntactic generalization occurs in verb argument structure.

REFERENCES

Balota, D.A., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

Barbieri, E., Luzzatti, C., Den Ouden, D.-B., Aggujaro, S., Verga, R., & Thompson, C. K. (2011). An investigation of the Argument Structure Complexity Hypothesis (ASCH) in English and Italian speakers: data from aphasic and healthy participants. *Structuring the Argument: A multidisciplinary workshop on the mental representation of verbal argument structure.* Paris, September 5-6.

Bastiaanse, R. & van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics, 18,* 57-66.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review, 16*(4), 161-169.

Berndt, R.S., Mitchum, C.C., Haendiges, A.N., & Sandson, J. (1997). Verb retrieval in aphasia: 1. Characterizing single word impairments. *Brain and Language, 56*(1), 68-106.

Black, M. & Chiat, S. (2008). Interfaces between cognition, semantics, and syntax. In Ball, M.J., Perkins, M.R., Muller, N., and Howard, S. (Eds), *The handbook of clinical linguistics*. Malden, MA: Blackwell Publishing Ltd, 290-307.

Coltheart, M. (2005). Modeling reading: the dual-route approach. In Snowling, M.J. & Hulme, C. (Eds.), *The science of reading: a handbook* (pp. 6-23). Malden: Blackwell Publishing.

Conti-Ramsden, G., & Jones, M. (1997). Verb use in Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 40,* 1298-1313.

Dabul, B. (2000). *Apraxia battery for adults - second edition.* Austin, TX: Pro-Ed.

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

Edmonds, L.A. (2014). Tutorial for Verb Network Strengthening Treatment (VNeST): Detailed description of the treatment protocol with corresponding theoretical rationale. *SIG 2 Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders, 24,* 78-88.

Edmonds, L.A. (2016). A review of Verb Network Strengthening Treatment: Theory, methods, results, and clinical implications. *Topics in Language Disorders, 36*(2), 123-135.

Edmonds, L.A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology, 20*, 131-145.

Edmonds, L.A., Mammino, K., & Ojeda, J. (2014). Effect of Verb Network Strengthening Treatment (VNeST) in persons with aphasia: extension and replication of previous findings. *American Journal of Speech Language Pathology, 23*(2), S312-329.

Edmonds, L.A., Nadeau, S.E., & Kiran, S. (2009). Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology, 23*(3), 402-424.

Ferretti, T.R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*(4), 516-547.

Friedmann, N., Taranto, G. Shapiro, L.P., & Swinney, D. (2008). The left fell (the leaf): The online processing of unaccusatives. *Linguistic Inquiry, 39*(3), 355-377.

Furnas, D.W., & Edmonds, L.A. (2014). The effect of Computerised Verb Network Strengthening Treatment on lexical retrieval in aphasia. *Aphasiology, 28*(4), 401-420.

Gierut, J.A. (2001). Complexity in phonological treatment: Clinical factors. *Language, Speech, and Hearing Services in Schools, 32*, 229-241.

Helm-Estabrooks, N., Nicholas, M., & Helm, S.A. (2000). *Sentence production program for aphasia.* Pro-ed.

Hilari, K., Byng, S., Lamping, D.L., & Smith, S.C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability, and validity. *Stroke, 34,* 1944-1950.

Hula, W.D., Doyle, P.J., Stone, C.A., Austermann Hula, S.N., Kellough, S., Wambaugh, J.L., Ross, K.B., Schumacher, J.G., & St. Jacque, A. (2015). The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and internal validation. *Journal of Speech, Language, and Hearing Research, 58*, 906-919.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston naming test.* Pro-ed.

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

Kertesz, A. (1982). *Western aphasia battery.* Austin, TX: Pro-Ed.

Kim, M. & Thompson, C.K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language 74*(1), 1-25.

Kiran, S. (2007). Complexity in the treatment of naming deficits. *American Journal of Speech-Language Pathology*, *16*(1), 18-29.

Kiran, S. & Thompson, C. (2003). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research, 46*(3), 608-622.

Kwag, E.J., Sun, J.E., Kim, Y., & Cheon, H. (2014). Effects of Verb Network Strengthening Treatment on retrieval of verbs and nouns in persons with aphasia. *Communication Sciences and Disorders, 19*(1), 89-98.

Linebarger, M.C., Schwartz, M.F., & Saffran, E.M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*(3), 361-392.

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The Communication Effectiveness Index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech Language, and Hearing Research, 46*, 113-124.

Loverso, F.L., Prescott, T.E. and Selinger, M. (1988). Cueing verbs: a treatment strategy for aphasic adults (CVT). *Journal of Rehabilitation Research*, 25(2), 47-60.

Maas, E., Robin, D.A., Austermann Hula, S.N., Freedman, S.K., Wulf, G., Ballard, K.J., & Schmidt, R.A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*(3), 277-298.

McAllister, T., Bachrach, A., Waters, G., Michaud, J. & Caplan, D. (2009). Production and comprehension of unaccusatives in aphasia. *Aphasiology, 23*(7-8), 989-1004.

McRae, K., Hare, M., Elman, J.L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33*(7), 1174-1184.

Meltzer-Asscher, A., Mack, J.E., Barbieri, E., & Thompson, C.K. (2015). How the brain processes different dimensions of argument structure complexity: evidence from fMRI. *Brain and Language, 142*, 65-75.

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

Shapiro, L.P. & Levine, B.A. (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language, 38*(1), 21-47.

Shapiro, L.P., Zurif, E., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition, 27*(3), 219-246.

Swinburn, K., Porter, G., & Howard, D. (2005). *The Comprehensive Aphasia Test*. Hove, UK: Psychology Press.

Thompson, C.K. (2002). *Northwestern verb production battery*. Unpublished.

Thompson, C.K. (2003). Unaccusative verb production in agrammatic aphasia: The argument structure complexity hypothesis. *Journal of Neurolinguistics*, *16*(2-3), 151-167.

Thompson, C.K. (2007). Complexity in language learning and treatment. *American Journal of Speech-Language Pathology, 16*, 3-5.

Thompson, C.K. (2011). *Northwestern assessment of verbs and sentences.* Evanston, IL.

Thompson, C.K., Lange, K.L., Schneider, S.L., & Shapiro, L.P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4-5), 473-490.

Thompson, C.K. & Shapiro, L.P. (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech Language Pathology, 16*(1), 30-42.

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

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6-10.

Winans-Mitrik, R.L., Hula, W.D., Dickey, M.W., Schumacher, J.G., Swoyer, B., & Doyle, P.J. (2014). Description of an intensive residential aphasia treatment program: rationale, clinical processes, and outcomes. *American Journal of Speech-Language Pathology, 23*, S330-342.

Zingeser, L.B., & Berndt, R.S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language. 39*, 14–32.