

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Training verb argument structure production in agrammatic aphasia: Behavioral and neural recovery patterns

Cynthia K. Thompson^{a,b,c}, Ellyn A. Riley^d, Dirk-Bart den Ouden^e, Aya Meltzer-Asscher^{f,g,*} and Sladjana Lukic^a

^a Aphasia and Neurolinguistics Research Laboratory, Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA

^b Department of Neurology, Northwestern University, Evanston, IL, USA

^c Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Evanston, IL, USA

^d Department of Communication Sciences and Disorders, Bowling Green State University, USA

^e Department of Communication Sciences and Disorders, University of South Carolina, USA

^f Linguistics Department, Tel Aviv University, Israel

^g Sagol School of Neuroscience, Tel Aviv University, Israel

ARTICLE INFO

Article history:

Received 26 March 2012

Reviewed 7 June 2012

Revised 6 September 2012

Accepted 4 February 2013

Action editor Roberto Cubelli

Published online xxx

Keywords:

Neuroplasticity

Agrammatic aphasia

Verb argument structure

fMRI

Aphasia treatment

ABSTRACT

Introduction: Neuroimaging and lesion studies indicate a left hemisphere network for verb and verb argument structure processing, involving both frontal and temporoparietal brain regions. Although their verb comprehension is generally unimpaired, it is well known that individuals with agrammatic aphasia often present with verb production deficits, characterized by an argument structure complexity hierarchy, indicating faulty access to argument structure representations for production and integration into syntactic contexts. Recovery of verb processing in agrammatism, however, has received little attention and no studies have examined the neural mechanisms associated with improved verb and argument structure processing. In the present study we trained agrammatic individuals on verbs with complex argument structure in sentence contexts and examined generalization to verbs with less complex argument structure. The neural substrates of improved verb production were examined using functional magnetic resonance imaging (fMRI).

Methods: Eight individuals with chronic agrammatic aphasia participated in the study (four experimental and four control participants). Production of three-argument verbs in active sentences was trained using a sentence generation task emphasizing the verb's argument structure and the thematic roles of sentential noun phrases. Before and after training, production of trained and untrained verbs was tested in naming and sentence production and fMRI scans were obtained, using an action naming task.

Results: Significant pre- to post-training improvement in trained and untrained (one- and two-argument) verbs was found for treated, but not control, participants, with between-group differences found for verb naming, production of verbs in sentences, and production of argument structure. fMRI activation derived from post-treatment compared to pre-

* Corresponding author. Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel.

E-mail address: ameltzer@post.tau.ac.il (A. Meltzer-Asscher).

0010-9452/\$ – see front matter © 2013 Elsevier Ltd. All rights reserved.

<http://dx.doi.org/10.1016/j.cortex.2013.02.003>

treatment scans revealed upregulation in cortical regions implicated for verb and argument structure processing in healthy controls.

Conclusions: Training verb deficits emphasizing argument structure and thematic role mapping is effective for improving verb and sentence production and results in recruitment of neural networks engaged for verb and argument structure processing in healthy individuals.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Individuals with agrammatic aphasia often present with verb production deficits. Notably, recent research indicates that verbs with greater (*vs* lesser) linguistic complexity are more difficult for these patients to produce. Verbs with more complex argument structure entries (i.e., a greater number of thematic roles and/or theta grids encoded within the verbs' representation) are more difficult compared to verbs with less complex entries. For example, ditransitive verbs like *deliver* and transitive verbs like *fix* are more difficult to produce than intransitive verbs like *laugh*. *Deliver* selects for three arguments: someone who delivers (i.e., an agent), something being delivered (i.e., the theme), and the receiver of the thing being delivered (i.e., the goal). In contrast, the verb *kick* only requires two arguments: an agent and a theme, and the verb *laugh* only requires one argument: an agent. This pattern (i.e., a verb argument structure complexity hierarchy) has been found in English, Dutch, German, Italian, and Russian agrammatic speakers (Bastiaanse and Jonkers, 1998; De Bleser and Kauschke, 2003; Dragoy and Bastiaanse, 2010; Kemmerer and Tranel, 2000; Kim and Thompson, 2000, 2004; Kiss, 2000; Luzzatti et al., 2002; Thompson et al., 1997), leading to the Argument Structure Complexity Hypothesis, stating that as the number of arguments increases for a verb, the more difficult it becomes to produce (Thompson, 2003).

In contrast to verb production deficits, individuals with agrammatic aphasia show retained ability to understand verbs in off-line auditory comprehension tasks (Kim and Thompson, 2000, 2004; but see Miceli et al., 1983) and show normal access to the subcategorization frames of verbs (i.e., encoded information pertaining to the syntactic environments in which the verb may appear) in on-line sentence processing tasks. That is, reaction times (RTs) are longer for verbs with multiple subcategorization options (e.g., the verb *send*) versus those with only one such option (e.g., the verb *fix*), as they are in healthy volunteers (Shapiro et al., 1993; Shapiro and Levine, 1990). Further, in a neuroimaging study using functional magnetic resonance imaging (fMRI) we (Thompson et al., 2010a) found normal activation patterns associated with argument structure complexity in four (of five) individuals with agrammatism, albeit some showed unilateral (right hemisphere – RH) activation because of necrosed tissue in relevant left hemisphere (LH) regions. Notably, patients with anomic aphasia typically present with greater difficulty producing nouns (objects) compared to verbs (actions), and in on-line sentence processing tasks, Wernicke's aphasic individuals do not show differential reaction times (RTs) to verbs based on linguistic complexity, indicating a lack of sensitivity to subcategorization information

associated with verbs (Kim and Thompson, 2004; Shapiro et al., 1993; Shapiro and Levine, 1990). Given the convention that patients with agrammatic aphasia present with lesions in frontal regions and those with anomic aphasia present with temporoparietal lesions, these neurolinguistic studies point to unique roles for anterior and posterior portions of the LH in verb and verb argument structure processing.¹

Neuroimaging studies examining verb processing in healthy individuals coincide with these general aphasic deficit patterns. Studies of verb (*vs* noun) processing using positron emission tomography (PET) and fMRI, as well as repetitive transcranial magnetic stimulation (rTMS) indicate left frontal convexity activation for verb processing and left temporal activation for nouns (e.g., Damasio and Tranel, 1993; Shapiro et al., 2005; Tyler et al., 2004), although inconsistent results across studies have been noted (see Crepaldi et al., 2011, for review). In addition, recent studies examining brain mechanisms involved in verb argument structure computation (i.e., processing of thematic roles encoded within the lexical representation of verbs) suggest that the posterior perisylvian region is also part of the network involved in verb processing. Using fMRI and a lexical decision task, Thompson et al. (2007) and Thompson et al., 2010a found that an increase in the number of arguments engenders increased activation in the angular and supramarginal gyri, bilaterally, in both young and older normal listeners. Ben-Shachar et al. (2003) found a similar pattern using a sentence processing task, with posterior superior temporal sulcus activation associated with increases in the number of arguments selected by verbs embedded in sentences. Further, in a verb production study, Den Ouden et al. (2009) found posterior activation associated with argument structure complexity (i.e., angular and supramarginal gyri, as well as bilateral fusiform, middle occipital, and superior parietal cortex). In addition, they identified activation for transitive compared to intransitive verbs in LH Broca's area – Brodmann areas (BAs) 44 and 45 – and surrounding areas. These findings suggest a LH network for verb processing, involving both frontal and temporoparietal regions.

Despite pervasive verb deficits in patients with agrammatic aphasia, few studies have addressed recovery of verb processing, either in behavioral or neuroimaging studies. Of the few behavioral studies published, most have used semantic

¹ We recognize that many patients who present with behavioral deficits consistent with agrammatic or anomic aphasia do not show these clear-cut lesion patterns. For example, agrammatism may result from large lesions that include frontal as well as temporal and parietal regions (see Caplan et al., 1996; Vanier and Caplan, 1990; also see Wilson and Saygin, 2004).

and/or phonological word retrieval cues to improve action naming, reporting little (or no) generalization from trained to untrained verbs resulting from this approach (Raymer and Ellsworth, 2002; Wambaugh et al., 2004). Studies training verbs in sentence contexts also report limited across-verb generalization. However, improved sentence production has been noted following treatment (e.g., Bastiaanse et al., 2006; Conroy et al., 2009; Marshall et al., 1998; Schneider and Thompson, 2003). One variable that has not been manipulated (or controlled) in treatment for verb deficits is verb argument structure. Indeed, studies focused on recovery of other aphasic deficits (e.g., object naming, sentence comprehension and production deficits) have shown that treatment of complex structures results in generalization to untrained, linguistically related, structures of lesser complexity (Kiran and Thompson, 2003; Thompson et al., 2003; also see Riley, 2011; Maas et al., 2002), resulting in a principle known as the Complexity Account of Treatment Efficacy (CATE; Thompson et al., 2003). Given the relevance of verb argument structure complexity to verb deficits in aphasia as well as the CATE principle, it is possible that training complex verbs may not only enhance production of trained verbs but also improve untrained verbs of lesser complexity. Further, because of the central role that verbs play in sentence computation, such treatment also should impact sentence production.

Treatment which improves production of verbs with complex argument structures also may impact the neural networks that subserve verb processing. In recent years, a number of studies have shown that language training affects brain processing. Although mixed findings have derived from these studies, this literature shows shifts in activation from pre- to post-training in contralesional (typically RH) homologs of damaged left brain regions and/or in perilesional (typically LH) regions (e.g., Belin et al., 1996; Breier et al., 2007; Cao et al., 1999; Crosson et al., 2005; Davis and Harrington, 2006; Leger et al., 2002; Meinzer et al., 2008; Perani et al., 2003; Raboyeau et al., 2008; Rosen et al., 2005; Saur et al., 2006; for reviews, see Crinion and Leff, 2007; Thompson and den Ouden, 2008). Notably, studies examining the neural mechanisms of treatment-induced improvements in word retrieval are constrained to noun or object naming studies (see Thompson and den Ouden, 2008, for review), with no studies to date exploring the neural correlates of improved verb production. However, even for object naming variable recovery patterns have been seen, a finding which, as noted by Thompson and den Ouden (2008), likely relates to a number of factors, including differences in lesion size and location, as well as the language deficits that participants present across studies. In addition, differences in the type of treatment provided, the neuroimaging task used and methods of data analysis utilized have led to variability in this literature (see Crinion et al., 2012; Kiran et al., 2012; Meinzer et al., 2012; Rapp et al., 2012, a recent series of papers addressing factors relevant to research examining the neural mechanisms of language recovery in aphasia).

The present study examined the effects of training three-argument verbs on production of trained verbs as well as untrained verbs of lesser complexity, i.e., two- and one-argument verbs, testing pre- to post-training performance in verb naming and sentence production conditions. We also

examined the neural correlates of improved verb production using fMRI, comparing pre- to post-training activation patterns to those derived from a healthy age-matched cohort performing a verb naming task. We predicted that treated, but not untreated (control) participants, would show improved three-argument verb naming as well as production of sentences with correct argument structure. In addition, following CATE, we predicted generalization to less complex verbs. We also surmised that, if training results in improved verb naming and sentence production, this should be reflected in changes in patterns of neural activation, anticipating treatment-induced recruitment of spared left and right perisylvian tissue associated with verb and verb argument structure processing in healthy individuals [i.e., the posterior perisylvian regions bilaterally, in particular the angular and supramarginal gyri and the superior parietal region, as well as left frontal regions (i.e., the pars triangularis and pars opercularis)].

2. Method

2.1. Examining the behavioral effects of treatment

2.1.1. Participants

Eight participants with agrammatic aphasia were recruited from the greater Chicago area. Four were assigned to the treatment group (T1–T4) (3 females) and four served as controls (C1–C4) (1 female). Although participants in the treatment group were younger, there was no significant difference between groups for age (treatment group age range 46–60 years, $M = 55$ years; control group age range 57–79 years, $M = 67$ years, Mann–Whitney $Z = -1.162$, $p = .245$). All aphasic participants were right-handed, monolingual English speakers, had normal or corrected-to-normal vision, hearing within normal limits, and at least a high school education. None reported any history of psychiatric, developmental speech-language, learning or neurological disorders other than stroke. All participants suffered a single, LH stroke, and were enrolled in the study at least 1 year post-stroke (control group: 1.5–15 years, $M = 6.8$ years; treatment group: 1.5–11 years, $M = 4.5$ years). All participants provided informed consent, and the study was approved by the Institutional Review Board at Northwestern University.

Structural magnetic resonance (MR) (T1-weighted) scans obtained for participants in the treatment group, who also participated in the imaging portion of the study, and for two of the participants in the control group (C3 and C4)² revealed differences in lesion size and localization in the LH across patients. Selected T1 slices are shown in Fig. 1 for these participants. Lesions were outlined and measured using MRIcro. Lesion size measurement was carried out independently by two investigators, with reliability greater than 95%. We also conducted a region of interest (ROI) analysis for participants T1–T4 to examine the extent of lesioned tissue in the left IFG [BA 44 and 45], middle and superior temporal gyri [MTG (BA 21)

² C1 was unable to participate in MR scanning for medical reasons; scanning was attempted with C2, however, she was unable to tolerate the procedure.

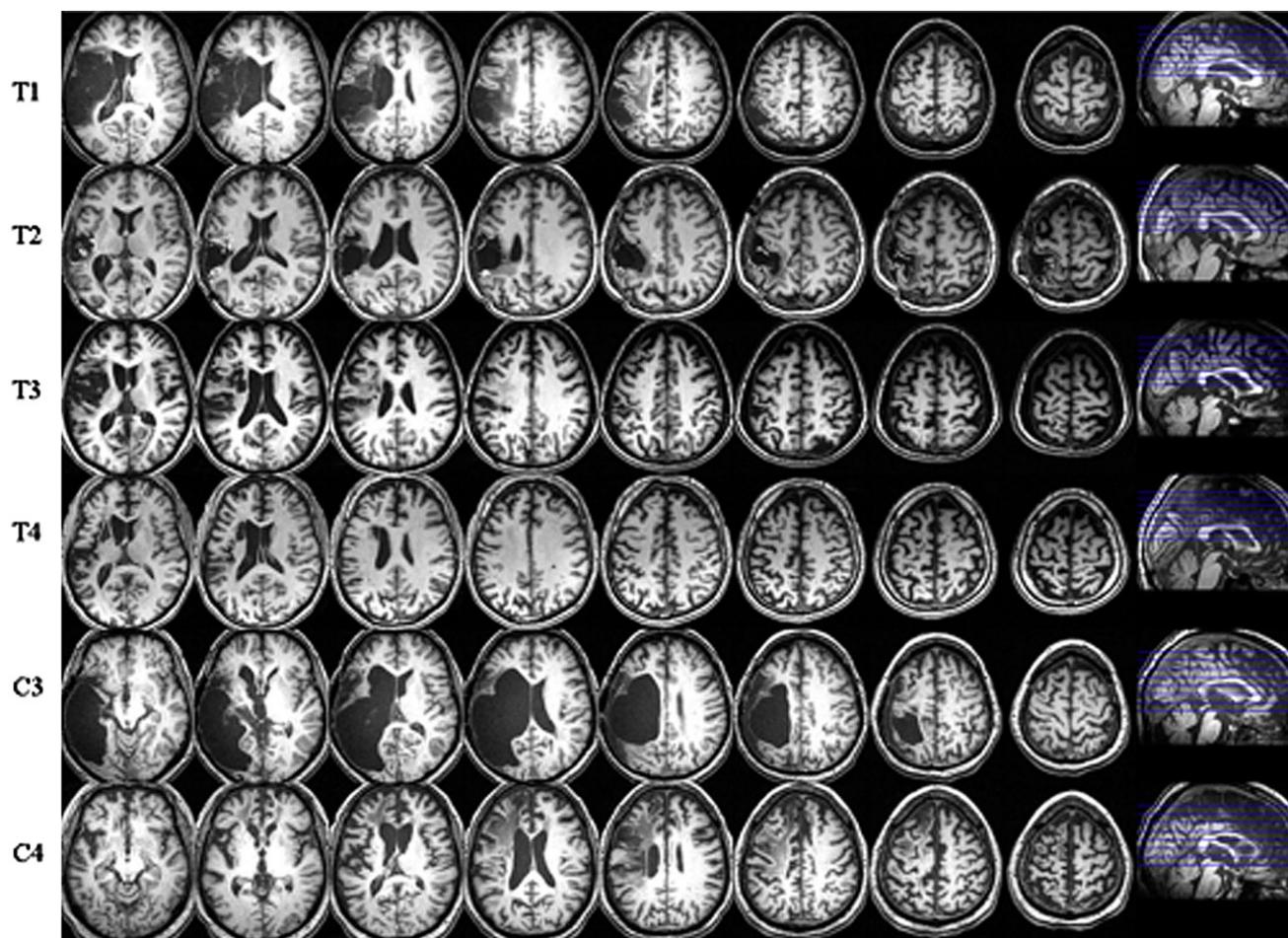


Fig. 1 – T1 scans of the four participants in the treatment group and two of the participants in the control group.

and STG (BA 22), respectively], angular (AG, BA 39) and supramarginal (SMG, BA 40) gyri and superior parietal lobule (SPL, BA 7). To do this we imported ROIs from the WFU-Pickatlas v2.4 (Maldjian et al., 2003) and measured the volume of the lesion within each. This number was divided by the volume of each ROI in a standard brain to compute the percent of lesioned tissue in each ROI. The results of this analysis are presented in Table 1. Participant 1 (T1) presented with a large lesion, including both anterior and posterior regions of the brain: parts of the IFG, MTG, STG, AG, and SMG. T2's lesion involved mostly posterior regions, in particular the STG, SMG, and SPL, with the lesion encompassing only 2.1% of tissue in the IFG. T3 presented with a lesion incorporating parts of the IFG, MTG and STG, extending to the SMG, and T4's lesion included the IFG, MTG and STG as well as subcortical tissue.

As for the control participants, C3 evinced a large, mostly posterior lesion, including portions of the inferior, middle and superior temporal gyri as well the pars opercularis; whereas C4's lesion was subcortical, extending to frontal regions, including the premotor cortex and pars opercularis.

2.1.2. Language testing

All participants with aphasia presented with a diagnosis of mild to moderately severe Broca's aphasia based on performance on the Western Aphasia Battery (WAB, Kertesz, 2007), with Aphasia Quotients (AQs) ranging from 54 to 81 ($M = 69.15$), and no significant differences found between groups [Treatment group: $M = 64.8$ ($SD = 9.09$); Control group: $M = 73.5$ ($SD = 6.59$); $Z = -1.732$, $p = .114$ (Mann–Whitney)]. They also showed language patterns

Table 1 – Number of lesioned voxels and percent lesioned tissue of treated participants in selected ROIs.

Participant	IFG (BA 44 & 45)	MTG (BA 21)	STG (BA 22)	AG (BA 39)	SMG (BA 40)	SPL (BA 7)
# Voxels (MNI brain)	38,900	40,340	42,852	3230	6823	6741
T1	18,056 (46.4%)	22,420 (55.6%)	38,789 (90.5%)	468 (14.5%)	6091 (89.3%)	58 (.8%)
T2	824 (2.1%)	587 (1.4%)	14,506 (33.9%)	78 (2.4%)	4573 (67.0%)	2362 (35.0%)
T3	23,684 (60.9%)	14,575 (36.1%)	31,407 (73.3%)	0 (.0%)	581 (8.5%)	0 (.0%)
T4	9121 (23.4%)	11,024 (27.3%)	22,273 (52.0%)	0 (.0%)	0 (.0%)	0 (.0%)

IFG = inferior frontal gyrus; MTG = middle temporal gyrus; STG = superior temporal gyrus; AG = angular gyrus; SMG = supramarginal gyrus.

Table 2 – Baseline performance (percent correct) on language tests for treated (T1–T4) and untreated control participants (C1–C4): NAVS verb naming and verb argument structure production scores by verb type.

Participant	Verb naming				Argument structure production			
	1-Arg	2-Arg	3-Arg	Total	1-Arg	2-Arg	3-Arg	Total
T1	60%	70%	14%	50%	100%	87%	33%	69%
T2	88%	56%	40%	59%	0%	0%	0%	5%
T3	63%	31%	0%	29%	100%	100%	67%	88%
T4	60%	80%	57%	68%	100%	100%	100%	100%
Mean (SD)	67.75% (13.6%)	59.25% (21.3%)	27.75% (25.6%)	51.5% (16.7%)	75% (50.0%)	71.75% (48.2%)	50% (43.1%)	64.25% (44.7%)
C1	0%	40%	58%	33%	100%	87%	67%	85%
C2	60%	50%	29%	46%	100%	93%	100%	98%
C3	100%	100%	100%	100%	100%	80%	75%	85%
C4	100%	70%	86%	85%	100%	100%	83%	94%
Mean (SD)	65% (47.3%)	65% (26.5%)	68.25% (31.5%)	66% (31.7%)	100% (.0%)	90% (8.5%)	81.25% (14.1%)	90.5% (6.6%)

consistent with agrammatism based on administration of the *Northwestern Assessment of Verbs and Sentences (NAVS, Thompson, 2011)*. All showed better comprehension and production of canonical sentences (i.e., actives and subject relatives) compared to noncanonical sentences (i.e., passives and object relatives) [canonical sentence comprehension $M = 77.50\%$ correct, noncanonical sentence comprehension $M = 55.88\%$ correct, $t(7) = 5.0$, $p = .002$; canonical sentence production $M = 45.83\%$ correct, noncanonical sentence production $M = 22.18\%$ correct, $t(14) = 2.097$, $p = .05$]. Additionally, as shown in Table 2, verb naming and argument structure production deficits were prevalent for all participants (verb naming: $M = 58.75\%$ correct; argument structure production: $M = 77.38\%$ correct). Treatment and control participants did not differ significantly on their verb naming or argument structure production scores (verb naming: treatment group $M = 51.5\%$ correct, control group $M = 66\%$ correct, Mann–Whitney $Z = -.577$, $p = .686$; argument structure production: treatment group $M = 64.25\%$ correct, control group $M = 90.50\%$ correct, Mann–Whitney $Z = -.581$, $p = .686$). In addition, a complexity effect was found in the argument structure production task such that sentences with 3 arguments were produced correctly less than those with 2 arguments and 1 argument [1-arg $M = 87.5\%$ correct, 2-arg $M = 80.88\%$ correct, 3-arg $M = 65.63\%$ correct, Friedman test $\chi^2_{(2)} = 7.714$, $p = .021$]. A trend in the same direction was seen in the verb naming condition, such that 3-argument

verbs were named accurately to a lesser degree than 2- and 1-argument verbs [1-arg $M = 66.38\%$ correct, 2-arg $M = 62.13\%$ correct, 3-arg $M = 48.0\%$ correct, Friedman test $\chi^2_{(2)} = 3.714$, $p = .156$]. Importantly, all treated and control participants showed lower accuracy on 3-argument verbs than on both 2- and 1-argument verbs in at least one of the tasks (i.e., verb naming/argument structure production).

Narrative language samples were obtained by asking participants to tell the story of Cinderella and analyzed using a coding system developed by Thompson et al. (1995; also see Thompson et al., 2012). All participants presented with effortful speech, consisting of simple phrases and sentences as well as ungrammatical word strings, with omission of grammatical morphemes. As shown in Table 3, all participants produced verb-less utterances (e.g., ‘Cinderella bye bye’, ‘stepsisters terrible’), with a significant difference between patients and a sample of eight healthy participants on the proportion of verb-less utterances produced (two-tailed t-test, $p < .001$). Aphasic participants also exhibited an argument structure effect in verb production, with most producing no 3-argument verbs in their narratives. Again, a significant difference was found between healthy and aphasic participants, such that aphasic individuals produced fewer 3-argument verbs than controls (two-tailed t-test, $p < .001$). There were no significant differences between the treatment and control aphasic groups on proportion of verb-less utterances produced or use of different verb types (Mann–Whitney, all p 's $> .05$).

Table 3 – Spontaneous speech features of control and treated aphasic participants pre-treatment, and of a sample of healthy speakers ($n = 8$): proportion (percent) verb-less utterances out of the total number of utterances; proportion of 1, 2 and 3-argument verbs out of all verbs produced.

Participant	% Verb-less utterance	% 1-Arg verbs	% 2-Arg verbs	% 3-Arg verbs
T1	51.2%	50%	50%	0%
T2	36%	25%	75%	0%
T3	81.1%	33.3%	66.7%	0%
T4	3%	25.6%	69.2%	5.1%
Mean (SD)	42.8% (32.5%)	33.5% (11.6%)	65.2% (10.7%)	1.3% (2.5%)
C1	33.3%	50%	50%	0%
C2	6.2%	19.7%	78.9%	1.3%
C3	22.6%	77.8%	22.2%	0%
C4	33.3%	0%	100%	0%
Mean (SD)	23.9% (12.8%)	36.9% (34.2%)	62.8% (33.9%)	.3% (.6%)
Mean (all patients) (SD)	33.3% (25.0%)	35.2% (23.7%)	64.0% (23.3%)	.8% (1.8%)
Healthy controls mean (SD)	.5% (1.3%)	31.6% (8.0%)	60.7% (7.8%)	7.7% (3.0%)

Screening for motor speech impairments revealed a mild apraxia of speech for one participant (C1), with no other participants showing any such deficits (either speech apraxia or dysarthria).³ Finally, a reading screening test showed that all participants had good oral reading and reading comprehension for single words and simple phrases.⁴

2.1.3. Materials

For the behavioral portion of the study, 50 verbs were selected: 10 three-argument verbs, 20 two-argument verbs and 20 one-argument verbs (see Appendix A). Verbs by argument structure type were matched for lexical frequency, word length, and animacy of arguments. Action pictures depicting each target verb were developed to elicit verb naming and sentence generation. Pictures were matched for visual complexity such that, in addition to depiction of all verb arguments, one- and two-argument action scenes included additional participants. For each three-argument verb utilized for treatment, two picture stimuli were developed, providing two sentence contexts for each. For example, the verb *give* was trained in the following contexts with corresponding pictures: “the boy is giving the flowers to the woman” and “the boy is giving the apple to the teacher”. The one- and two-argument verbs were used only for generalization testing. All trained and generalization items were arranged into two sets of probes for verb naming and two sets of probes for sentence production. All four sets contained the same 60 items [20 three-argument targets (two picture stimuli for each of 10 trained three-argument verbs), 20 two-argument targets, and 20 one-argument targets], but were designed to prevent participants from always receiving these items in the same order. Within each probe set, items were presented in a pseudorandomized order such that no more than two verbs from the same argument structure category were presented consecutively. Probe stimulus pictures consisted of black and white line drawings of actions presented either on a computer screen (verb naming probes) or on an 8.5 × 11” card (sentence generation probes). Pictures used for verb naming probes contained only line drawings, whereas pictures used for sentence generation probes contained the same line drawings as well as all content nouns printed along the bottom of the card to reduce lexical retrieval difficulties.

2.1.4. Baseline testing

For all participants with aphasia, two sets of behavioral probes (one for verb naming and one for sentence production), each consisting of the same 60 items described above, were administered during a single session. During a second session, two additional sets of behavioral probes (one for verb naming and one for sentence production) were administered. These two baseline probe scores were

averaged to provide an initial measurement of performance on trained and untrained verbs for the treated and control participants. For the verb naming task, pictures were presented on a computer screen, using Superlab (Cedrus Corporation, version 2.0, Phoenix, AZ). Each trial began with a 1000 msec blank screen, followed by a fixation cross presented for a 1000 msec, before the picture was presented. Participants were required to provide a single action word for each target picture within 10 sec and, following this, to press a button to advance to the next trial. To examine sentence production, participants were asked to produce an active sentence for each examiner-presented picture stimulus. For each item, participants were given 15 sec to respond. Both naming and sentence responses were recorded using Praat (Boersma and Weenink, 2005). Naming responses were scored as correct if participants produced the target verb, or a semantically plausible substitution with the same number of arguments as the target (e.g., *donate* for *give*). For the sentence generation task, responses were scored for: (a) production of the verb (or a licit substitution as described above) and (b) argument structure production, scored as correct if all verb arguments were produced in the correct order. Responses with inflection errors and phonological paraphasias were scored as correct.

2.1.5. Verb argument structure training

Treatment, focused only on production of three-argument verbs in sentence contexts, was provided for two 1.5 h sessions per week for a maximum of 20 sessions or until participants reached a criterion of 80% correct production of trained verbs on both verb naming and sentence generation probes for two consecutive sessions. These probe tasks, testing all trained verbs as well as a subset of untrained (two- and one-argument) verbs, were administered prior to each treatment session. Untrained verbs were pseudorandomly distributed across four lists containing five verbs of each type, and participants were tested with one list for verb naming and a different list for sentence production per session. Participants were exposed to each of the untrained items in the probe tasks up to five times throughout treatment and no feedback regarding response accuracy was provided for any items during probe tasks.

During each training session line drawings for three-argument verbs were presented together with a sentence production template consisting of place-holders for all verb arguments and the target verb (see Appendix B). For each item, the examiner first identified all nouns in the picture and the participant was asked to name the action and generate a sentence. Following this, word cards for the target verb and its arguments were presented and thematic role training was provided. For example, for the sentence “the boy is giving the apple to the teacher”, separate word cards (i.e., *is giving*, *the boy*, *the apple*, and *to the teacher*) were presented. The examiner named the verb, the agent, the theme, and the goal, explaining the role of each argument as it was presented, e.g., for the aforementioned target sentence, the examiner explained: “The action is *give*. Giving requires someone doing the giving: *the boy is giving*. Giving also involves something being given: *the boy is giving the apple*. Giving also involves someone being given something: *the boy*

³ Motor speech screening was based on methods espoused by Dabul (1979) and Wertz et al. (1984).

⁴ Reading ability was tested using a task requiring participants to orally read a list of single words and simple phrases used in the treatment protocol. Comprehension was assessed using a written/spoken phrase matching task, using phrases included in the treatment protocol.

is giving the apple to the teacher.” As each argument was named, the examiner placed the corresponding word card in the appropriate template slot. The participant then was asked to point to the verb, agent, theme, and goal (word cards) as named by the examiner and to read the sentence aloud. Next, the word cards were removed from the template slots and scrambled for the participant to rearrange in correct order and read aloud once again. Finally, the word cards were removed from the picture and the participant was asked to produce the target sentence.

2.1.6. Post-treatment testing

Behavioral probes examining verb and sentence production, identical to those used to evaluate baseline performance, were administered within one week following completion of treatment as a post-treatment measurement for treated participants. As in the baseline, two full probes were obtained and scored on separate days (occurring within the same week). For the control participants, scores on two full sets of treatment probes were obtained 10 weeks after baseline testing (the maximum length of treatment) as a final measurement.

2.1.7. Data analysis

For each treated and control participant, the two baseline probe scores for each dependent variable (verb naming, verb production in sentence context, and argument production in sentence context) were averaged to produce a single baseline pre-training score for trained verbs and a single baseline pre-training score for untrained verbs. The same procedure was followed for the two final probe scores for each participant. For each participant, the chi-square test was used to compare pre- to post-training performance, and Cramer's phi effect size was calculated. A paired-samples Wilcoxon test was used to compare pre- to post-training group mean scores and a corresponding effect size r was computed as Z/\sqrt{N} , where Z is the test statistic obtained in the Wilcoxon test, and N is the number of observations.

2.1.8. Reliability

Thirty percent of recorded naming and sentence production probe responses were scored for inter-rater reliability by the second or last author. Across all samples, agreement between the raters was 96% for verb naming, 95% for verb production in sentences and 97% for verb argument structure production in sentences. All disagreements were discussed to achieve 100% agreement.

2.2. Examining the neural mechanisms of treatment-induced recovery

2.2.1. Participants

The four aphasic participants in the treatment group also participated in the fMRI portion of the study, undergoing pre- and post-treatment structural and functional MR scans to evaluate activation patterns associated with action verb naming. Post-treatment fMRI scanning occurred one week after post-treatment behavioral testing. In addition, 13 healthy individuals, recruited from Northwestern University and the surrounding community, were included in the study

to ascertain normal activation patterns (5 females; age 51–69 years, $M = 60.4$ years). There was no significant difference in age between the treatment and control groups (Mann–Whitney, $Z = -1.25$, $p = .211$). All control participants were right-handed, monolingual English speakers with normal or corrected-to-normal vision, and normal hearing. None reported any history of psychiatric, developmental speech-language, language or neurological disorders. All participants provided informed consent, and the study was approved by the Institutional Review Board at Northwestern University.

2.2.2. fMRI scanning procedures

An action (verb) video naming task was utilized in the scanner. The task included 18 two- and 18 one-argument verbs (see Appendix C). The two verb sets were controlled for frequency, number of syllables and body–part association. Each verb was depicted by a 2-sec action video, controlled for visual complexity (i.e., number of objects and people in the scene), and animacy of arguments across verb types (see Den Ouden et al., 2009, for details). Verbs from the different conditions appeared in a randomized order in an event-related design. Stimuli were repeated over two runs, for a total of 36 trials per condition. Each trial lasted 10 sec, beginning with a jittered Inter-stimulus interval between 0 and 2000 msec, which preceded the video presentation onset. The 2000 msec video was then presented, followed by a centered fixation cross presented for the remainder of the 10-sec trial. Participants were thus allowed between 6 and 10 sec for action naming in each trial. The total time of each run was 6:22 min. Participants were instructed to overtly produce the verb depicted in each video clip. Stimuli were presented using Superlab software (Cedrus Corporation, version 2.0, Phoenix, AZ). Verbal responses produced in the scanner were recorded using custom-built software (running under Matlab) for scanner noise reduction. The denoised files then were converted into Praat for analysis of response accuracy and RT.

Scanning was carried out on a Siemens 3T TIM Trio scanner. A T1-weighted anatomical scan was obtained at the start of each protocol with the following parameters: TR = 2300 msec; TE = 3.36 msec; flip angle = 9°; image matrix = 224 × 256; FOV = 256; voxel size = 1 × 1 × 1 mm. During experimental runs, functional volumes with BOLD contrast were obtained using gradient echo-planar imaging sequences (TR = 2000 msec; TE = 30 msec; flip angle = 80°; matrix size = 64 × 64; FOV = 220 mm; voxel size = 3.44 × 3.44 × 3 mm; 32 slices). To allow for image saturation, the first 10 volumes of each run were discarded.

2.2.3. Data analysis

Data preprocessing and statistical analysis were performed using SPM5 (<http://www.fil.ion.ucl.ac.uk/spm>). Functional scans were corrected for slice-acquisition timing and realigned to a mean functional volume. The anatomical volume was coregistered to the mean image and normalized to the Montreal Neurological Institute (MNI) 152-subject template brain (ICBM, NIH P-20 project). Functional volumes were then normalized using the same transformation, resliced at a resolution of 3 × 3 × 3 mm, and spatially smoothed using a 6 mm

(full-width, half-maximum) isotropic Gaussian kernel. Effects of global signal were removed from the functional time-series using the method described by Macey et al. (2004). In the first-level analysis, a high-pass filter of 256 sec was used to eliminate scanner drift. For each run, six movement parameters obtained during pre-processing were entered as regressors.

Response accuracy and RTs were calculated using Praat files for each participant. Accuracy was scored according to the same criteria used for verb naming in the behavioral task (see Section 2.1.4). RT was measured from the onset of the stimulus video to the onset of production of target phonological material (e.g., for the response ‘uh bite, fishing’ following a video of fishing, RT was calculated to the onset of ‘fishing’). Incorrect (e.g., ‘fall’ for kick, ‘water’ for drink) and abandoned responses were not scored for RT. For reliability purposes an independent coder scored all responses for accuracy and 50% of responses for each participant for RT. Point-to-point inter-rater agreement for accuracy was 84.7 and for RT was 94.2. Disagreements were resolved by the first author.

For healthy control participants, activation maps for the main effects of two-argument and one-argument verb naming were entered into a second-level factorial model. Second-level statistics were evaluated at a voxelwise significance threshold of $p < .05$, corrected for multiple comparisons per false discovery rate (FDR; Benjamini and Hochberg, 1995; Genovese et al., 2002), with a minimum cluster size of 15 contiguous voxels (405 mm^3).

The aphasic individuals’ fMRI data were analyzed on a single-subject basis. In order to model the variance in response times to items between as well as within participants, the response times of each participant for each individual event were used as event durations. Naming accuracy was modeled as a parametric value with each individual trial.⁵ At the first level, each participant’s pre- and post-treatment data were combined into a single model. The main effects of pre- versus post-treatment as well as the interaction between two- and one-argument conditions and pre- versus post-treatment activation were then evaluated at an uncorrected voxelwise significance threshold of $p < .001$, with a cluster threshold of 15 contiguous voxels (405 mm^3).

In addition, we performed ROI analyses on the aphasic participant’s pre- versus post-treatment main effect changes, in six LH regions and their contralateral homologs: the IFG (BAs 44 and 45), the superior and middle temporal gyri (BA 22, 21), the angular and supramarginal gyri (BA 39, 40), and the superior parietal cortex (BA 7). As in the lesion volume analysis described in Section 2.1.1, ROIs were based on the WFU-Pickatlas toolbox, version 2.0 (Maldjian et al., 2003). For these analyses, we used the participant’s locally specific hemodynamic response functions (HRFs), as obtained through a long-trial version of the naming task, following the same methodology as in Thompson et al. (2010b) (also see Bonakdarpour et al., 2007). In a 10-min run, participants performed the video action naming task, with a fixed stimulus onset asynchrony of 30 sec and volume acquisition (TR = 2 sec) locked to the events,

⁵ For the healthy control participants, this was deemed irrelevant, as their naming accuracy was high and response time was consistent and fast.

allowing us to measure the BOLD response up to 30 sec post-stimulus, at a frequency of .5 Hz (1 data point every 2 sec). The long-trial data were analyzed using Brain Voyager (QX 1.4, Maastricht, The Netherlands). HRF latency maps were formed using linear correlation lag analysis of the stimulation onsets and the time series on a voxel by voxel basis. Within a suprathresholded ROI, a stimulus-locked average formed the HRF curve for that particular cluster. For each participant, we made alternative pre- versus post-treatment data models in SPM, using Finite Impulse Response (FIR) functions, instead of modeling the responses as canonical HRFs. In order to keep these models relatively simple, we did not model RTs or accuracy. For each ROI, the obtained hemodynamic response patterns from 0 to 20 sec post-stimulus were fed as weights into relevant FIR-modeled post-event time bins (10 time bins of 2 sec each) in the statistical model in SPM. Specific contrasts were compared across conditions. These t-tests were evaluated at an alpha level of $p < .05$, FDR corrected for multiple comparisons.

3. Results

3.1. Response to treatment: behavioral performance

All participants who received treatment reached criterion within 20 sessions (range 17–20), with all showing improved naming of trained three-argument verbs, production of these verbs in sentences and production of their arguments in sentences from pre- and post-treatment. The untreated controls, however, showed stable performance across test points in all conditions for three-argument verbs (see Fig. 2; individual pre- and post-training scores on all tasks are provided in Appendix D). For the treated participants pre-treatment verb naming ranged from 20% to 40% correct ($M = 27\%$ correct, $SD = 9\%$), with post-treatment scores ranging from 60% to 78% correct ($M = 72\%$ correct, $SD = 8.1$). In contrast, verb naming performance for the controls at baseline ranged from 25% to 70% correct ($M = 46\%$ correct, $SD = 21.7$), with little change noted in follow-up testing (range = 13%–75%, $M = 48\%$ correct, $SD = 26.6$).

Similar improvement in sentence production as well as verb argument structure was noted for the treated, but not for the control participants. Verb production in sentences for the treated participants pre-treatment ranged from 8% to 38% correct ($M = 24\%$ correct, $SD = 14$) and post-treatment ranged from 53% to 78% correct ($M = 65\%$ correct, $SD = 11$). For the controls verb production in sentences at baseline ranged from 15% to 70% correct ($M = 56\%$ correct, $SD = 28$), with similar scores at follow-up (range = 10%–80% correct, $M = 43\%$ correct, $SD = 35$). Finally, for verb argument structure production the treated participants showed pre-treatment performance ranging from 0% to 33% correct ($M = 12\%$ correct, $SD = 15$) and post-treatment scores ranging from 38% to 68% correct ($M = 54\%$ correct, $SD = 12\%$), whereas baseline performance for the control participants ranged from 3% to 65% correct ($M = 45\%$ correct, $SD = 29\%$), with follow-up scores ranging from 5% to 55% correct ($M = 23\%$ correct, $SD = 24$).

Statistical analysis examining pre- to post-treatment change in production of three-argument verbs for the treatment group approached significance for all conditions (verb

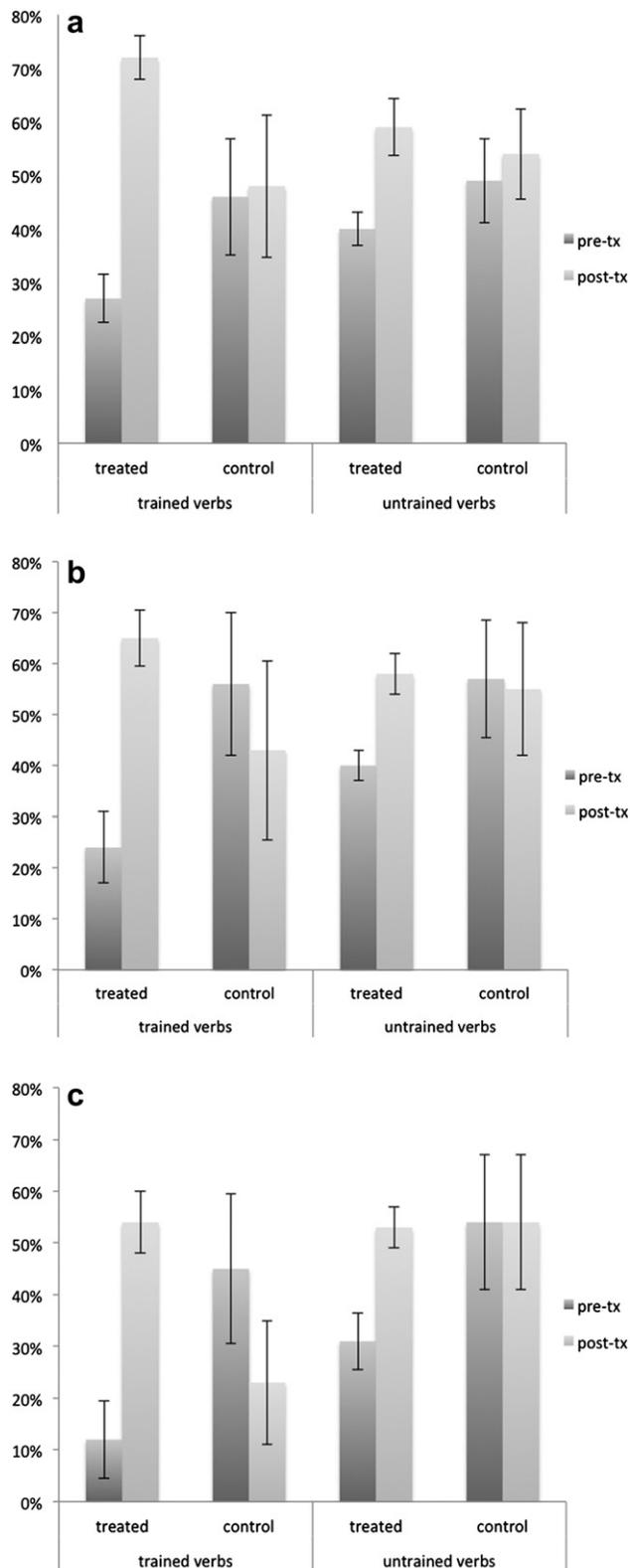


Fig. 2 – Mean percent correct production pre- and post-treatment (tx) for treated and control participant groups on (a) verb naming (b) verb production in sentences, and (c) argument structure production. Error bars represent $\pm 1SE$.

naming, verb production in sentences, and argument structure production in sentences) ($Z = -1.826, p = .068$ for all), with a medium to large effect size ($r = .646$ for all), whereas control participants showed no significant difference between baseline and follow-up probes (verb naming: $Z = -.730, p = .465$; verb production in sentences: $Z = -.816, p = .414$; verb argument structure production: $Z = -1.461, p = .144$) and small or negative effect sizes were found across conditions (verb naming: $r = .258$; verb production in sentences: $r = -.288$; verb argument structure production: $r = -.516$).⁶

This same pattern was found for untrained two- and one-argument verbs. In verb naming, treated participants showed an increase from a mean of 40% correct (range: 36%–49%, $SD = 6.2\%$) to a mean of 59% correct (range: 45%–70%, $SD = 10.6\%$), whereas untreated participants were at 49% (range: 28%–61%, $SD = 15.7\%$) pre-treatment, and 54% (range: 38%–75%, $SD = 16.9\%$) post-treatment. For verb production in sentences, trained participants showed a similar increase from 40% correct pre-treatment (range: 36%–48%, $SD = 6\%$) to 58% correct post-treatment (range: 48%–61%, $SD = 8\%$), while control participants showed a slight decrease, from 57% correct pre-treatment (range: 23%–73%, $SD = 23\%$) to 55% correct post-treatment (range: 26%–83%, $SD = 26\%$). For argument structure production, treated participants exhibited an increase from 31% (range: 15%–39%, $SD = 11\%$) to 53% correct (range: 46%–61%, $SD = 8\%$) following training, whereas control participants' performance did not change from pre- (mean: 54% correct, range: 15%–71%, $SD = 26\%$) to post-treatment (mean: 54% correct, range: 25%–83%, $SD = 26\%$).

Statistical analyses showed that treated participants demonstrated a trend toward significant pre- to post-treatment improvement ($Z = -1.826, p = .068, r = .646$ for all conditions). Control participants, in contrast, showed no significant difference in performance across probes ($Z = -.730, p = .465, r = .258$ in verb naming and verb production in sentences conditions; $Z = -.365, p = .715, r = .12$ for argument structure production).

3.2. fMRI results

3.2.1. Behavioral performance

Data from the 13 healthy participants indicated high verb naming accuracy (mean correct responses for one-argument verbs: 90%, for two-argument verbs: 99%), with two-argument verbs produced more accurately than one-argument verbs [$t(463.338) = -5.175, p < .001, df$ corrected for unequal variances]. RT means for naming one-argument and two-argument verbs were 2022 and 1915 msec, respectively, with no significant RT difference between the two verb types [$t(710) = 1.445, p = .149$].

For the treated participants improved accuracy from pre- to post-treatment was noted for all participants, for both one- and

⁶ The negative effect sizes were mainly driven by C2, who showed decreases in performance from baseline to follow-up, due to a deterioration in general health. However, even after taking C2's scores out, percent correct verb production in sentences by control participants did not show any increase (mean correct pre-treatment = 53.3%, mean correct post-treatment = 53.3%), and percent correct argument structure production dropped (mean correct pre-treatment = 37.6%, mean correct post-treatment = 28.3%).

two-argument verbs [one-argument: pre-treatment range = 29.7%–54.2% correct [$M = 41.3\%$ correct], post-treatment range = 39.5%–63.9% correct [$M = 50.05\%$ correct], $t(3) = -8.064$, $p = .004$; two-argument verbs: pre-treatment range = 32.2%–55.6% correct [$M = 38.75\%$ correct], post-treatment range = 39.8%–58.3% correct [$M = 46.47\%$ correct], $t(3) = -3.672$, $p = .035$]. The RT data, however, showed an increase from pre- to post-training for all participants, albeit not significant for either verb type [one-argument: pre-treatment range 1182.8–4381.5 msec, $M = 2328.6$ msec, post-treatment, range 1191–4235.3 msec, $M = 2922.4$ msec, $t(3) = -1.238$, $p = .304$; two-argument: pre-treatment, range 1111–3953.1 msec, $M = 2277.6$ msec, post-treatment, range 1215.6–4937.3 msec, $M = 3130.2$ msec, $t(3) = -2.133$, $p = .123$].⁷

3.2.2. Neural activation patterns

3.2.2.1. CONTROL PARTICIPANTS. Results for the contrast of two-argument versus one-argument action naming for the healthy volunteers are presented in Table 4. At the threshold of $p < .05$ (FDR corrected), large clusters of increased activation associated with two-argument verbs, relative to one-argument verbs, were found in the LH postcentral gyrus and inferior parietal lobule and in the RH inferior frontal gyrus (BA 9). Smaller clusters were found in the LH insula and in RH Heschl's gyrus and thalamus. The opposite contrast, of one-argument versus two-argument verb naming, showed no significant regions of activation.

3.2.2.2. AGRAMMATIC PARTICIPANTS. Pre- versus post-treatment main effects and transitivity effects (i.e., two-argument verbs contrasted with one-argument verbs) for each agrammatic aphasic participant, as modeled with a canonical HRF over the whole brain are shown in Tables 5 and 6, respectively, and in Fig. 3. In addition, Table 7 presents the results of the ROI analysis of differential main effects, modeled with regionally specific true HRFs for each participant.

3.2.2.2.1. PARTICIPANT 1. Main effects analyses for Participant 1 using a canonical HRF [with time-to-peak (TTP) of approximately 6 sec] showed no significant regions of cortical activation for the pre- versus post-treatment contrast. However, significant post- versus pre-treatment upregulation was found in the LH angular gyrus. For the two-argument versus one-argument verb contrast no regions of activation were found at pre-treatment or at post-treatment. When we examined activation patterns using Participant 1's true HRF in the ROI analysis, greater pre- versus post-treatment activation

⁷ A reviewer of this paper suggested that the fact that some of the stimuli in the fMRI task were also included in the behavioral probes affected participants' performance on these items. Out of the transitive verbs in the fMRI task, 10 appeared in the probes and 8 did not. We compared these two verb sets on two measures: i) change in accuracy pre- to post-treatment; ii) change in RT pre- to post-treatment. In both cases, the difference between sets was not statistically significant ($p = .152$ and $p = .294$, respectively). These comparisons suggest that exposure to some actions in the probes did not significantly alter participants' ability to name them. This is expected since during the probes, participants were only exposed to pictures depicting these verbs, and received no feedback on their naming.

Table 4 – Suprathreshold activation for the naming of 2-argument over 1-argument verbs, in the healthy control group [FDR corrected, $p < .05$, cluster size (k) > 25].

Hem.	Cluster extent	BAs	k	Peak MNI coordinates			t-Max
				x	y	z	
L	Insula	13	49	-42	-3	9	4.83
L	Postcentral gyrus, inferior parietal lobule	2, 40	178	-57	-21	30	6.88
R	Inferior frontal gyrus	9	173	48	0	24	6.18
R	Transverse temporal gyrus	41	44	33	-30	6	5.34
R	Thalamus		52	21	-18	6	5.03

was found in the LH Broca's area and the middle temporal gyrus. Greater post- versus pre-treatment activation also was found in the LH Broca's area. Notably, the HRF TTP was delayed (10–12 sec) in Broca's area, but not in posterior brain regions.

3.2.2.2.2. PARTICIPANT 2. For Participant 2 the main effects analysis showed pre- over post-treatment activation only in the primary visual cortex (BA 18) and no differential activation in any brain regions for two- compared to one-argument verbs. However, greater post- compared to pre-treatment main effects were found with significant activation in motor and sensory cortices bilaterally, and, notably, in the RH superior temporal gyrus (BA 22), as well as the fusiform gyrus (BA 37). For the contrast between verb types, i.e., two-argument > one-argument verb naming, greater post-compared to pre-treatment activation was found in the LH insula. The HRF analysis showed TTP within the normal range (ranging from 4 to 8 sec across regions) for Participant 2. Using these TTP, the ROI analyses reveal greater pre-compared to post-treatment activation in the LH angular gyrus and in the supramarginal gyrus, bilaterally, but no significant post- versus pre-treatment activation.

3.2.2.2.3. PARTICIPANT 3. For Participant 3, main effects analysis in the pre-treatment > post-treatment contrast showed bilateral activation in the inferior frontal gyrus and superior parietal regions, as well as in superior frontal gyrus and motor cortex in the LH, and middle frontal gyrus, fusiform gyrus and visual cortex in the RH. Post-treatment contrasted with pre-treatment showed upregulation in the middle and superior temporal gyri as well as the angular gyri and superior parietal cortex bilaterally, with additional activations in LH supramarginal gyrus and RH sensory and motor cortices. The interaction between time (i.e., pre- vs post-treatment) and two-argument over one-argument verb naming showed activation in the LH insula and in RH premotor cortex (BA 6). Post-compared to pre-treatment, the two- over one-argument contrast showed activation in RH superior temporal gyrus. The ROI analyses further revealed activation post- compared to pre-treatment in the LH supramarginal gyrus, using the participant's specific, normal-like TTP of 6–8 sec.

3.2.2.2.4. PARTICIPANT 4. Finally, Participant 4 showed no regions of significant activation prior to treatment (i.e., in the pre-

Table 5 – Significant regions of pre- versus post-treatment activation (main effects) for agrammatic participants as measured using a canonical HRF, $p < .001$, uncorrected.

Participant	Contrast	L/R	Region	BA	k	Peak MNI coordinates			t-Max
						x	y	z	
T1	Pre > post		None	N/A					
	Post > pre	L	Angular Gyrus	39	20	-42	-69	15	3.77
T2	Pre > post	R	Cuneus	18	34	42	-87	-9	4.77
	Post > pre	L	Precentral gyrus	6	92	-39	-6	54	5.04
			Postcentral gyrus	3		-39	-21	54	4.31
		R	Precentral gyrus	4	432	18	-27	60	6.99
				6		36	-9	57	4.58
			Postcentral gyrus	3		42	-27	60	5.36
		R	STG	22	24	60	-45	18	4.31
		R	Fusiform gyrus	37	42	36	-63	-6	4.72
T3	Pre > post	L	IFG	44		-51	15	18	5.59
				46		-42	39	12	5.22
		L	SFG	10	652	-24	48	21	5.7
		L	Precentral gyrus	4	41	-51	-12	39	4.94
		R	IFG	45	27	42	21	15	4.87
				46		51	30	21	3.20
		R	Anterior cing. MFG	24 10	408	12 9	30 54	15 12	5.36 5.03
		R	Fusiform gyrus	37	59	45	-63	3	4.28
		R	Cuneus	18	102	33	-54	6	4.23
		BL	SPL	7	696	18	-57	60	5.21
	Post > pre	L	MTG	21	24	-63	-54	-3	3.98
				20		-57	-45	-15	3.71
		L	STG	22	79	-63	-42	6	4.31
			SMG	40		-63	-36	24	4.24
		L	AG	39	129	-42	-60	27	4.87
			Cuneus	19		-30	-84	36	5.36
		R	SFG	8	26	21	24	51	5.17
		R	Precentral gyrus	6	21	42	-9	54	4.45
		R	AG	39	64	45	-75	33	5.71
		R	Postcentral gyrus	43	646	66	-18	15	6.66
			MTG	21		54	-12	-15	6.54
			STG	22		60	6	3	5.96
		BL	SPL	7	165	-3	-81	42	4.95
T4	Pre > post		None	N/A					
	Post > pre	L	SMG	40	68	-39	-57	54	3.96
					41	-57	-36	36	3.86
			SPL	7		-36	-72	45	4.55

Note. L = left; R = right; BL = bilateral; MFG = middle frontal gyrus; STG = superior temporal gyrus; SFG = superior frontal gyrus; cing. = cingulate; IFG = inferior frontal gyrus; MTG = middle temporal gyrus; SMG = supramarginal gyrus.

treatment vs post-treatment contrast), however, the opposite contrast (post-treatment vs pre-treatment) revealed greater activation in the LH supramarginal gyrus and superior parietal cortex. In addition, for the two- versus one-argument verb naming contrast, greater post- compared to pre-treatment activation was found in the LH sensory cortex and cuneus, and RH motor cortex, as well as superior temporal and Heschl's gyri. ROI analyses revealed upregulation for main effects in the LH supramarginal gyrus and superior parietal cortex, with TTP between 10 and 12 sec.

4. Discussion

Results of the present study suggest that verb deficits can be improved using treatment that explicitly exploits the argument structure of verbs. All treated participants improved in their ability to produce trained, three-argument verbs both in

the verb naming task and in sentence contexts following treatment, whereas untreated participants showed no improvements in either task. Moreover, training of verbs with more complex argument structure resulted in generalized production of untrained verbs with less complex argument structures, both as singletons and in sentences. This finding is in line with CATE, indicating that generalization occurs from more complex to less complex, linguistically related structures. Importantly, the generalized verb naming improvement noted during treatment also was noted during performance of the fMRI task in that all participants showed increased naming accuracy on post-treatment scans for both one- and two-argument verbs. Notably, accuracy in the scanner was poorer than that seen on behavioral (out-of-scanner) probes. This is likely due to scanner noise (even though we used custom software to reduce noise) and other variables that are not at play in clinical testing. Interestingly, RT was slower on post-treatment compared to pre-treatment scans, increasing from

Table 6 – Differential activation patterns for interactions of argument structure complexity (two-argument > one-argument) with scan time, for each agrammatic participant, as measured using a canonical HRF, with TTP of approximately 6 sec.

Participant	Contrast	L/R	Region	BA	k	Peak MNI coordinates			t-Max
						x	y	z	
T1	Pre > post		None	NA					
	Post > pre		None	NA					
T2	Pre > post		None	NA					
	Post > pre	L	Insula	13	23	–42	15	18	4.22
T3	Pre > post	L	Insula	13	20	–48	–15	15	3.87
	Post > pre	R	MFG	6	34	27	–6	48	3.72
		R	STG	22	93	54	–30	3	5.02
T4	Pre > post		None	NA					
	Post > pre	L	Postcentral gyrus	2	83	–36	–39	66	5.09
				3		–24	–33	75	4.74
		L	Cuneus	19	22	–39	–81	9	3.85
		R	Precentral gyrus	6	44	30	–24	75	4.52
		R	STG	22	21	51	12	–6	4.83
		R	Heschl's gyrus	42	234	66	–18	12	6.01

Note. L = left; R = right; MFG = middle frontal gyrus; STG = superior temporal gyrus.

a mean of 2329 msec to a mean of 2966 for one-argument verbs and from a mean of 2527 to a mean of 3134 for two-argument verbs from pre- to post-treatment scans. Although not significant, these findings suggest a trade-off between accuracy and response time, with accurate responding requiring greater processing time. Additionally, the increase in post-training RT may have resulted from the treatment itself, which emphasized production of the verb as well as its arguments. This perhaps enhanced verb argument structure encoding during naming in the scanner task, hence, increasing RT.⁸

As noted in the introduction, treatments attempting to improve verb production by providing phonological or semantic cues have not resulted in generalization from trained to untrained verbs. Phonological cuing strategies may not promote generalization because the source of verb naming deficits in agrammatic aphasia is likely not related to a phonological access deficit (see Bastiaanse and van Zonnefeld, 2004; Luzzatti et al., 2002; Thompson, 2003 and others). Semantic cuing strategies, on the other hand, may not result in generalization because of fundamental differences in the organization of the lexicon in the verbal domain compared to the nominal domain. Whereas for nouns, items form semantic categories with hierarchical structure and shared features (e.g., *terrier, golden retriever* → *dog* → *animal*); for verbs, such strong featural overlap between lexical entries does not exist (see Huttenlocher and Lui, 1979; Kable et al., 2002; Kemmerer and Gonzalez Castillo, 2010; Levin, 1993; Van Valin, 2005 and many others). Instead, verb categories are formed based on their event structure and argument structure. Hence, exploiting the argument structure information associated with verbs during treatment leads to significant generalization.

One caution with regard to the interpretation of the present data pertains to the fact that although the control

participants were tested on two occasions separated by the treatment time provided to the experimental participants, they received no intervention between test points, whereas participants in the treatment group received direct one-on-one training. It is possible, therefore, that the improvement exhibited by the treated participants resulted from social interaction rather than from the treatment provided, per se. We point out, however, that previous verb treatment studies (i.e., those employing phonological and/or semantic treatment), also providing several hours of training for aphasic speakers, have not shown improvements in verb and sentence production and generalization, as in the current study. Thus, it is unlikely that the treatment effects found in the present experiment resulted solely from social interaction. Nevertheless, future studies should provide control participants with placebo treatment to rule out this possibility.

There also were some age and gender differences between the experimental and control participants. On average the controls were older (although group differences for age were not significant) and there were fewer females in the control group. However, the relationship between age and response to treatment in aphasia is unclear, with most studies suggesting that age has no significant effect on recovery (Kertesz and McCabe, 1977; Lazar et al., 2008; Pedersen et al., 1995), some showing an effect (Laska et al., 2001), and others presenting mixed findings (Lenderm et al., 1988). It is also important to note that participants in both groups were well into the chronic stage of aphasia. According to Robey's (1998) meta-analysis of clinical outcomes in aphasia treatment, spontaneous language improvement during the chronic stage is likely to be very small (average effect size = .05). We also point out that several studies have found no gender differences in aphasia recovery (Lazar and Antonello, 2008; Lazar et al., 2008; Pedersen et al., 1995) and in studies showing a difference, greater improvement for males has been noted (Holland et al., 1989).

Turning next to discuss the imaging data, the healthy speakers showed several activation clusters in response to naming two-argument versus one-argument verbs. A large

⁸ We thank an anonymous reviewer for suggesting this possibility.

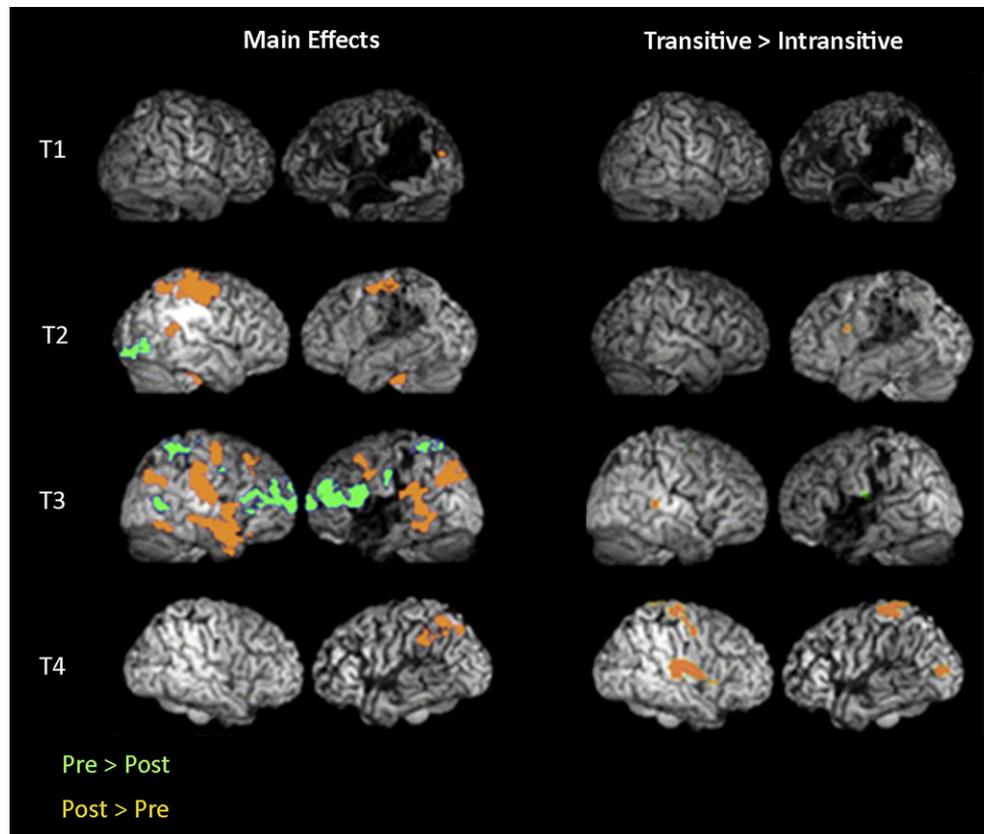


Fig. 3 – Pre- versus post-training fMRI results for the agrammatic participants ($p < .001$, uncorrected, $k \geq 15$). Shown on the left are main effects of verb processing; shown on the right are effects of the two-argument > one-argument contrast. Note: Participant 1 had no pre > post or post > pre activations for the two-argument > one-argument comparison.

cluster was found in the left hemisphere postcentral and supramarginal gyri, similar to the one found in Den Ouden et al. (2009) for the same contrast. Supramarginal gyrus activation in response to complex argument structures was found also in Thompson et al. (2007, 2010a) and in Ben-Shachar et al. (2003). These results are in line with the many findings showing

involvement of the inferior parietal lobule, particularly in the left hemisphere, in tasks implicating semantic computation (Binder et al., 2009). We also found bilateral IFG activations (with a peak in BA 9 in the right hemisphere, and in the insula in the left hemisphere), which also parallel the activation found for video naming of two-argument over one-argument verbs in

Table 7 – Results of ROI analysis, conducted by modeling each participant's regionally specific true HRF, for pre- versus post training differential activation for verb naming (FDR corrected; $p < .05$). Also given is the regionally specific TTP of the hemodynamic response curve.

Participant	Contrast	L/R	ROI	TTP	k	Peak MNI coordinates			t-Max
						x	y	z	
T1	Pre > post	L	Broca's area	10–12 sec ^a	6	–36	18	33	4.61
	Post > pre	L	Middle temporal gyrus	6–8 sec	3	–51	–30	–9	5.01
		L	Broca's area	10–12 sec ^a	5	–39	36	3	3.79
T2	Pre > post	L	Supramarginal gyrus	6–8 sec	7	–54	–51	33	5.03
	Post > pre	L	Angular gyrus	6–8 sec	2	–54	–51	33	4.80
		R	Supramarginal gyrus	4–6 sec	7	48	–39	27	4.05
T3	Pre > post		None						
T4	Post > pre	L	Supramarginal gyrus	6–8 sec	11	–45	–33	33	3.97
	Pre > post		None						
T4	Post > pre	L	Supramarginal gyrus	10–12 sec	36	–51	–48	27	3.78
		L	Superior parietal cortex	10–12 sec	26	–33	–63	60	4.12
					8	–30	–66	45	3.45
					3	–12	–72	39	3.26

a HRF with negative slope.

Den Ouden et al. (2009), where a small cluster of activation was found in the left IFG, and a bigger one in the right homolog of this region. The healthy participants also showed a cluster of activation in the right thalamus. The precise role of the thalamus in language processing is not clear (e.g., Crosson and Haaland, 2003). However, a number of studies suggest that it is involved in semantic processing (Raymer et al., 1997; Kraut et al., 2002). Further, the thalamus may be involved in word generation (Crosson et al., 2003), in line with our findings.

With regard to the effects of treatment on brain activation in the participants with aphasia, we first point out the variability in lesion volume and location found across patients in the study. Importantly, our inclusionary criteria were based solely on the participants' linguistic profile, i.e., all presented with a diagnosis of aphasia with agrammatism, rather than on their lesion properties. Lesion analyses showed that all evinced damage in BA 44/45 and the posterior perisylvian region. T4's lesion also extended subcortically, a feature that has been found in other patients with Broca's aphasia, including those studied by Broca (see Dronkers et al., 2007; Signoret et al., 1984). This heterogeneity is not surprising given the growing body of evidence showing that agrammatism is not necessarily correlated with lesions in Broca's area (see e.g., Caplan et al., 1996; Dronkers et al., 1992; Kaan and Swaab, 2002; also see Thompson et al., 2010b; Kiehl et al., 2012, for sets of patients with agrammatic aphasia and associated lesions).

The aphasic participants also showed considerable variability in neural activation patterns, as has been found in previous studies of treatment-induced language recovery. This variability undoubtedly stems, at least in part, from differences in lesion parameters as discussed above as well as related factors, including regions of hypoperfused tissue (see Thompson et al., 2010b). A limitation of the present study is that the aphasic control participants did not participate in the imaging portion of the experiment. It is thus not possible to be certain that the observed changes in activation from pre- to post-treatment were associated with treatment-induced improvement. Rather it is possible that these changes reflect variability associated with repeat scans. We note, however, that previous studies examining activation across repeat scans in healthy speakers have found relatively stable patterns (Fridriksson et al., 2006; Meltzer et al., 2009). Meltzer et al. (2009), for example, reported fMRI signal stability across multiple sessions in terms of regions of activation, but with reduced activation across scans, possibly related to repetition effects. Based on these findings, Meltzer et al. (2009) suggested that although reduced activation across scans in patients may be difficult to interpret, "activation increases occurring over a time course in aphasia recovery are quite likely to reflect genuine reorganization and neural recruitment" (p. 753). Hence, it is likely that increases in signal from pre- to post-treatment in our participants reflected changes in their language ability.

We also observed several interesting shifts in activation from pre- to post-treatment, which were consistent across participants, and in line with the activations found for this task in our healthy control participants and in Den Ouden et al. (2009). All participants who received treatment in the present study showed greater activation on post-treatment compared to pre-treatment scans, both for verb naming in general and for the two- versus one-argument verb contrast.

For all treated participants, main effects analysis revealed upregulation in the angular, supramarginal and/or superior posterior temporal gyri. For three participants (T1, T3, T4) activation was perilesional. The analysis also showed post-treatment recruitment of the superior parietal cortex for T3 and T4 and sensory-motor cortices for T2 and T3. As mentioned above, posterior perisylvian regions have been implicated in the processing of complex argument structures in several studies (Thompson et al., 2007, 2010a; Meltzer-Asscher et al., 2012), and we also found activation in this region in the healthy controls in the present study. SPL activation was also found in Den Ouden et al. (2009) and in Thompson et al. (2010b), in a study examining the effects of syntactic treatment highlighting verb properties. The superior parietal cortex also has been found to play a role in working memory in general (Koenigs et al., 2009) and verbal memory in particular (Tsukiura et al., 2001), suggesting, as noted by Thompson et al. (2010b), that participants relied more heavily on verbal working memory post- relative to pre-treatment. Notably, the left parietal lobe activation, associated with verb processing in our study as well as others (see Silveri et al., 2003; Tettamanti et al., 2005), also may be associated with action recognition (Shmuelof and Zohary, 2005; Buccino et al., 2001). With regard to post-treatment activation in sensory-motor areas in two of our participants, Den Ouden et al. (2009) also found motor activation in their verb naming task, suggesting possible recruitment of the mirror neuron system (Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004).

Interestingly, for the interaction between treatment and argument structure, treatment led to greater activation post-compared to pre-treatment in three of four participants (T2, T3, T4). Whereas for T2, upregulation was noted only in the insula, participants T3 and T4, once again, showed upregulation in the right superior temporal gyrus, implicated in previous studies with healthy controls for processing argument structure (Ben-Shachar et al., 2003; Den Ouden et al., 2009; Thompson et al., 2007, 2010a). This analysis thus indicates that T2, T3 and T4 recruited brain regions engaged in complex argument structure processing in healthy individuals.

We also performed an ROI analysis, preceded by modeling each participant's regionally specific true HRF. Individual variation was observed in HRF curves, with two participants (T2 and T3) showing normal-like hemodynamic responses, and the other two showing delayed responses, peaking at around 10–12 sec post-stimulus in both pre- and post-treatment scans. These patterns are similar to those noted in the aphasic participants studied by Bonakdarpour et al. (2007) and Thompson et al. (2010b). The ROI analysis showed upregulation following treatment for three of the four participants, constrained to the left hemisphere. T3 and T4 showed upregulation in the left supramarginal gyrus and T4 showed additional upregulation in left superior parietal cortex, regions involved in complex argument structure processing in healthy individuals. T1 evinced a different pattern, showing upregulation in Broca's area. It is possible that T1's large posterior lesion precluded activation in posterior ROIs, resulting in recruitment of anterior brain structures. Although T1's activation peak was more frontal than the peak observed in the insula in our healthy controls, the finding of upregulation in the left IFG is in line with healthy participants' general activation patterns for the same contrast.

The post-treatment upregulation of activation, rather than downregulation (i.e., post-treatment decrease in activation) on some grounds may seem unusual. That is, because some studies suggest that upregulation of neural activity is associated with decreases in efficiency of language processing (Sonty et al., 2003), it follows that a shift from greater, more widespread activation, to more focal activation may be expected when language improves following stroke. This latter view is consistent with the learning literature showing that skill acquisition correlates with spatially more restricted areas of activation and also with decreased BOLD response (Chein and Schneider, 2005; Haier et al., 1992; Raichle et al., 1992; Wong et al., 2007). Notably, however, although some studies have found this recovery pattern in neuroimaging studies of aphasia treatment (e.g., Cherney and Small, 2006; Richter et al., 2008), several studies have shown the opposite pattern: increases in activation following treatment (e.g., Crosson et al., 2005; Fridriksson et al., 2006; Thompson et al., 2001; see Thompson and den Ouden, 2008 for review). Indeed, the neurobiology of recovery is a complex issue and further research is needed to understand the relation between both cognitive and brain impairments and recovery of function. However, the present findings indicate that brain regions that are inactive prior to treatment and within the language network associated with language processing in healthy people have the potential to become actively engaged in linguistic processing following treatment. The treated participants in this study showed increased activation in temporoparietal regions, consistent with imaging data on argument structure processing in healthy speakers. These findings suggest that regions of the brain activated in healthy individuals performing a language function, and contralateral homologs of these regions, are areas capable of supporting that function in an automatized fashion, perhaps because of their redundant or “latent” capacity (Toga et al., 2006; Zilles and Amunts, 2009).

5. Conclusion

The findings reported here suggest that treatment aimed to improve verb (and sentence) production deficits that exploit the argument structure properties of verbs is efficacious for individuals with agrammatic aphasia. Treatment not only improved production of trained verbs, but it also resulted in generalization to untrained verbs, a finding that has not been forthcoming in studies examining the effects of semantic and phonological cuing approaches to verb deficits. In addition, treatment resulted in increased recruitment of posterior perisylvian, as well as superior parietal and sensory-motor cortices bilaterally. These regions are associated with verb processing and, in particular, complex argument structure processing in healthy individuals, suggesting that recovery was accompanied by restoration of normal-like brain activation patterns.

Acknowledgments

This research was supported by NIH grants R01-DC001948-19 and R01-DC007213-05 (Dr. Cynthia K. Thompson).

Appendix A

Verbs used in the behavioral portion of the study.

1-Argument	2-Argument	3-Argument
Cry	Kiss	Give
Yawn	Lick	Deliver
Smile	Watch	Write
Laugh	Sniff	Send
Wink	Bite	Build
Cough	Stir	Sell
Scream	Grab	Show
Bowl	Wash	Pour
Clap	Lift	Read
Point	Hit	Throw
Fish	Pull	
Box	Pinch	
Fence	Tickle	
Jump	Hug	
Hop	Push	
Tiptoe	Cut	
Skate	Zip	
Kneel	Kick	
Run	Chase	
Ski	Spill	

Appendix B



the boy

is giving

the apple

to the teacher

Example of picture stimuli used in training.

Appendix C

Verbs used in the fMRI task.

1-Argument	2-Argument
Cry	Kiss
Yawn	Lick
Laugh	Watch
Wink	Sniff
Cough	Bite
Scream	Pull
Bowl	Pinch
Clap	Push
Point	Zip
Fish	Kick
Box	Follow
Fence	Whisper
Jump	Spit
Run	Eat
Smile	Squeeze
Sneeze	Drink
Frown	Type
Nod	Smoke

Appendix D

Percent correct responses on verb naming probes, statistical significance and effect sizes for treated (T1–T4) and untreated control (C1–C4) participants.

Participant	Trained verbs					Untrained verbs				
	Pre	Post	χ^2	p	ϕ	Pre	Post	χ^2	p	ϕ
T1	25%	78%	4.95	.026	.4	36%	58%	1.16	.281	.22
T2	23%	75%	9.03	.003	.53	49%	70%	1.09	.296	.22
T3	20%	60%	5.1	.024	.41	36%	45%	.06	.806	.09
T4	40%	75%	3.68	.055	.35	39%	63%	1.48	.224	.24
Mean (SD)	27% (9%)	72% (8.1%)		.068 ^a		40% (6.2%)	59% (10.6%)		.068 ^a	
C1	30%	43%	.24	.62	.13	28%	43%	1.37	.242	.16
C2	25%	13%	– ^b	1	–.16	46%	38%	.32	.572	–.09
C3	58%	60%	.03	.862	.03	61%	60%	.01	.920	–.01
C4	70%	75%	0	1	.06	61%	75%	1.17	.279	.15
Mean (SD.)	46% (21.7%)	48% (26.6%)		.465 ^a		49% (15.7%)	54% (16.9%)		.465 ^a	

a Statistical significance based on a paired-sample Wilcoxon test.

b The chi-square test could not be applied due to small cell frequencies; Fisher's exact test was used instead.

Percent correct responses on verb production in sentence context, statistical significance and effect sizes for treated (T1–T4) and untreated control (C1–C4) participants.

Participant	Trained verbs					Untrained verbs				
	Pre	Post	χ^2	p	ϕ	Pre	Post	χ^2	p	ϕ
T1	18%	78%	12.13	<.001	.6	36%	58%	1.06	.303	.21
T2	33%	68%	2.51	.113	.3	48%	66%	.72	.396	.18
T3	8%	60%	10.09	.001	.56	39%	48%	.06	.806	.09
T4	38%	53%	.4	.527	.15	36%	61%	1.41	.235	.24
Mean (SD)	24% (14%)	65% (11%)		.068 ^a		40% (6%)	58% (8%)		.068 ^a	
C1	15%	15%	– ^b	1	0	23%	26%	.02	.887	.04
C2	65%	10%	10.67	.001	–.57	65%	40%	4.06	.044	–.25
C3	75%	65%	.12	.729	–.11	65%	70%	.06	.806	.05
C4	70%	80%	.13	.718	.12	73%	83%	.65	.420	.12
Mean (SD)	56% (28%)	43% (35%)		.414 ^a		57% (23%)	55% (26%)		.465 ^a	

a Statistical significance based on a paired-sample Wilcoxon test.

b The chi-square test could not be applied due to small cell frequencies; Fisher's exact test was used instead.

Percent correct responses on argument structure in sentence context, statistical significance and effect sizes for treated (T1–T4) and untreated control (C1–C4) participants.

Participant	Trained verbs					Untrained verbs				
	Pre	Post	χ^2	p	ϕ	Pre	Post	χ^2	p	ϕ
T1	0%	68%	17.47	<.001	.71	15%	46%	3.24	.072	.34
T2	8%	53%	7.62	.006	.49	35%	61%	1.6	.206	.25
T3	5%	55%	9.64	.002	.55	39%	46%	.03	.862	.08
T4	33%	38%	.01	.920	.04	35%	60%	1.6	.206	.25
Mean (SD)	12% (15%)	54% (12%)		.068 ^a		31% (11%)	53% (8%)		.068 ^a	
C1	3%	5%	– ^b	1	.07	15%	25%	.7	.403	.13
C2	65%	5%	13.3	<.001	–.63	65%	40%	4.06	.044	–.25
C3	45%	25%	.99	.320	–.21	63%	66%	.01	.920	.04
C4	65%	55%	.1	.752	–.1	71%	83%	.86	.354	.13
Mean (SD)	45% (29%)	23% (24%)		.144 ^a		54% (26%)	54% (26%)		.715 ^a	

a Statistical significance based on a paired-sample Wilcoxon test.

b The chi-square test could not be applied due to small cell frequencies; Fisher's exact test was used instead.

REFERENCES

- Bastiaanse R, Hurkmans J, and Links P. The training of verb production in Broca's aphasia: A multiple-baseline across-behaviours study. *Aphasiology*, 20(2): 298–311, 2006.
- Bastiaanse R and Jonkers R. Verb retrieval in action naming and spontaneous speech in agrammatic and anomia aphasia. *Aphasiology*, 12: 951–969, 1998.
- Bastiaanse R and van Zonneveld R. Broca's aphasia, verbs and the mental lexicon. *Brain and Language*, 90: 198–202, 2004.
- Belin P, Van Eackhout P, Zilbovicius M, Remy P, Francois C, Guillaume S, et al. Recovery from nonfluent aphasia after melodic intonation therapy: A PET study. *Neurology*, 47: 1504–1511, 1996.
- Ben-Shachar M, Hendlar T, Kahn I, Ben-Bashat D, and Grodzinsky Y. The neural reality of syntactic transformations. *Psychological Science*, 14: 433–440, 2003.
- Benjamini Y and Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57: 289–300, 1995.
- Binder JR, Desai RH, Graves WW, and Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19: 2767–2796, 2009.
- Boersma P and Weenink D. *Praat: Doing Phonetics by Computer [Computer Program]*.
- Bonakdarpour B, Parrish T, and Thompson CK. Hemodynamic response changes in patients with stroke-induced aphasia: Implications for fMRI data analysis. *NeuroImage*, 36: 322–331, 2007.
- Breier JI, Maher LM, Schmadeke S, Hasan KM, and Papanicolaou AC. Changes in language-specific brain activation after therapy for aphasia using magnetoencephalography: A case study. *Neurocase*, 13: 169–177, 2007.
- Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, et al. Action observation activated premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, 13: 400–404, 2001.
- Cao Y, Vikingstad EM, George KP, Johnson AF, and Welch KMA. Cortical language activation in stroke patients recovering from aphasia with functional fMRI. *Stroke*, 30: 2331–2340, 1999.
- Caplan D, Hildebrandt N, and Makris N. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain*, 119: 933–949, 1996.
- Chein JM and Schneider W. Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, 25: 607–623, 2005.
- Cherney LR and Small SL. Task-dependent changes in brain activation following therapy for nonfluent aphasia: Discussion of two individual cases. *Journal of the International Neuropsychological Society*, 12: 828–842, 2006.
- Conroy P, Sage K, and Lambon Ralph MA. A comparison of word versus sentence cues as therapy for verb naming in aphasia. *Aphasiology*, 23: 462–482, 2009.
- Crepaldi D, Berlinger M, Paulesu E, and Luzzatti C. A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain and Language*, 116: 33–49, 2011.
- Crinion J, Holland AL, Copland DA, Thompson CK, and Hillis AE. Neuroimaging in aphasia treatment research: Quantifying brain lesions after stroke. *NeuroImage*. <http://dx.doi.org/10.1016/j.neuroimage.2012.07.044>, 2012.
- Crinion JT and Leff AP. Recovery and treatment of aphasia after stroke: Functional imaging studies. *Current Opinion in Neurology*, 20: 667–673, 2007.
- Crosson B, Bacon Moore A, Gopinath K, White KD, Wierenga CE, Gaiefsky ME, et al. Role of the right and left hemispheres in recovery of function during treatment of intention in aphasia. *Journal of Cognitive Neuroscience*, 17: 392–406, 2005.
- Crosson B, Benefield H, Cato MA, Sadek JR, Moore AB, Wierenga CE, et al. Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*, 9: 1061–1077, 2003.
- Crosson B and Haaland KY. Subcortical functions in cognition: Toward a consensus. *Journal of the International Neuropsychological Society*, 9: 1027–1030, 2003.
- Dabul BL. *Apraxia Battery for Adults*. Austin, TX: Pro-Ed, 1979.
- Damasio AR and Tranel D. Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Science in the USA*, 90: 4957–4960, 1993.
- Davis CH and Harrington G. Intensive semantic intervention in fluent aphasia: A pilot study with fMRI. *Aphasiology*, 20: 59–83, 2006.
- De Bleser R and Kauschke C. Acquisition and loss of nouns and verbs: Parallel or divergent patterns? *Journal of Neurolinguistics*, 16: 213–229, 2003.
- Den Ouden DB, Fix S, Parrish TB, and Thompson CK. Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22: 196–215, 2009.
- Dragoy O and Bastiaanse R. Verb production and word order in Russian agrammatic speakers. *Aphasiology*, 24: 28–55, 2010.
- Dronkers NF, Plaisant O, Iba-Zizen MT, and Cabanis EA. Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130: 1432–1441, 2007.
- Dronkers NF, Shapiro JK, Redfern B, and Knight RT. The role of Broca's area in Broca's aphasia. *Journal of Clinical and Experimental Neuropsychology*, 14: 52–53, 1992.
- Fridriksson J, Morrow-Odom L, Moser D, Fridriksson A, and Baylis G. Neural recruitment associated with anomia treatment in aphasia. *NeuroImage*, 32: 1403–1412, 2006.
- Genovese CE, Lazar NA, and Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15: 870–878, 2002.
- Haier RJ, Siegel BV, MacLauchlan A, Soderling E, Lottenberg S, and Buchsbaum MS. Regional glucose metabolic changes after learning a complex visuospatial/motor task: A positron emission tomographic study. *Brain Research*, 570: 134–143, 1992.
- Holland AL, Greenhouse JB, Fromm D, and Swindell CS. Predictors of language restitution following stroke: A multivariate analysis. *Journal of Speech and Hearing Research*, 32: 232–238, 1989.
- Huttenlocher J and Lui F. The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior*, 18: 141–162, 1979.
- Kaan E and Swaab TY. The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, 6: 350–356, 2002.
- Kable JW, Lease-Spellmeyer J, and Chatterjee A. Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, 14: 795–805, 2002.
- Kemmerer D and Gonzalez Castillo J. The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*, 112: 54–76, 2010.
- Kemmerer D and Tranel D. Verb retrieval in brain-damaged subjects. *Brain and Language*, 73: 347–420, 2000.
- Kertesz A. *Western Aphasia Battery-revised*. San Antonio, TX: PsychCorp, 2007.
- Kertesz A and McCabe P. Recovery patterns and prognosis in aphasia. *Brain*, 100: 1–18, 1977.
- Kielar A, Meltzer-Asscher A, and Thompson CK. Electrophysiological responses to argument structure

- violations in healthy adults and individuals with agrammatic aphasia. *Neuropsychologia*, 50: 3320–3337, 2012.
- Kim M and Thompson CK. Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, 74: 1–25, 2000.
- Kim M and Thompson CK. Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. *Brain and Language*, 88: 1–20, 2004.
- Kiran S, Ansaldo A, Bastiaanse R, Cherney LR, Howard D, Farooq-Shah Y, et al. Neuroimaging in aphasia treatment research: Standards for establishing the effects of treatment. *NeuroImage*. <http://dx.doi.org/10.1016/j.neuroimage.2012.10.011>, 2012.
- Kiran S and Thompson CK. 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: 773–787, 2003.
- Kiss K. Effect of verb complexity on agrammatic aphasics' sentence production. In Bastiaanse R and Grodzinsky Y (Eds), *Grammatical Disorders in Aphasia: A Neurolinguistic Perspective*. London: Whurr, 2000: 123–151.
- Koenigs M, Barbey AK, Postle BR, and Grafman J. Superior parietal cortex is critical for the manipulation of information in working memory. *Journal of Neuroscience*, 29: 14980–14986, 2009.
- Kraut MA, Kremen S, Segal JB, Calhoun V, Moo LR, and Hart J. Object activation from features in the semantic system. *Journal of Cognitive Neuroscience*, 14: 24–36, 2002.
- Laska AC, Hellblom A, Murray V, Kahan T, and Von Arbin M. Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, 249: 413–422, 2001.
- Lazar RM and Antonello D. Variability in recovery from aphasia. *Current Neurology and Neuroscience Reports*, 8: 497–502, 2008.
- Lazar RM, Speizer AE, Festa JR, Krakauer JW, and Marshall RS. Variability in language recovery after first-time stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, 79: 530–534, 2008.
- Leger A, Demonet JF, Ruff S, Aithamon B, Touyeras B, Puel M, et al. Neural substrates of spoken language rehabilitation in an aphasia patient: An fMRI study. *NeuroImage*, 17: 174–183, 2002.
- Lenderm W, McGuirk E, and Lincoln NB. Factors affecting language recovery in aphasic stroke patients receiving speech therapy. *Journal of Neurology, Neurosurgery and Psychiatry*, 51: 1103–1104, 1988.
- Levin B. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: University of Chicago Press, 1993.
- Luzzatti C, Raggi R, Zonca G, Pistarini C, Contardi A, and Pinna GD. Verb-noun double dissociation in aphasic lexical impairments: The role of word frequency and imageability. *Brain and Language*, 81: 432–444, 2002.
- Maas E, Barlow J, Robin D, and Shapiro L. Treatment of sound errors in aphasia and apraxia of speech: Effects of phonological complexity. *Aphasiology*, 16(4–6): 609–622, 2002.
- Macey PM, Macey KE, Kumar R, and Harper RM. A method for removal of global effects from fMRI time series. *NeuroImage*, 22: 360–366, 2004.
- Maldjian JA, Laurienti PJ, Kraft RA, and Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19: 1233–1239, 2003.
- Marshall J, Pring T, and Chiat S. Verb retrieval and sentence production in aphasia. *Brain and Language*, 63: 159–183, 1998.
- Meinzer M, Beeson PM, Cappa S, Crinion J, Kiran S, Saur D, et al. Neuroimaging in aphasia treatment research: Consensus and practical guidelines for data analysis. *NeuroImage*. <http://dx.doi.org/10.1016/j.neuroimage.2012.02.058>, 2012.
- Meinzer M, Flaisch T, Breitenstein C, Wienbruch C, Elbert T, and Rockstroh B. Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. *NeuroImage*, 39: 2038–2046, 2008.
- Meltzer JA, Postman-Caucheteau WA, McArdle JJ, and Braun AR. Strategies for longitudinal neuroimaging studies of overt language production. *NeuroImage*, 47: 745–755, 2009.
- Meltzer-Asscher A, Schuchard J, den Ouden DB, and Thompson CK. The neural substrates of complex argument structure representations: Processing 'alternating transitivity' verbs. *Language and Cognitive Processes*. <http://dx.doi.org/10.1080/01690965.2012.672754>, 2012.
- Miceli G, Mazzucchi A, Menn L, and Goodglass H. Contrasting cases of Italian agrammatic aphasia without comprehension disorder. *Brain and Language*, 19: 65–97, 1983.
- Pedersen PM, Jorgensen HS, Nakayama H, Raaschou HO, and Olsen TS. Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology*, 38: 659–666, 1995.
- Perani D, Cappa SF, Tettamanti M, Rosa M, Scifo P, Miozzo A, et al. A fMRI study of word retrieval in aphasia. *Brain and Language*, 85: 357–368, 2003.
- Raboyeau G, de Boissezon X, Marie N, Balduyck S, Puel M, Bézy D, et al. Right hemisphere activation in recovery from aphasia: Lesion effect or function recruitment? *Neurology*, 70: 290–298, 2008.
- Raichle ME, Fiez JA, Videen TO, MacLeod AMK, Pardo JV, Fox PT, et al. Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, 4: 8–26, 1992.
- Rapp B, Caplan D, Edwards S, Visch-Brink E, and Thompson CK. Neuroimaging in aphasia treatment research: Issues of experimental design for relating cognitive and neural changes. *NeuroImage*. <http://dx.doi.org/10.1016/j.neuroimage.2012.09.007>, 2012.
- Raymer AM and Ellsworth TA. Response to contrasting verb retrieval treatments: A case study. *Aphasiology*, 16: 1031–1045, 2002.
- Raymer AM, Moberg PJ, Crosson B, Nadeau SE, and Gonzalez-Rothi LJ. Lexical-semantic deficits in two cases of thalamic lesion. *Neuropsychologia*, 35: 211–219, 1997.
- Richter M, Miltner WHR, and Straube T. Association between therapy outcome and right-hemispheric activation in chronic aphasia. *Brain*, 131: 1391–1401, 2008.
- Riley EA. Effects of phonological complexity on error production and pseudoword training in acquired phonological dyslexia. Dissertation. Evanston, IL: Northwestern University, 2011.
- Rizzolatti G and Craighero L. The mirror-neuron system. *Annual Review of Neuroscience*, 27: 169–192, 2004.
- Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, Paulesu E, Perani D, et al. Localization of grasp representation in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, 111: 246–252, 1996.
- Robey RR. A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*, 41: 172–187, 1998.
- Rosen HJ, Petersen SE, Linenweber MR, Snyder AZ, White DA, Chapman L, et al. Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology*, 55: 1883–1894, 2005.
- Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, et al. Dynamics of language reorganization after stroke. *Brain*, 129: 1371–1384, 2006.
- Schneider SL and Thompson CK. Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalization. *Aphasiology*, 17: 213–241, 2003.
- Shapiro L, Gordon B, Hack N, and Killackey J. Verb argument processing in complex sentences in Broca and Wernicke's aphasia. *Brain and Language*, 45: 423–447, 1993.

- Shapiro L and Levine B. Verb processing during sentence comprehension in aphasia. *Brain and Language*, 38: 21–47, 1990.
- Shapiro K, Mottaghy FM, Schiller NO, Poeppel TD, Fluss MO, Muller HW, et al. Dissociating neural correlates for nouns and verbs. *NeuroImage*, 24: 1058–1067, 2005.
- Shmuelof L and Zohary E. Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, 47: 457–470, 2005.
- Silveri MC, Perri R, and Cappa A. Grammatical class effects in brain-damaged patients: Functional locus of noun and verb deficit. *Brain and Language*, 85: 49–66, 2003.
- Signoret J, Castaigne P, Lehermitte F, Abelanet R, and Lavorel P. Rediscovery of Leborgne's brain: Anatomical description with CT scan. *Brain and Language*, 22: 303–319, 1984.
- Sonty SP, Mesulam MM, Thompson CK, Johnson NA, Weintraub S, Parrish TB, et al. Primary progressive aphasia: PPA and the language network. *Annals of Neurology*, 53: 35–49, 2003.
- Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, et al. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17: 273–281, 2005.
- Thompson CK. *Northwestern Assessment of Verbs in Sentences*. Evanston, IL: Northwestern University, 2011.
- Thompson CK. Unaccusative verb production in agrammatic aphasia: The argument structure complexity hypothesis. *Journal of Neurolinguistics*, 16: 151–167, 2003.
- Thompson CK, Lange KL, Schneider SL, and Shapiro LP. Agrammatic and non-brain-damaged subjects' verb and verb argument production in constrained elicitation conditions. *Aphasiology*, 11: 473–490, 1997.
- Thompson CK, Bonakdarpour B, and Fix SF. Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. *Journal of Cognitive Neuroscience*, 22: 1993–2011, 2010a.
- Thompson CK, Bonakdarpour B, Fix SF, Blumenfeld HD, Parrish TB, Gitelman DR, et al. Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19: 1753–1767, 2007.
- Thompson CK, Cho S, Ju-Hsu C, Wieneke C, Rademaker A, Weitner BB, et al. Dissociations between fluency and agrammatism in primary progressive aphasia. *Aphasiology*, 26: 20–43, 2012.
- Thompson CK and den Ouden DB. Neuroimaging and recovery of language in aphasia. *Current Neurology and Neuroscience Reports*, 8: 475–483, 2008.
- Thompson CK, den Ouden DB, Bonakdarpour B, Garibaldi K, and Parrish TB. Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia*, 48: 3211–3227, 2010b.
- Thompson CK, Fix S, Gitelman DR, Parrish TB, and Mesulam MM. The neurobiology of recovered sentence comprehension in aphasia: Treatment-induced fMRI activation patterns. *Journal of Cognitive Neuroscience*, S164, 2001.
- Thompson CK, Shapiro LP, Kiran S, and 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*, 46: 591–607, 2003.
- Thompson CK, Shapiro LP, Tait ME, Jacobs B, Schneider S, and Ballard K. A system for the linguistic analysis of agrammatic language production. *Brain and Language*, 51: 124–129, 1995.
- Toga AW, Thompson PM, Mori S, Amunts K, and Zilles K. Towards multimodal atlases of the human brain. *Nature Neuroscience*, 7: 952–966, 2006.
- Tsukiura T, Fujii T, Takahashi T, Xiao R, Inase M, Iijima T, et al. Neuroanatomical discrimination between manipulating and maintaining processes involved in verbal working memory: A functional MRI study. *Cognitive Brain Research*, 11: 13–21, 2001.
- Tyler LK, Bright P, Fletcher P, and Stamatakis EA. Neural processing of nouns and verbs: The role of inflectional morphology. *Neuropsychologia*, 42: 512–523, 2004.
- Vanier M and Caplan D. CT-scan correlates of agrammatism. In Menn L and Obler LK (Eds), *Agrammatic Aphasia*. Amsterdam: Benjamins, 1990: 97–114.
- Van Valin RD. *Exploring the Syntax–Semantics Interface*. Cambridge: Cambridge University Press, 2005.
- Wambaugh J, Cameron R, and Kalinyak-Fliszar M. Retrieval of action names in aphasia: Effects of two cueing treatments. *Aphasiology*, 18: 979–1004, 2004.
- Wertz RT, LaPointe LL, and Rosenbek JC. *Apraxia of Speech in Adults: The Disorder and Its Management*. Orlando: Grune and Stratton, 1984.
- Wilson SM and Saygin AP. Grammatical judgment in aphasia: Deficits are not specific to specific syntactic structures, aphasic syndromes, or lesion sites. *Journal of Cognitive Neuroscience*, 16: 238–252, 2004.
- Wong PCM, Perrachione TK, and Parrish T. Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, 28: 995–1006, 2007.
- Zilles K and Amunts K. Receptor mapping: Architecture of the human cerebral cortex. *Current Opinion in Neurology*, 22: 331–339, 2009.