



Published in final edited form as:

*Am J Speech Lang Pathol.* 2012 May ; 21(2): S88–S102. doi:10.1044/1058-0360(2012/11-0106).

## Development of a Theoretically Based Treatment for Sentence Comprehension Deficits in Aphasia

Swathi Kiran<sup>1</sup>, David Caplan<sup>1,2</sup>, Chaleece Sandberg<sup>1</sup>, Joshua Levy<sup>1,2</sup>, Alex Berardino<sup>1</sup>, Elsa Ascenso<sup>1</sup>, Sarah Villard<sup>1</sup>, and Yorghos Tripodis<sup>3</sup>

<sup>1</sup>Speech and Hearing Sciences, Boston University, Boston, MA

<sup>2</sup>Neuropsychology Laboratory, Massachusetts General Hospital, Boston, MA

<sup>3</sup>School of Public Health, Boston University, Boston, MA

### Abstract

**Purpose**—We piloted two new treatments, one based on sentence to picture matching (SPM) and the other based on object manipulation (OM), that train participants on the thematic roles of sentences using pictures or by manipulating objects.

**Methods**—Using a single subject multiple baseline design, sentence comprehension was trained on the affected sentence type in one task-related protocol in 15 participants with aphasia. The two tasks were SPM and OM and the treatment stimuli were Object relatives; Object clefts; Passives and Unaccusatives as well as two control structures; Object relatives with a complex NP, active sentences with three NPs.

**Results**—The criteria for efficacious treatment was an increase in the level of performance from the pre-treatment probes to post-treatment probes for the treated structure such that accuracy rose from at or below chance to above chance and either a) accuracy on rose by 33% or b) the effect size was 2.6. Based on these criteria, the success rate for training the target structure was 4/7 participants in the OM condition and 2/6 participants in the SPM condition.

**Conclusions**—The outcome of this study illustrates the utility of this theoretically-motivated and efficacious treatment for sentence comprehension deficits in participants with aphasia.

### Introduction

The goal of this paper is to describe the development of a clinically applicable treatment for disorders affecting the ability to use syntactic structure to understand sentences. Our project builds on work that characterizes the nature of sentence comprehension disorder and work that explores the factors that influence the effects of treatment. Many researchers maintain that syntactic comprehension deficits result from loss of knowledge regarding certain syntactic constituents or from loss of the ability to use that knowledge in comprehension ('specific deficit' models). One example of this type of theory is the Trace Deletion Hypothesis (Grodzinsky, 2000), which relies on Chomsky's theory of syntactic structure (Chomsky, 1995). The Trace Deletion Hypothesis postulates that certain participants with aphasia have lost the ability to relate certain types of these moved constituents to their sites of origin through the co-indexation of traces, resulting in the inability to assign thematic roles to these constituents. Trace deletion leads to abnormal sentential comprehension in

instances where compensatory mechanisms of thematic role assignment (i.e. 'heuristic' mechanisms) also fail to accurately assign thematic roles.

Other descriptions of aphasic comprehension deficits also are of the specific deficit type. For instance, several alternative characterizations of a deficit affecting traces have been proposed (Beretta, 2001; Beretta, Harford, Patterson, & Piñango, 1996; Maunder, 1995; Maunder, Fromkin, & Cornell, 1993). The mapping deficit account (e.g., Linebarger, 1995) proposes a specific deficit in the syntactic comprehension process that is wider-ranging than the trace deletion account and argues that participants with aphasia assign syntactic structures but have lost the ability to map them onto sentential semantic meanings.

A second view of syntactic comprehension deficits states that participants display reductions in the resources (e.g., working memory) necessary for the normal operation of processes associated with syntactic comprehension. These models maintain that participants with syntactic comprehension problems do not lose specific syntactic constituents or parsing/interpretive operations but rather have pathological limitations in their abilities to undertake the computations that assign and interpret syntactic structure (a resource limitation). These limitations could be due to a reduced capacity to maintain representations in memory while computations are applied to them (i.e., a working memory deficit: Caplan & Waters, 1999), a slow rise time or fast decay rate for representations, which makes them unavailable when needed (i.e., a speed of processing deficit: Haarman & Kolk, 1994), or some other impairment that affects syntactic processing across the board (Tseng, McNeil, & Milenkovic, 1993). Evidence for a capacity reduction comes from the finding that sentences that are more complex linguistically and require more processing are more affected than sentences that require less syntactic (Caplan & Futter, 1986; Caplan & Hildebrandt, 1988 a,b; Caplan, Baker & Dehaut, 1985; Caplan, Waters & Hildebrandt, 1997).

Specific deficits are rare. Caplan and colleagues obtained multiple measures (accuracy, RTs, on-line data) regarding participants' performances on two constructions instantiating each of three syntactic operations (WH-movement, NP-movement, co-indexation of reflexives) across different tasks including sentence comprehension and grammaticality judgment in both sentence to picture matching and object manipulation modalities (Caplan, Waters, DeDe, Michaud, & Reddy, 2007). Results showed that every participant with a task-independent deficit affecting one structure also had abnormal performances affecting other structures. Further, sentences with non-canonical thematic role order were more difficult than sentences with canonical thematic role order, and OM tasks were more difficult than SPM tasks. These results point to resource reduction as the major cause of syntactic comprehension disorders and suggest that OM is a more demanding task than SPM (Caplan & Waters, 2006; Caplan, DeDe & Michaud, 2006, Caplan et al., 2007; DeDe & Caplan, 2006). We should note that other studies that have also implemented tasks similar to sentence to picture matching and enactment and have also observed task specific effects in comprehension of canonical and noncanonical sentences (Salis & Edwards, 2009; Cupples & Inglis, 1993).

Each of the above mentioned theoretical viewpoints are relevant to therapy as they generate suggestions regarding the structures that need to be targeted in training. Theories that argue for a specific deficit in syntactic comprehension may suggest that improvements may be restricted to trained structures, and consequently participants will have to be trained on a large number of structures to achieve overall improvements in comprehension. The resource view of syntactic comprehension deficits is more optimistic about generalization of successful training from one structure to another; if training targets an ability to utilize resources in the service of syntactic comprehension, successful training might generalize to all structures, or to all structures that require the same or fewer resources.

Few treatment studies have specifically targeted sentence comprehension as a goal of rehabilitation. Mitchum, Greenwald and Berndt (2000) divided studies into three broad categories --verb centered treatments, active feedback treatments and prepositional/locative treatments. They reviewed data from 17 participants across these studies and reported variable outcomes across different participants. In a series of studies that focused on syntactic structure, Berndt, Mitchum, Haediges and colleagues trained active and passive reversible sentences, emphasizing the association between the surface sentence structure and the meaning of the sentence (Berndt, Mitchum, Burton, & Haediges, 2004; Haediges, Berndt, & Mitchum, 1996; Mitchum, Haediges, & Berndt, 1995). Across these studies, participants demonstrated improvements on trained items, but continued to have persisting deficits in other aspects of sentence comprehension and generalization effects were variable. One group study (Crerar, Ellis, & Dean, 1996) examined treatment of sentence comprehension using a computer-based written sentence-to-picture matching task. Fourteen aphasic participants, divided into two equal groups, received either verb comprehension followed by preposition comprehension training or the reverse order of therapy. Both groups showed improved comprehension, and the group receiving verb comprehension therapy first demonstrated generalization to preposition comprehension.

Schwartz and colleagues have argued that therapy aimed at the mapping of a sentence's thematic roles onto its syntactic structure achieved better results than traditional repetition-based therapy (Schwartz et al., 1994). In this study, eight participants with agrammatism received therapy that focused on thematic role assignment on sentence structures that ranged from simple actives to more complex structures and found improvements in several participants. Several similar approaches have been reported in which treatment targeted the mapping of thematic roles onto a variety of structures in either comprehension or, more frequently, production. For instance, some studies have targeted thematic role mapping by manipulating the word order of sentences using either noun phrases or prepositions (Byng, 1988; Nickels, Byng, & Black, 1991). Across these studies, participants showed variable sentence-specific improvements and generalization to untrained structures, with some notable successes (e.g., Byng, 1988) and other outcomes that were much less favorable.

There is a more developed literature pertaining to the combination of training with respect to sentence production and with sentence comprehension. Several trends seem to emerge from this literature. Pertinent to the present study, Thompson, Shapiro and their colleagues (Jacobs & Thompson, 2000; Thompson, Ballard, & Shapiro, 1998; Thompson & Shapiro, 2005; Thompson, Shapiro, Kiran, & Sobecks, 2003) developed the 'treatment of underlying forms', an approach that is similar to mapping therapy in that participants are taught aspects of the relation between thematic roles and syntactic structure, focusing on the syntactic realization of verbal argument structure in more complex sentences. One study using this approach examined the effect training comprehension on production and vice-versa using object clefts (OC) and passives (PA) and found that training comprehension improved production whereas training production was less effective at improving comprehension (Jacobs & Thompson, 2000).

A second feature of these studies that is relevant to the present study is the Complexity Account of Treatment Efficacy (CATE) (Thompson & Shapiro, 2007; Thompson, et al., 2003). The CATE maintains that successful training on more complex but similar structures leads to generalization to less complex structures but not vice versa. For instance, Thompson et al. (1998) reported that successful training in production of cleft object structures generalized to WH-questions, but not vice versa. Similarly, Thompson et al. (2003) extended these results to object relative clauses (OR), which are more complex than object clefts (OC); the same pattern of generalization only from complex to simple WH-movement structures was found. Thompson et al (1997) reported a CATE effect in training NP-

movement, with generalization from successfully trained “raising” structures to passives (which are argued to be less complex). Jacobs and Thompson (2000) and Thompson et al (1997, 1998) reported that successful training on either WH-movement or NP-movement did not generalize to the other type of structure.

Taken together, mapping treatment, TUF treatment and other syntactically constrained treatment studies provide an important foundation for developing a sentence comprehension treatment for participants with aphasia, but new therapeutic techniques are important to explore. Importantly, many therapies that have examined comprehension outcomes require a combination of production and comprehension, making it difficult to determine the extent to which they would be effective for sentence comprehension regardless of the production capacity in these participants.

In this project, we piloted two treatments (one based on sentence-to-picture matching (SPM) and the other based on object manipulation (OM)) and systematically manipulated the syntactic structures of sentences trained, to train participants on the relation between features of syntactic structure and the meanings of sentences by having participants indicate thematic roles pictured in drawings or by manipulating objects. Although the treatment methods reviewed above are directed at relating semantic values to syntactic structure, the methods consist of having participants arrange written words to indicate thematic roles or syntactic categories. The goal of our project was to develop a single therapy approach that could be implemented in two tasks (OM & SPM) and therefore be able to compare the enactment-based therapy (OM) with one that required identification of thematic roles using pictures (SPM). We hypothesized that displaying thematic roles while manipulating objects will be a more engaging and effective therapy approach than identifying thematic roles using picture stimuli. This is because demonstrating accurate sentence comprehension during the OM task requires the participant to develop a mental visual schema of the action and subsequently generate the motor plan to implement the enactment. In contrast, in SPM, the participant is required to recognize a visual schema of the action depicted and match that image to one of two picture options. In general, retrieval of items from memory is more difficult than recognition, leading to the assumption that OM is more demanding than SPM.

We also systematically manipulated the treatment dosage to understand the stimulus and treatment parameters that are optimal for promoting improvements in participants with comprehension deficits. Previous treatment studies that have targeted sentence comprehension have provided therapy at various frequencies, but frequency of treatment has been confounded with intensity of treatment and in some cases type of treatment. For instance, Byng’s patient was seen for treatment only for 2 weeks but showed generalization both to untrained items and from comprehension to production (Byng, 1988); whereas Schwartz et al., (1994) trained their participants three times a week for 60–90 minutes up to four months, and reported a variety of outcomes in different participants. Jacobs and Thompson (2000) trained their participants two-to-three times per week for 60–90 minutes for approximately 10 weeks. Mitchum et al. (1995) also trained participants for extensive periods, and their participants did not show generalization. In this study, we examined whether a more or less treatment dosage at the same frequency (twice a week) was more effective in determining treatment outcomes.

Finally, we also examined generalization from the trained to untrained sentences within the trained task. Our choice of stimuli was motivated by previous work that has shown OR structures to be more difficult to understand relative to OC structures and passives were found to be more difficult than unaccusatives (UNACC; e.g., Caplan et al., 1985, 2006, 2007; MacAllister et al., 2009; Salis & Edwards, 2009). We hypothesized that if structure is not a barrier to generalization, because participants have resource reductions serving

sentence comprehension and not specific deficits, generalization may be expected from successful training of both WH-movement structures to both NP-movement structures and vice versa. Alternatively, based on the CATE hypothesis, generalization may occur within WH- and NP-movement structures but not across structures, and only from the complex structures to the less complex structures (e.g., OR->OC and PA->UNACC).

## Methods

### Participants

Fifteen monolingual English-speaking individuals with aphasia were recruited from local hospitals within the Boston area to participate in the study. Participant ages ranged from 29 to 73 years (mean = 57 years). Time post-onset ranged from 6 to 147 months (mean = 55 months). Several initial selection criteria were met, including (a) a single left hemisphere stroke, (b) onset of stroke at least six months prior to participation in the study, (c) premorbid right-handedness, and (d) at least a high school diploma (see Table 1). There was one exception, BUMA05 was pre- pre-morbidly left-handed and had a right hemisphere stroke. All participants had received varying amounts of traditional language treatment during the initial months following their stroke.

Screening measures were administered prior to treatment to determine eligibility. The first of these assessed participants' auditory comprehension of both the nouns and verbs included in the treatment and monitoring stimuli (see section on stimuli for details). Participants were shown two depictions of either nouns or verbs and asked to choose the picture that matched the auditorily-presented target; comprehension of nouns and verbs was assessed separately. Accuracy scores on the noun portion of this screener ranged from 86.9% to 100% correct (mean = 95.1%) and from 68.8% to 100% correct (mean = 88.3%) on the verb portion and these scores were considered to be within the range of acceptability for eligibility for the study. Additionally, both an SPM and an OM sentence comprehension screener was administered to assess each participant's comprehension level for both NP-movement and WH-movement sentences, as well as other structures. Accuracy scores on the SPM screener ranged from 42.7% to 85.5% correct (mean = 63.2%), and accuracy scores on the OM screener ranged from 20.0% to 90.0% correct (mean = 47.5%)(see Table 3). Eligibility for entrance into the study was performance that was at or below chance on any of the target sentence types.

### Stimuli

All stimuli used in the study were comprised of reversible sentences that were implemented into a screening battery, a monitoring battery and a treatment set. All stimuli were drawn by the same artist, recorded by the same male voice at a consistent pace, and all items were presented using E-prime software. Half of all sentences for each sentence type used animate nouns while the remaining half used inanimate nouns. The animacy of nouns and the lexical frequency of nouns and verbs were balanced across batteries and across versions of each battery.

For the SPM tasks, a target picture depicting the actions in the sentence and a foil picture depicting the same items and actions with reversed thematic roles were drawn for each sentence. The frequency of the position of the target picture on either side of the screen was counterbalanced across trials. For the OM tasks, paper dolls representing the nouns in the sentence were created for use by the participants. The dolls were laid in front of the patient in random order to control for order effects of agent and theme within and across sentence types.

**Screeners**—Two versions of each type of screener (SPM, OM) were developed that included both the target sentence types (PA, UNACC, OR, OC), two control sentence types (sentential complement (3NP) and object relative with embedded subject modified by a prepositional phrase (ORCNP)), and a variety of other sentence types (active, reflexive, pronoun-as-object, object control, subject control, and noun phrase raising). Ten tokens of each sentence type were presented, except for the control sentences which had five tokens each, for a total of 110 items per screener version (see Appendix A for examples of these sentences). The same stimuli were used for each version of the screener, but the order of presentation was reversed between versions one and two.

**Monitoring batteries**—Four versions of each type of monitoring battery (SPM, OM) were developed that included 15 tokens of each target sentence type (OC, OR, PA) except unaccusatives (10 Tokens) and 10 tokens of each control sentence type (3NP and ORCNP) for a total of 75 items per version. The 3NP sentence structures were included to account for the length of OR sentences, whereas ORCNP structures were included as a measure of added complexity. The thematic roles assigned to particular nouns within the stimuli were counterbalanced across the four versions such that thematic roles of nouns in version 2 & 4 were reverse of thematic roles of the same nouns in version 1 & 3. Versions 1 and 2 were differentiated from versions 3 and 4 by placing the target item on a different side of the screen for each sentence. For the OM task, versions 1 and 2 had the same sentences as SPM versions 1 and 2. Version 3 contained new sentences with the same verbs and version 4 contained the same nouns/verbs as version 3, but the roles of agent/theme were reversed.

**Treatment materials**—Twenty sentences for each sentence type, except unaccusatives (UNACC), which only had 14 sentences, were used in training for both the SPM and the OM training methods. These sentences were mostly original, but less than 1/4 were copied from the screeners and one (for each sentence type) was copied from monitoring batteries for all sentence types except UNACC. For the UNACC training sentences, 9 of the 14 sentences were taken from the screeners and 2 were taken from the monitoring batteries. This was unavoidable due to the scarcity of English unaccusative verbs and because the same verb must be able to be used both unaccusatively and transitively (to allow plausible foils). Since participants were only trained on one sentence type in one modality (SPM or OM), the same sentences for each sentence type were used across each training modality.

## Experimental Design

A single subject multiple baseline design (Thompson, 2006) was employed to examine treatment effects. Following three baseline sessions, for each participant, one treatment task was used to train comprehension of one affected sentence structure. The treatment protocol utilized either an SPM or an OM task and the treated sentence type was one of the following: OR, OC, PA or UNACC. In addition to the target structures, two untrained control structures (ORCNP and 3NP) were included in the monitoring batteries. Treatment was discontinued when the participant exceeded 80% accuracy on three consecutive monitoring batteries or after 10 weeks of treatment, whichever came first. Versions of both the SPM and OM Monitoring Batteries were administered weekly during the treatment and three additional times after treatment to establish the effect size of the treatment. Two training session schedules were used. The first consisted of 2 one-hour sessions per week. The second consisted of 2 two-hour sessions per week.

The criteria for efficacious treatment was deemed to be an increase in the level of performance from the pre-treatment baseline probes to post-treatment probes for the treated structure such that accuracy rises from at or below chance to above chance and either a) accuracy on the treated structure rises by 33% or b) the effect size of training on the treated

structure is 2.6. These criteria were selected for the following reasons. First, participants were required to improve from below/at chance to above chance performance rather than normal performance because the 95% confidence interval for normal performance includes chance for OR sentences (Caplan et al, 2006), and we believe that performance must be above chance to demonstrate that participants can reliably understand target sentences. The other two criteria were included because it would be possible for a patient to meet criterion by increasing the number of correct responses by one sentence, which we believe is too small an improvement. The level of 33% improvement was selected so that a patient performing at the upper range of chance in the pre-test period would have to improve by four items – an improvement that seems possible and meaningful. Finally, with regards to the criteria for effect sizes, existing data suggest that the effect size seen in aphasia treatment studies differs depending upon the language process under study. Based on an overall review of research on treatment studies in aphasia (Robey, Schutlz, Crawford, & Skinner, 1999), Beeson and Robey (2006) have suggested that effect sizes of 2.6, 2.9, and 5.8 be used as initial benchmarks for small, medium and large respectively for single subject treatment studies in aphasia. Thus, in this study we considered treatments that result in an effect size of at least 2.6 to be successful as a first step in understanding the effectiveness of a treatment for sentence comprehension.

The average baseline probe scores were subtracted from the average post-treatment scores and the result was divided by the standard deviation of the baseline scores to calculate effect sizes for each patient. Chance performance for the SPM and OM task on each structure was calculated using the binomial test. For SPM there was a 50% chance of selecting the correct picture for all sentence structures; whereas for the OM task, all structures except OR and ORCNP also had a 50% chance of enacting the thematic roles accurately. For OR and ORCNP structures, there was a 33% chance of enacting the accurate thematic roles.

### **Treatment protocol**

The therapy used an approach that explicitly demonstrates the thematic role of each constituent of the sentence. Both the SPM and OM treatments were similar in terms of the number of steps as well as the basic procedures involved, but differed in terms of the nature of thematic role mapping. See Appendix B for detailed protocol descriptions. In the SPM task, the clinician used pictures to demonstrate the thematic roles in the trained sentence to the patient whereas in the OM task, the participant enacted the thematic roles using paper dolls. The protocols for OR, OC, PA and UNACC were similar in terms of the basic procedures, but differed in terms of the number of steps to facilitate sentence comprehension.

### **Reliability**

All of the baseline, training and monitoring sessions were digitally video recorded. Responses were scored by the treating clinician and 30% of these sessions were verified by a student trained to code each training step for adherence to the protocol and to code participants' enactment of thematic roles as depicting the target meaning. In terms of the reliability on the dependent variable, responses to SPM tasks were recorded on E-prime while responses to OM were recorded by the clinician. Object manipulation has proven to be a highly feasible means of having participants demonstrate comprehension, with hundreds of participants having been tested using this method and inter-observer reliability of over 90% for interpretation of depicted thematic roles (Caplan et al, 1985). Initial disagreements among the primary and secondary coder were discussed and either resolved or the trial eliminated. The overall inter-observer reliability on the dependent variable was 99%. Reliability on the independent variable (i.e., administration of treatment) was 89% between the two scorers.

## Results

We trained 15 participants, two on each of the target structures. Only one patient entered into the OM-UNACC treatment condition because several other candidate participants improved on this structure during baseline testing. Two participants dropped out of the study before completion of data collection in the SPM phase; hence, we cannot report efficacy data on these two participants. Of the remaining 13 participants, six participants met all three criteria set for treatment (see Table 2 for a summary of results). It should be noted that of the seven participants assigned to the OM treatment, four (BUMA10, BUMA20, BUMA11, BUMA24) met the stringent treatment criteria and two met partial criteria (BUMA21, BUMA05). BUMA14 showed no response to treatment at all; there was a decline in performance as a function of treatment. Results were mixed for participants receiving the SPM treatment. Two participants (BUMA16, BUMA23) met both the criteria; one (BUMA03) met one of two criteria and three patients (BUMA08; BUMA13; BUMA15) showed no improvement. Both BUMA15 and BUMA03 were already above chance before treatment (although both patients made gains in treatment). Two patients (BUMA08; BUMA13; BUMA15) showed no change as a function of therapy, and closer inspection of these patient's time-series data revealed great variability on a session-by-session basis. Recall that we selected two control structures, 3NP and ORCNP, in order to establish experimental control in the study. It was reasoned that ORCNP would not change as a function of treatment as they were significantly more complex than any other sentence type. While 3NP sentence structures were included to control for the length of the OR sentence structures; these sentences were canonical structures and thus could change as a function of treatment, especially if the treatment was provided to the more complex OR sentence structures. Results from Table 2 show that of the 13 patients that completed treatment, three participants showed changes on the ORCNP structures based on our criteria set for improvement on trained items. BUMA10 and BUMA03 were trained on OR structures and BUMA20 was trained on OM-OC. Two other participants showed changes on 3NP based on our criteria (BUMA24 and BUMA08); both participants were trained on UNACC. Neither of these two patients changed on ORCNP. Isolated improvements in the control structures in some patients do not pose problems for experimental control because for each patient, there were several untrained structures that showed no change in the face of improvements for the trained structure (see Table 2). As an example, patients trained on OR did not improve on PA (BUMA 10; BUMA21, BUMA03) and patients trained on PA did not improve on OR or ORCNP (BUMA11, BUMA16).

Several statistical tests were performed on the data and are organized according to (a) statistical tests done on the monitoring probes that examine differences between the two treatment types and the two treatment dosages and (b) statistical tests done on the screener probe data that examine differences between the two treatment types. Non parametric statistics were performed due to the relatively small and unequal sample sizes in several analyses.

A Wilcoxon test examining average percent change on the trained structure of the monitoring probes found a significant effect ( $z = 2.90$ ,  $p = 0.0037$ ) across participants, indicating the overall positive effect of treatment. We then separated the patient data by trained task (SPM/OM); a Mann-Whitney U-test examining the effect of treatment type (SPM/OM) revealed no significant effect on treatment effect size ( $U = 15.0$ ,  $p = 0.432$ ; Means: SPM = 2.84, OM = 5.97). Another Mann-Whitney U-test examining the effect of treatment type (SPM/OM) revealed no significant effect on change in % accuracy ( $U = 16.0$ ,  $p = 0.520$ ; Means: SPM = 0.252, OM = 0.347). While there were numerical differences between patients who received 2 hours of treatment each week (Mean ES = 2.9; Mean percent change = 20%) and four hours of treatment each week (Mean ES = 5.89; Mean



percent change = 36%) the difference was not statistically significant for effect size ( $U = 19.0$ ,  $p = 0.830$ ) or for average percent change ( $U = 16.0$ ,  $p = 0.520$ ) (see Table 2).

Next, the pre-post screener data was examined (see Table 3). A Wilcoxon test on improvement on all of the twelve structures within the trained task, pre to post screener, across patients, revealed a significant effect of treatment ( $z = 3.57$ ,  $p = 0.0003$ ). Next, a Wilcoxon test on improvement on the trained structure within the trained task, pre to post screener, across patients, revealed a significant effect of treatment ( $z = 2.48$ ,  $p = 0.013$ ). A Wilcoxon test on improvement on improvement on 3NP within the trained task, pre to post screener, across patients, revealed no significant effect of treatment ( $z = 0.652$ ,  $p = 0.515$ ). Finally a Wilcoxon test on improvement on improvement on ORCNP within the trained task, pre to post screener, across patients, revealed no significant effect ( $z = 1.47$ ,  $p = 0.142$ ). In addition, we performed Wilcoxon tests to examine the effect of each treatment on changes on the post-treatment screeners. Patients who received SPM treatment ( $n=6$ ) improved on their trained SPM structure ( $z = 2.20$ ,  $p = .028$ ) whereas OM patients ( $n=7$ ) did not improve significantly on their trained OM structure ( $z = 1.35$ ,  $p = .18$ ).

Given that some untrained structures showed changes as a function of treatment (see Table 2), we conducted an exploratory cross-correlation analyses using the **autoregressive integrated moving average** (ARIMA) procedure for each patient to examine if improvements on the untrained structure were associated with improvements in the trained structure/task. ARIMA cross-correlation models are calculated for time-series data as follows. For each time series, a regression line is fit to the actual data and the residuals are calculated for that data. Then cross-correlations are calculated on the residuals and averaged over time (Box, Jenkins & Reinsel, 1994). Therefore a correlation at a negative lag would indicate that changes in the second structure precede changes in the first structure; a correlation at 0 lag would indicate changes are concurrent in both structures and a correlation at a positive lag would indicate changes in the first structure precede changes in the second structure. In this study, for each patient, we correlated the time series between all possible pairs of structures at 10 lag points ( $-5$  to  $5$ ). Correlations that exceeded  $.60$  and exceeded two standard errors were deemed statistically significant. In the interest of space, we only report significant correlations (either positive or negative) between the trained structure and specific untrained structures and the corresponding effect sizes for the untrained structures for each patient. For a complete list of all significant correlations please see submitted supplemental data.

When combined with data presented in Table 2, Table 4 reveals several important observations. First, participants such as BUMA14, BUMA07, BUMA13 and BUMA08 who either dropped out or did not show improvements on the trained structure did not show any generalization to untrained structures. Several other participants showed positive and negative changes on the untrained structures that correspond to the data regarding effect size, percent accuracy and change to above chance levels presented in Table 2. Specifically, BUMA10 was trained on OR and showed concurrent changes in ES and percent accuracy on ORCNP and in percent accuracy for UNACC (correlation coefficient was  $.58$ ). BUMA21 was also trained on OR but did not show generalization to untrained structures, because several of these structures were above chance accuracy prior to treatment. BUMA05 was trained on OC but no other structures showed changes. BUMA20 was trained on OC; positive changes in ES, chance level accuracy and percent accuracy on OR and ORCNP coincided with improvements on OC. BUMA11 was trained on PA, slight improvements ES, percent accuracy and chance levels were observed in OC that preceded changes on the trained structure. In contrast, ORCNP and 3NP declined relative to positive changes on PA. BUMA24 was trained on UNACC, changes in ES and percent accuracy were observed on

3NP that preceded changes on UNACC. Also, slight changes in ES and percent accuracy were observed on PA coincided with improvements as the trained UNACC (see Figure 1a).

With regards to participants receiving the SPM treatment, BUMA03 was trained on OR, and changes in ES and percent accuracy were observed on 3NP that preceded improvements on OR as well as changes in ES and percent accuracy on UNACC and ORCNP that coincided with changes in OR. Additionally, changes in UNACC preceded changes in ORCNP and PA indicating that several untrained structures improved when OR was trained in this patient (see supplemental data). BUMA15 was trained on OC and also showed changes in ES and percent accuracy in PA concurrent with OC. This patient also showed changes in ES and percent accuracy for OR and UNACC but the cross correlations were not significant. The lack of congruency between the cross-correlations and the outcome measures for the untrained structures seen in BUMA15 is not problematic because recall that cross-correlations are averaged over time and, thus, the coefficients are reflective of similarities in trend lines between the two structures. BUMA16 was trained on PA; there were coinciding improvements in ES and percent accuracy in both UNACC and OC. Also, negative correlations were observed for both 3NP & OR indicating that as comprehension of PA structures improved, comprehension of OR and 3NP declined (See Figure 1b). BUMA08 was trained on UNACC which did not improve, however, changes in ES and percent accuracy were observed in 3NP, OR and ORCNP; none of the correlations were significant due to the lack of improvements in UNACC. Lastly BUMA23 was trained on UNACC and slight changes in ES and percent accuracy were observed on OR that lagged changes on UNACC.

## Discussion

The goal of this project was to develop a theoretically motivated and effective treatment for sentence comprehension deficits in individuals with aphasia. Our basic hypothesis was that a therapy that trained participants on the relationship between syntactic structure and the meanings of sentences by having participants indicate thematic roles on pictures or by manipulating objects may result in an improved ability to utilize resources in the service of syntactic comprehension.

In general, the results support our hypothesis that the SPM and OM treatments, both based on strengthening thematic role mapping, were generally successful in improving sentence comprehension of noncanonical sentences. Treatment was clearly effective, based on the stringent criteria we set for efficacy using three different metrics. When the combination of the three criteria was taken into account, there were more participants who met our criteria in the OM treatment (4/7) than in the SPM treatment (2/6). While neither percent change nor effect size independently captured the nature of the effect of treatment, it could be argued that these criteria are somewhat arbitrary and not particularly indicative of any psychological aspect of sentence comprehension. In contrast, the third and main criterion; improvement from below/at chance accuracy to above chance accuracy on the trained structure may be more reflective of the ability of an individual to comprehend syntactically complex sentences. Based on this metric, with the exception of three participants (BUMA14; BUMA13 and BUMA08), all participants comprehended the structure they were respectively trained on reliably above chance at the end of treatment. We should note that even though other participants showed above chance comprehension before treatment, each of them improved to greater levels of accuracy subsequent to treatment BUMA21 (53% to 88% accuracy) and BUMA03 (75% to 95% accuracy).

Interestingly, even though there were no significant differences between the two treatment dosage (4hrs/week vs 2 hrs/week), four out of the seven participants receiving the therapy

for four hours a week met treatment effectiveness criteria, whereas only three out of the eight participants who received the therapy for 2 hours a week met our criteria for efficacious treatment. In addition, numerical trends in the data indicated that patients receiving 4hrs/week showed mean effect sizes and percent change accuracy that was larger than the mean effect sizes and percent change accuracy for patients receiving treatment 2hrs/week. Therefore, the dosage of therapy is arguably has some sort of effect on treatment outcome; however, given the lack of significant findings, any conclusions regarding the optimal treatment dosage for our sentence comprehension treatment are at best speculative at this point. We should note that our initial experimental design aimed to compare therapy provided for 1 hour 4 times a week with therapy provided for 2 hours 2 times a week. However, participants did not attend therapy 4 times a week consistently, requiring us to modify our treatment dosage comparison. This speaks to the tractability of actually prescribing and carrying out daily therapy, at least in a research/clinical setting.

Other than the observation that more participants met our criteria in the OM treatment than in the SPM treatment, differences between SPM and OM therapy were nonsignificant with respect to effect size and percent change on the trained structure, indicating that both treatments were equally effective in improving comprehension of the trained structure. This result is hardly surprising; participants were trained until they achieved criterion and the treatment protocols for SPM and OM were identical except for utilization of pictures in SPM and object enactment in OM. There are some initial speculations for why more participants met the criteria on the OM therapy than the SPM therapy. It is possible that enacting the thematic roles during therapy may engage resources required for sentence processing differently than while marking thematic roles during picture matching. Consequently, enacting thematic roles with consistent feedback may lend itself to greater accuracy at identifying thematic roles for noncanonical sentences than marking thematic roles for picture matching. Clearly, the notion of a presumed relation between perceptual, motor and language systems that is equally active in the execution and observation of actions has received extensive support (i.e., embodied communication, Rizzolatti & Arbib, 1998) and the difference between the demands of an enactment based task and sentence to picture matching task has been observed by other researchers (Salis & Edwards, 2009). Several recent rehabilitation approaches of aphasia are based on the notion of capitalizing on this interaction between action and language and are mostly geared towards noun and/or verb retrieval (Marangolo et al., 2010; Rose & Douglas, 2001; Raymer et al., 2006). The OM treatment described here provides preliminary evidence for the benefit of enactment in sentence comprehension and provides an important contribution to therapies for sentence comprehension deficits which have until now mostly used picture or written word stimuli.

It is interesting to note that even though fewer participants receiving SPM treatment met criteria for improvement on the monitoring probes, when changes on the post versus pre-treatment screeners were compared for participants trained on the SPM task relative to participants trained on the OM, the opposite results were found. Participants trained on the SPM task showed significant improvements on their trained and untrained sentences in SPM tasks whereas participants trained on OM tasks did not show significant improvements on the OM task screeners. Therefore, it appears that although OM treatment is more effective as a therapy approach, the SPM treatment results in greater overall changes in comprehending a variety of syntactically manipulated sentences. These observations are preliminary and require further careful examination with larger groups of participants receiving both SPM and OM treatments.

Finally, we examined patterns of generalization from the trained structure to untrained structures within the trained task. Results showed that when generalization occurred, structure was not a barrier to generalization; participants showed generalization within and

across NP and WH-movement structures specifically between PA and OC (e.g., BUMA16; BUMA11). Other participants showed generalization within movement structures (e.g., BUMA20 from OC to OR; BUMA24 from UNACC to PA) other participants showed generalization across movement structures. These results suggest that because participants have resource reductions and not specific deficits; training participants on the relationship between syntactic structure and the meanings of sentences by having participants indicate thematic roles on pictures or by manipulating objects may have resulted in an improved ability to utilize resources to facilitate comprehension of other noncanonical sentences. Two participants (BUMA 24, BUMA08) showed positive changes on 3NP which are canonical sentences, suggesting that the increased sentence length with less complex syntactic structure (such as canonical sentences) require fewer resources than the sentence with complex syntactic structures (such as noncanonical sentences). Changes on ORCNP were observed in two participants (BUMA10, BUMA03) who received training on OR and one (BUMA20) who received training on OC again suggesting that improved ability to utilize resources for noncanonical sentences facilitated comprehension of other noncanonical sentences.. With regards to CATE's prediction about generalization only from complex to simple structures, the results are equivocal at this point; some participants showed generalization from complex to simpler structures whereas other showed improvements from simple to more complex structures. For instance, BUMA03 was trained on OR and improved on OC, with OR being more complex than OC. BUMA16 was trained on PA and improved on UNACC, with PA being more complex than UNACC. In contrast, BUMA 24 showed improvements from UNACC to PA and BUMA 20 showed improvements from OC to OR.

Nonetheless, these generalization results provide preliminary but promising evidence for the utility of the therapy program in terms of facilitating improvements to untrained sentence structures. From a theoretical standpoint, if what is trained is an ability to utilize resources in the service of syntactic comprehension, these results suggest that successful training of the trained structures results in improvements to structures that require the same or fewer resources. Clearly, the next step is to establish the extent to which treatment generalizes to untrained structures and tasks, savings in further treatment, maintenance, and transfer to discourse which is more comparable to real world comprehension settings.

## Conclusions

In conclusion, we have successfully developed and implemented two treatments for syntactic comprehension on 15 participants with sentence comprehension deficits. In participants where treatment was efficacious, a significant increase in the level of performance from pre- to-post-treatment probes was observed for the trained structure and resulted in above chance performance. Results also show the efficacy of the treatment approach in training two syntactic structures (WH-movement structures: OR and OC, NP-movement structures: PA and UNACC). Our continued work in this area involves a larger group of participants to systematically examine the nature of generalization, maintenance and transfer resulting from this sentence comprehension treatment.

## Acknowledgments

This project is funded by NIDCD 1R21DC010461-01. The authors thank Marissa Simms, Rebecca Hufford, and Balaji Rangarathnam for their assistance in data collection and Jennifer Richardson for her assistance in data analysis.

## References

- Beretta A. Linear and structural accounts of theta-role assignment in agrammatic aphasia. *Aphasiology*. 2001; 15:515–531.
- Beretta A, Harford C, Patterson J, Piñango M. The derivation of post-verbal subjects: Evidence from agrammatic aphasia. *Natural Language and Linguistic Theory*. 1996; 14:725–748.
- Berndt RS, Mitchum CC, Burton MW, Haendiges AN. Comprehension of reversible sentences in aphasia: The effects of verb meaning. *Cognitive Neuropsychology*. 2004; 21(2–4):229–244. [PubMed: 21038202]
- Byng S. Sentence processing deficits: Theory and therapy. *Cognitive Neuropsychology*. 1988; 5(6): 629–676.
- Box, G.; Jenkins, GM.; Reinsel, G. *Time Series Analysis: Forecasting & Control*. Prentice Hall; New Jersey: 1994.
- Caplan D, Baker C, Dehaut F. Syntactic determinants of sentence comprehension in aphasia. *Cognition*. 1985; 21(2):117–175. [PubMed: 2419022]
- Caplan D, DeDe G, Michaud J. Task-independent and task-specific syntactic deficits in aphasic comprehension. *Aphasiology*. 2006; 20:893–920.
- Caplan D, Futter C. Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language*. 1986; 27(1):117–134. [PubMed: 3947937]
- Caplan D, Hildebrandt N. Disorders affecting comprehension of Syntactic Form. Preliminary results and their implications for theories of syntax and parsing. *Canadian Journal of Linguistics*. 1988a; 33:477–505.
- Caplan D, Hildebrandt N. Specific deficits in syntactic comprehension. *Aphasiology*. 1988b; 2:255–258.
- Caplan D, Waters GS. Syntactic processing in sentence comprehension under dual-task conditions in aphasic participants. *Language and Cognitive Processes*. 1996; 11(5):525–551.
- Caplan, D.; Waters, GS. Comprehension disorders in aphasia: The case of sentences that require syntactic analysis. In: Traxler, M., editor. *The Handbook of Psycholinguistics*. Oxford: OUP; 2006.
- Caplan D, Waters G, DeDe G, Michaud J, Reddy A. A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain and Language*. 2007; 101(2):103–150. [PubMed: 16999989]
- Caplan D, Waters GS, Hildebrandt N. Determinants of sentence comprehension in aphasic patients in sentence–picture matching. *Journal of Speech, Language, and Hearing Research*. 1997; 40:542–555.
- Cupples L, Inglis AL. When task demands induce “asyntactic” comprehension: A study of sentence interpretation in aphasia. *Cognitive Neuropsychology*. 1993; 10:201–234.
- Chomsky, N. *The minimalist program*. MIT Press; 1995.
- Crerar MA, Ellis AW, Dean EC. Remediation of Sentence Processing Deficits in Aphasia Using a Computer-Based Microworld. *Brain and Language*. 1996; 52(1):229–275. [PubMed: 8741982]
- DeDe G, Caplan D. Factor Analysis of Syntactic Deficits in Aphasic Comprehension. *Aphasiology*. 2006; 20:123–135.
- Grodzinsky Y. The neurology of syntax: Language use without Broca’s area. *Behavioral and Brain Sciences*, 2000. 2000; 23:1–21.
- Haarmann HJ, Kolk HH. On-line sensitivity to subject-verb agreement violations in Broca’s aphasics: the role of syntactic complexity and time. *Brain and Language*. 1994; 46(4):493–516. 10.1006/brln.1994.1028 [PubMed: 8044674]
- Haendiges AN, Berndt RS, Mitchum CC. Assessing the elements contributing to a “mapping” deficit: A targeted treatment study. *Brain and Language*. 1996; 52(1):276–302. [PubMed: 8741983]
- Jacobs BJ, Thompson CK. Cross-modal generalization effects of training noncanonical sentence comprehension and production in agrammatic aphasia. *Journal of Speech, Language and Hearing Research*. 2000; 43(1):5–20.

- Linebarger MC. Agrammatism as evidence about grammar. *Brain and Language*. 1995; 50:52–91. [PubMed: 7552230]
- Maurer G. Examining the empirical and linguistic bases of current theories of agrammatism. *Brain and Language*. 1995; 50:339–368. [PubMed: 7583194]
- Maurer G, Fromkin VA, Cornell TL. Comprehension and acceptability judgements in agrammatism: Disruptions in the syntax of referential dependency. *Brain and Language*. 1993; 45:340–370. [PubMed: 8269330]
- McAllister T, Bachrach A, Waters G, Michaud J, Caplan D. Production and comprehension of unaccusatives in aphasia. *Aphasiology*. 2009; 23:989–1004.
- Marangolo P, Bonifazi S, Tomaiuolo F, Craighero L, Coccia M, Altoé G, Cantagallo A. Improving language without words: First evidence from aphasia. *Neuropsychologia*. 2010; 48(13):3824–3833.10.1016/j.neuropsychologia.2010.09.025 [PubMed: 20887740]
- Mitchum CC, Greenwald ML, Berndt RS. Cognitive treatments of sentence processing disorders: What have we learned? *Neuropsychological Rehabilitation*. 2000; 10(3):311–336.
- Mitchum CC, Haendiges AN, Berndt RS. Treatment of thematic mapping in sentence comprehension: Implications for normal processing. *Cognitive Neuropsychology*. 1995; 12(5):503–547.
- Mitchum CC, Haendiges AN, Berndt RS. Response strategies in aphasic sentence comprehension: An analysis of two cases. *Aphasiology*. 2004; 18(8):675–692.
- Nickels L, Byng S, Black M. Sentence processing deficits: A replication of therapy. *British Journal of Disorders of Communication*. 1991; 26(2):175–199. [PubMed: 1777398]
- Rizzolatti G, Arbib MA. Language within our grasps. *Trends in Neuroscience*. 1998; 21:188–194.
- Raymer AM, Singletary F, Rodriquez A, Ciampitti M, Heilman KM, Rothi LJG. Effects of gesture + verbal treatment for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*. 2006; 12:867–882. [PubMed: 17064449]
- Rose M, Douglas J. The differential facilitatory effect of gesture and visualization processes on object naming in aphasia. *Aphasiology*. 2001; 15:977–990.
- Robey RR, Schultz MC, Crawford AB, Sinner CA. Single-subject clinical-outcome research: Designs, data, effect sizes, and analyses. *Aphasiology*. 1999; 13(6):445–473.
- Salis C, Edwards S. Tests of syntactic comprehension in aphasia: An investigation of task effects. *Aphasiology*. 2009; 23:1215–1230.
- Schwartz MF, Saffran EM, Fink RB, Myers JL, et al. Mapping therapy: A treatment programme for agrammatism. *Aphasiology*. 1994; 8(1):19–54.
- Thompson CK. Single subject controlled experiments in aphasia: the science and the state of the science. *Journal Communication Disorders*. 2006; 39(4):266–291.
- Thompson CK, Ballard KJ, Shapiro LP. The role of syntactic complexity in training wh-movement structures in agrammatic aphasia: optimal order for promoting generalization. *Journal of International Neuropsychological Society*. 1998; 4(6):661–674.
- Thompson CK, Shapiro LP, Ballard KJ, Jacobs BJ, Schneider SS, Tait ME. Training and generalized production of wh- and NP-movement structures in agrammatic aphasia. *Journal of Speech, Language and Hearing Research*. 1997; 40(2):228–244.
- Thompson CK, Shapiro LP. Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology*. 2005; 19(10–11):1021–1036. [PubMed: 17410280]
- Thompson CK, Shapiro LP. Complexity in treatment of syntactic deficits. *American Journal of Speech Language Pathology*. 2007; 16(1):30–42. [PubMed: 17329673]
- Thompson CK, Shapiro LP, Kiran S, Sobecks J. 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*. 2003; 46(3):591–607.
- Tseng C-I, McNeil MR, Milenkovic P. An investigation of attention allocation deficits in aphasia. *Brain and Language*. 1993; 45:276–296. [PubMed: 8358600]

## Appendix A Sample Stimuli Used in Treatment

<b>Screeners</b>		
<b>Structure</b>	<b>Number of Items</b>	<b>Example</b>
Active	10	The helicopter trapped the airplane.
Noun Phrase Raising	10	The niece seemed to the wife to be pushing the boy.
Cleft Object	10	It was the bike that the tank chased.
Object Control	10	The man convinced the king to follow the wife.
Object Relative	10	The daughter who the nephew rescued poked the uncle.
Passive	10	The boy was nudged by the aunt.
Pronoun-as-object	10	The man said that the boy kissed him.
Reflexive	10	The mother said that the girl kicked herself.
Subject Control	10	The nephew pledged to the grandmother to hug the niece.
Unaccusative	10	The woman was drowning.
3 Noun Phrase	5	The girl said that the aunt fed the man.
Object Relative w/Complex NP	5	The girl who the husband of the woman kicked touched the boy.

<b>Monitors</b>		
<b>Structure</b>	<b>Number of Items</b>	<b>Example</b>
3 Noun Phrase	10	The father said that the boy hugged the aunt.
Object Relative w/Complex NP	10	The train that the door of the car scratched destroyed the helicopter.
Unaccusative	10	The man was spinning.
Passive	15	The bike was dragged by the car.
Cleft Object	15	It was the husband who the wife tickled.
Object Relative	15	The diamond that the saw cut destroyed the ball.

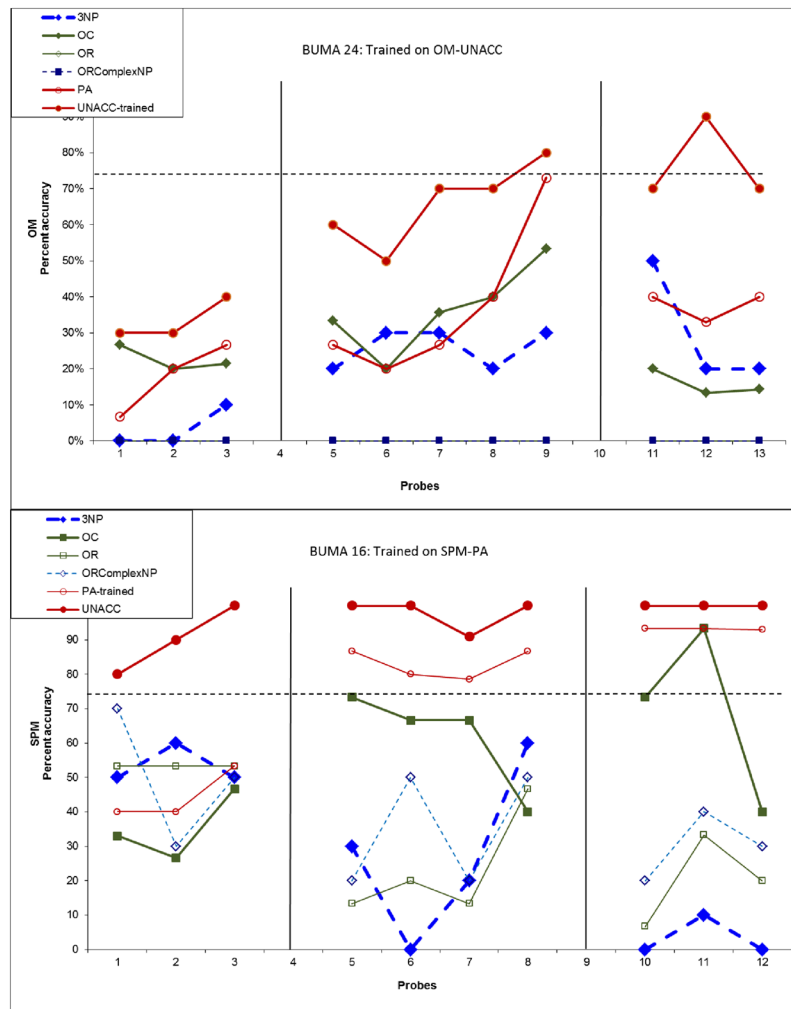
<b>Therapy Items (Each patient was assigned to one of these structures)</b>		
<b>Structure</b>	<b>Number of Items</b>	<b>Example</b>
Unaccusative	10	The king was healing.
Passive	20	The cannon was scratched by the truck.
Cleft Object	20	It was the water that the blanket covered.
Object Relative	20	The husband who the girl greeted bit the wife.

## Appendix B Sample Treatment Protocols for SPM and OM Passive Sentences

<b>SPM Treatment Protocol</b>	<b>OM Treatment Protocol</b>
<p>For example, the SPM-PA (e.g., <i>the aunt was kissed by the man</i>) treatment protocol included the following steps:</p> <ol style="list-style-type: none"> <li>1 Upon hearing the sentence read aloud by the clinician, the patient chose between the target picture depicting that action and a foil picture</li> </ol>	<p>Therefore, for the same sentence, (e.g., <i>the aunt was kissed by the man</i>) treatment protocol included the following steps:</p> <ol style="list-style-type: none"> <li>1 Upon hearing the sentence read aloud by the clinician, the patient enacted the target</li> </ol>

SPM Treatment Protocol	OM Treatment Protocol
<p>with reversed thematic roles (i.e. depicting <i>the man was kissed by the aunt</i>), and feedback regarding accuracy was provided;</p> <p>2 The clinician removed the foil picture, placed a strip of paper with the target sentence printed on it beneath the target picture, and further explained the roles of the agent and theme (e.g., “In this sentence, <i>the aunt was kissed by the man</i>, the man is kissing the aunt.”);</p> <p>3 The clinician emphasized the agent of the action by placing a version of the target picture in which the agent was bolded/emphasized and the theme was un-bolded/de-emphasized, as well as a sentence strip in which the agent was bolded/emphasized and the theme was un- bolded/de-emphasized, on the table and saying, “In this part of the sentence, <i>the aunt was kissed by the man</i>, the man is doing the action. Ok, now you show me who is doing the action”. Feedback regarding accuracy was again provided;</p> <p>4 The clinician emphasized the theme of the action in a similar manner, using both a picture and sentence strip in which the theme was bolded/emphasized and the agent was un- bolded/de-emphasized, again provided feedback as to accuracy;</p> <p>5 The clinician put the original target picture and target sentence strip back on the table and once again read the full sentence and explained the roles of the agent and theme;</p> <p>6 The clinician read the target sentence aloud and gave the patient the target picture and the foil picture to choose from.</p>	<p>sentence using dolls (<i>aunt, man</i>) and feedback regarding accuracy was provided;</p> <p>2 The clinician enacted the sentence using the dolls, placed a strip of paper with the target sentence printed on it beneath the target picture and further explained the roles of the agent and theme (e.g., “In this sentence, <i>the aunt was kissed by the man</i>, the man is kissing the aunt.”);</p> <p>3 The clinician emphasized the agent of the action as well as a sentence strip in which the agent was bolded/emphasized and the theme was un- bolded/de-emphasized, on the table and saying, “In this part of the sentence, <i>the aunt was kissed by the man</i>, the man is doing the action. Ok, now you show me who is doing the action”. Feedback regarding accuracy was again provided;</p> <p>4 The clinician emphasized the theme of the action in a similar manner (e.g., “In this part of the sentence, <i>the aunt was kissed by the man</i>, the woman is being kissed. Ok, now you show me who is receiving the action”) and then the theme of the action was emphasized in a similar manner</p> <p>5 The clinician put the original target picture and target sentence strip back on the table and once again read the full sentence and explained the roles of the agent and theme;</p> <p>6 The patient enacted the sentence with feedback from the clinician.</p>





**Figure 1.** Results for two participants (BUMA24, BUMA16) showing trends on trained structures and untrained structures within the trained task. Red lines indicate NP movement structures, green lines indicate Wh-movement structures, blue dashed lines indicate control structures.

**Table 1**  
Demographic Information and Pre-Treatment Screener Performance for 15 Participants Enrolled in the Treatment Study

Participant	Age	Sex	Race	Years of Education	MPO	Overall SPM Screener (pre-testing)	Overall OM Screener (pre-testing)	Noun Test	Verb Test
BUMA03	67	F	W	16	72	83.6%	90.0%	97.8%	91.2%
BUMA05	54	M	W	12	111	73.6%	54.5%	91.3%	100.0%
BUMA07	29	M	B	12	6	47.3%	26.4%	97.8%	87.5%
BUMA08	62	F	W	16	54	42.7%	41.8%	91.3%	89.1%
BUMA10	65	M	W	18	60	85.5%	55.5%	100.0%	92.2%
BUMA11	59	M	W	16	147	66.4%	39.1%	91.3%	79.4%
BUMA13	63	M	W	16	93	51.8%	26.4%	95.7%	90.6%
BUMA14	63	M	W	18	96	52.7%	30.9%	86.9%	73.5%
BUMA15	59	M	W	18	18	72.7%	62.7%	100.0%	89.1%
BUMA16	56	M	W	18	76	51.8%	41.8%	95.7%	92.2%
BUMA17	73	F	W	14	36	62.7%	44.6%	97.8%	90.6%
BUMA20	45	M	W	20	15	79.1%	77.3%	100.0%	95.3%
BUMA21	39	F	W	20	9	84.6%	89.1%	91.3%	92.2%
BUMA23	65	F	W	12	32	48.2%	27.3%	89.1%	68.8%
BUMA24	59	M	W	14	6	44.6%	20.0%	100.0%	92.2%

Note. MPO = months post onset, SPM = sentence to picture matching, OM = object manipulation, F = female, M = male, W = white, B = black.

**Table 2**

Patient Performance on the Three Criteria Set for Efficacious Treatment

Patient	Trained Task	Trained Structure	Frequency	Outcome Measures	3NP	OC	OR	OR Complex NP	PA	UNACC
BUMA10	OM	OR	2 × 2	ES: change in % acc:	0.00	0.56	<b>23.00</b>	4.21	0.00	0.58
				chance:	AC → AC	AC → AC	<b>BC → AC</b>	BC → C	AC → AC	AC → AC
BUMA21	OM	OR	2 × 1	ES: change in % acc:	0.00	0.00	<b>3.33</b>	0.00	0.58	0.00
				chance:	AC → AC	AC → AC	<b>AC → AC</b>	C → C	AC → AC	AC → AC
BUMA05	OM	OC	2 × 1	ES: change in % acc:	1.73	<b>1.37</b>	0.03	0.00	0.44	0.00
				chance:	10%	<b>18%</b>	0%	0%	5%	0%
BUMA20	OM	OC	2 × 2	ES: change in % acc:	-0.43	<b>4.88</b>	11.36	5.01	0.00	0.00
				chance:	AC → AC	<b>C → AC</b>	BC → BC	BC → BC	AC → AC	AC → AC
BUMA11	OM	PA	2 × 2	ES: change in % acc:	-2.79	2.28	-0.96	-0.87	<b>3.24</b>	-0.07
				chance:	-48%	23%	-6%	-5%	<b>47%</b>	-1%
BUMA14	OM	PA	2 × 1	ES: change in % acc:	0.58	-2.40	0.22	0.00	<b>-1.53</b>	0.67
				chance:	7%	-24%	2%	0%	<b>-16%</b>	7%
BUMA24	OM	UNACC	2 × 1	ES: change in % acc:	4.62	-1.94	0.00	0.00	<b>1.95</b>	<b>7.51</b>
				chance:	27%	-7%	0%	0%	<b>20%</b>	<b>43%</b>
BUMA03	SPM	OR	2 × 2	ES: change in % acc:	2.31	0.00	<b>2.60</b>	2.67	0.67	0.58
				chance:	13%	0%	<b>20%</b>	27%	4%	3%
BUMA17	SPM	OR	2 × 1	No effects	AC → AC	AC → AC	<b>AC → AC</b>	AC → AC	AC → AC	AC → AC
BUMA15	SPM	OC	2 × 2	ES: change in % acc:	0.41	<b>1.26</b>	1.09	-0.65	1.83	1.17
				chance:	3%	<b>11%</b>	15%	-8%	6%	18%
BUMA07	SPM	OC	2 × 1	No effects	AC → AC	<b>AC → AC</b>	C → C	C → C	AC → AC	AC → AC

Patient	Trained Task	Trained Structure	Frequency	Outcome Measures	3NP	OC	OR	OR Complex NP	PA	UNACC
BUMA13	SPM	PA	2 × 2	ES: change in % acc: chance:	-0.33 -3% C → C	0.48 7% BC → C	-0.29 -2% C → C	0.22 3% C → C	<b>0.50</b> 9% C → C	1.15 7% AC → AC
BUMA16	SPM	PA	2 × 1	ES: change in % acc: chance:	-8.66 -50% C → BC	3.27 33% BC → AC	-3.26 -33% C → BC	-1.00 -20% C → BC	<b>6.34</b> 48% C → AC	1.00 10% AC → AC
BUMA08	SPM	UNACC	2 × 1	ES: change in % acc: chance:	6.35 37% BC → C	-0.44 -4% C → C	2.31 9% C → C	2.31 13% C → C	-1.33 -9% C → C	<b>0.58</b> 3% C → C
BUMA23	SPM	UNACC	2 × 2	ES: change in % acc: chance:	-2.00 -3% C → C	-0.30 -2% C → C	0.70 7% C → C	-0.40 -13% C → BC	-2.70 -18% C → C	<b>5.77</b> 33% C → AC

Note. In order for treatment to be considered effective for each participant, accuracy should rise from at or below chance to above chance and either a) accuracy on the treated structure rises by 33% or b) the effect size of training on the treated structure is 2.6. BC → AC indicates improvement from below chance (BC) to above chance (AC) and so on. The trained structures are bolded.

Table 3

Patient Performance on the Screeners Pre- and Post-Treatment

Sentence Type	BUMA 10		BUMA 21		BUMA 05		BUMA 20		BUMA 11		BUMA 14		BUMA 24	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
3NP	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	60.0%	20.0%	0.0%	20.0%	0.0%	20.0%
active	80.0%	100.0%	100.0%	100.0%	100.0%	90.0%	50.0%	50.0%	60.0%	20.0%	70.0%	50.0%	30.0%	10.0%
raising NP	10.0%	80.0%	100.0%	90.0%	60.0%	80.0%	20.0%	80.0%	10.0%	0.0%	20.0%	50.0%	30.0%	10.0%
object cleft	100.0%	100.0%	100.0%	90.0%	40.0%	60.0%	100.0%	40.0%	40.0%	40.0%	20.0%	70.0%	40.0%	30.0%
object control	50.0%	100.0%	100.0%	90.0%	60.0%	100.0%	80.0%	80.0%	30.0%	20.0%	10.0%	30.0%	0.0%	30.0%
object relative	60.0%	40.0%	60.0%	50.0%	20.0%	0.0%	50.0%	0.0%	0.0%	10.0%	10.0%	0.0%	0.0%	0.0%
OR complex NP	20.0%	20.0%	0.0%	0.0%	0.0%	0.0%	20.0%	0.0%	0.0%	20.0%	0.0%	0.0%	0.0%	0.0%
passive	100.0%	100.0%	100.0%	80.0%	70.0%	100.0%	100.0%	100.0%	60.0%	80.0%	40.0%	50.0%	20.0%	20.0%
pronoun	50.0%	90.0%	80.0%	90.0%	30.0%	80.0%	90.0%	90.0%	10.0%	50.0%	30.0%	60.0%	20.0%	10.0%
reflexive	0.0%	100.0%	100.0%	90.0%	0.0%	100.0%	100.0%	100.0%	80.0%	60.0%	30.0%	0.0%	10.0%	0.0%
subject control	40.0%	60.0%	90.0%	90.0%	80.0%	10.0%	90.0%	30.0%	10.0%	10.0%	10.0%	70.0%	20.0%	20.0%
unaccusative	0.0%	100.0%	100.0%	90.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	50.0%	90.0%
TOTAL	55.5%	84.5%	89.1%	82.7%	54.5%	62.7%	77.3%	63.6%	39.1%	37.3%	30.9%	44.5%	20.0%	20.9%

Sentence Type	BUMA 03		BUMA 17		BUMA 15		BUMA 07		BUMA 13		BUMA 16		BUMA 23	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
3NP	100.0%	100.0%	20.0%	80.0%	80.0%	80.0%	60.0%	20.0%	20.0%	40.0%	80.0%	20.0%	80.0%	40.0%
active	90.0%	100.0%	60.0%	70.0%	30.0%	60.0%	60.0%	50.0%	50.0%	80.0%	80.0%	70.0%	40.0%	60.0%
raising NP	70.0%	100.0%	50.0%	80.0%	60.0%	60.0%	60.0%	80.0%	80.0%	70.0%	30.0%	60.0%	30.0%	30.0%
object cleft	100.0%	90.0%	50.0%	70.0%	50.0%	30.0%	30.0%	40.0%	40.0%	50.0%	50.0%	30.0%	60.0%	60.0%
object control	90.0%	90.0%	60.0%	60.0%	70.0%	30.0%	50.0%	50.0%	50.0%	70.0%	30.0%	50.0%	50.0%	60.0%
object relative	50.0%	90.0%	30.0%	80.0%	30.0%	40.0%	40.0%	60.0%	60.0%	60.0%	30.0%	60.0%	60.0%	40.0%
OR complex NP	100.0%	100.0%	20.0%	80.0%	80.0%	60.0%	60.0%	60.0%	60.0%	40.0%	40.0%	80.0%	80.0%	60.0%

SPM Training	BUMA 03		BUMA 17		BUMA 15		BUMA 07		BUMA 13		BUMA 16		BUMA 08		BUMA 23	
	OR		OR		OC		OC		PA		PA		UNACC		UNACC	
Sentence Type	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
passive	70.0%	<u>100.0%</u>	90.0%	<u>90.0%</u>	90.0%	<u>90.0%</u>	60.0%	<b>30.0%</b>	<b>40.0%</b>	<b>10.0%</b>	<b>50.0%</b>	60.0%	50.0%	70.0%	80.0%	
pronoun	90.0%	<u>100.0%</u>	90.0%	<u>90.0%</u>	90.0%	<u>90.0%</u>	40.0%	60.0%	50.0%	20.0%	50.0%	60.0%	50.0%	70.0%	70.0%	
reflexive	70.0%	<u>100.0%</u>	90.0%	<u>100.0%</u>	90.0%	<u>100.0%</u>	70.0%	70.0%	<u>100.0%</u>	<u>100.0%</u>	<u>100.0%</u>	10.0%	60.0%	60.0%	<u>90.0%</u>	
subject control	90.0%	<u>100.0%</u>	70.0%	70.0%	80.0%	<u>90.0%</u>	20.0%	30.0%	80.0%	30.0%	70.0%	30.0%	40.0%	20.0%	70.0%	
unaccusative	<u>100.0%</u>	<u>100.0%</u>	80.0%	80.0%	90.0%	<u>100.0%</u>	50.0%	80.0%	<u>100.0%</u>	80.0%	<u>100.0%</u>	<b>60.0%</b>	<b>100.0%</b>	<b>60.0%</b>	<b>100.0%</b>	
<i>TOTAL</i>	<i>83.6%</i>	<i>97.3%</i>	<i>62.7%</i>	<i>ND</i>	<i>72.7%</i>	<i>81.8%</i>	<i>47.3%</i>	<i>ND</i>	<i>51.8%</i>	<i>61.8%</i>	<i>51.8%</i>	<i>65.5%</i>	<i>42.7%</i>	<i>56.4%</i>	<i>48.2%</i>	<i>64.5%</i>

*Note.* The performance level for the trained structure for each patient is bolded. Performance levels above chance are underlined. All other values are at or below chance level comprehension. For SPM, all structures have a chance of 50% accuracy. For OM, depending on the structure, there is either a 50% change of guessing correctly or a 33% chance. Except ORCNP & 3NP which have 5 items. all other structures have 10 items. ND: indicates no data, as patients did not finish treatment.

**Table 4**

Cross-Correlation Results (R Values Reported in Parenthesis) for Untrained Structures That Show a Correlation with the Trained Structure for Each Patient

Patient	OM Training	Trained Structure	Positive Correlations			Negative Correlations			
			Negative Lag	0 Lag	Positive Lag	Negative Lag	0 Lag	Positive Lag	
BUMA10		OR		with ORCNP (.839)					
BUMA21		OR							
BUMA05		OC							
BUMA20		OC		with OR (.947) with ORCNP (.630)					
BUMA11		PA	with OC (.665)			with 3NP (-.723) with ORCNP (-.641)			
BUMA14		PA							
BUMA24		UNACC	with 3NP (.751)		with PA (.811)				
SFM Training		Trained Structure		Positive Correlations			Negative Correlations		
Patient			Negative Lag	0 Lag	Positive Lag	Negative Lag	0 Lag	Positive Lag	
BUMA03		OR	with 3NP (.900)	with UNACC (.606) with ORCNP (.627)					
BUMA17		OR	No effects						
BUMA15		OC		with PA (.594)					
BUMA07		OC	No effects						
BUMA13		PA							
BUMA16		PA		with UNACC (.740) with OC (.675)			with OR (-.758) with 3NP (-.707)		
BUMA08		UNACC							
BUMA23		UNACC	with OR (.635)						

Note. Only correlations that exceed .60 and two standard errors are reported. Negative lags indicate lags of -1, -2, -3 sessions, positive lags indicate 1, 2, or 3 sessions, 0 lag indicates a change on the same session. Positive correlations indicate both structures improved, negative correlations indicate that as one structure improved, the other decreased.