



Published in final edited form as:

Cortex. 2019 November ; 120: 394–418. doi:10.1016/j.cortex.2019.06.015.

## Recovery of offline and online sentence processing in aphasia: Language and domain-general network neuroplasticity

Elena Barbieri<sup>a,\*</sup>, Jennifer Mack<sup>a</sup>, Brianne Chiappetta<sup>a</sup>, Eduardo Europa<sup>a</sup>, Cynthia K. Thompson<sup>a,b,c</sup>

<sup>a</sup>Center for the Neurobiology of Language Recovery, Northwestern University, Evanston, IL, USA

<sup>b</sup>Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Chicago, IL, USA

<sup>c</sup>Department of Neurology, Northwestern University, Chicago, IL, USA

### Abstract

This paper examined the effects of treatment on both offline and online sentence processing and associated neuroplasticity within sentence processing and dorsal attention networks in chronic stroke-induced agrammatic aphasia. Twenty-three neurotypical adults and 19 individuals with aphasia served as participants. Aphasic individuals were randomly assigned to receive a 12-week course of linguistically-based treatment of passive sentence production and comprehension (N = 14, treatment group) or to serve as control participants (N = 5, natural history group). Both aphasic groups performed two offline tasks at baseline and three months following (at post-testing) to assess production and comprehension of trained passive structures and untrained syntactically related and unrelated structures. The aphasic participants and a healthy age-matched group also performed an online eyetracking comprehension task and a picture-verification fMRI task, which were repeated at post-testing for the aphasic groups. Results showed that individuals in the treatment, but not in the natural history, group improved on production and comprehension of both trained structures and untrained syntactically related structures. Treatment also resulted in a shift toward more normal-like eye movements and a significant increase in neural activation from baseline to post-testing. Upregulation encompassed right hemisphere regions homologs of left hemisphere regions involved in both sentence processing and domain-general functions and was positively correlated with treatment gains, as measured by offline comprehension accuracy, and with changes in processing strategies during sentence comprehension, as measured by eyetracking. These findings provide compelling evidence in favor of the contribution of both networks within the right hemisphere to the restoration of normal-like sentence processing patterns in chronic aphasia.

\* Corresponding author. elena.barbieri@northwestern.edu, elena.barbieri83@gmail.com (E. Barbieri).

CRedit authorship contribution statement

**Elena Barbieri:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Jennifer Mack:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Brianne Chiappetta:** Investigation, Data curation, Writing - original draft, Visualization. **Eduardo Europa:** Investigation, Formal analysis, Visualization. **Cynthia K. Thompson:** Conceptualization, Methodology, Investigation, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Open Practices

The study in this article earned an Open Materials badge for transparent practices.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2019.06.015>.

## Keywords

Aphasia; Sentence comprehension; Online sentence processing; fMRI; Eyetracking; Treatment of underlying forms

---

## 1. Introduction

Studies focused on recovery of language in chronic stroke-induced aphasia have shown that language treatment (therapy) improves language ability and that it results in measurable changes in brain function. Among these, many studies show increased activation (i.e., upregulation) following language treatment in non-lesioned cortical tissue in the left hemisphere (LH), including perilesional areas (Belin et al., 1996; Fridriksson, Bonilha, Baker, Moser, & Rorden, 2009; Fridriksson, Richardson, Fillmore, & Cai, 2012; Leger et al., 2002; Meinzer et al., 2008; Wierenga et al., 2006), although this may be dependent on the amount of preserved white matter within the ipsilesional hemisphere (Bonilha, Gleichgerrcht, Nesland, Rorden, & Fridriksson, 2016). Changes in neural activation following language treatment also have been reported in the contralesional, right hemisphere (RH; Abel, Weiller, Huber, Willmes, & Specht, 2015; Breier, Maher, Novak, & Papanicolaou, 2006; Blasi et al., 2002; Cherney & Small, 2006; Crosson et al., 2009; Kiran, Meier, Kapse, & Glynn, 2015; Mohr, Difrancesco, Harrington, Evans, & Pulvermüller, 2014; Raboyeau et al., 2008; Thompson, den Ouden, Bonakdarpour, Garibaldi, & Parrish, 2010; Thompson, Riley, den Ouden, Meltzer-Asscher, & Lukic, 2013), with some studies also finding changes in the functional or structural connections within the RH (Kiran et al., 2015; Schlaug, Marchina, & Norton, 2009; Wan, Zheng, Marchina, Norton, & Schlaug, 2014).

Although providing compelling evidence for treatment-induced neural plasticity in chronic aphasia, these studies are inconclusive with respect to the mechanisms underlying language recovery. Specifically, it is unclear whether recruitment of RH tissue in aphasia recovery is adaptive or mal-adaptive (Gainotti, 2015; Hartwigsen & Saur, 2019). Evidence in favor of a compensatory role of the RH comes from neuroimaging studies showing a positive relation between RH upregulation and/or increased connectivity and treatment outcome (Breier et al., 2006; Kiran et al., 2015; Raboyeau et al., 2008). Data in favor of a maladaptive role of the RH in aphasia recovery comes primarily from studies using non-invasive brain stimulation (e.g., repetitive transcranial magnetic stimulation, rTMS; transcranial direct current stimulation, tDCS), which have shown improved language performance following application of excitatory stimulation to the LH (e.g., Fridriksson et al., 2018) or inhibitory rTMS to the RH inferior frontal gyrus (IFG), either with (Thiel et al., 2013; Weiduschat et al., 2011) or without (Barwood et al., 2011; Martin et al., 2009) concomitant language treatment, although treatment gains are often minimal (see also Norise & Hamilton, 2017). Findings of a few neuroimaging studies further support this observation, by showing upregulation of the RH in non-recovered (vs recovered) participants (Marcotte et al., 2012), correlations between errors on a picture naming task and RH activation (Postman-Caucheteux et al., 2010), or a direct association between treatment efficacy and decreased activation (i.e., downregulation) in the RH (Abel et al., 2015; Baciú et al., 2016; Nardo, Holland, Leff, Price, & Crinion, 2017).

Inconsistencies across studies may result from differences in recruitment criteria and consequently in the participants' characteristics, including demographic variables (i.e., age and sex), where better recovery is found in younger (*vs older*) and female (*vs male*) individuals (e.g., Laska, Hellblom, Murray, Kahan, & Von Arbin, 2001; McGlone, 1977; Pedersen, Vinter, & Olsen, 2004), aphasia type and severity, with milder forms of aphasia generally resulting in a better outcome than more severe aphasia (Lazar et al., 2010; Pedersen et al., 2004), and time post-stroke (e.g., Bakheit, Shaw, Carrington, & Griffiths, 2007; Demeurisse et al., 1980; but see; Lazar et al., 2010). Lesion-related variables also must be considered when examining the neuroplasticity of language networks. While some studies have shown better outcomes and greater LH upregulation for smaller (*vs larger*) lesions (Maas et al., 2012), particularly in perilesional tissue, others have found no differences based on lesion size (Ansaldi, Arguin, & Roch Lecours, 2002; Heiss & Thiel, 2006; Hillis, 2007; Mattioli et al., 2014). Recent research also has shown that better language outcomes are predicted by the extent to which white matter tracts, such as the uncinate and the superior/inferior longitudinal fasciculi in the LH (Hope, Seghier, Leff, & Price, 2013) or the long segment of the arcuate fasciculus in the RH are compromised by stroke (Forkel et al., 2014). Further, studies suggest that recovery may be predicted by the integrity of tissue within specific brain regions, such as the middle/superior temporal gyrus and the basal ganglia (Bonilha et al., 2016; Fridriksson et al., 2012; Heiss, Thiel, Kessler, & Herholz, 2003; Naeser, Helm-Estabrooks, Haas, Auerbach, & Srinivasan, 1987). In a meta-analysis, Turkeltaub, Messing, Norise, and Hamilton (2011) found a relation between lesions within the left IFG and recruitment of the RH homologous IFG.

Several aphasia treatment studies also address the role of regions associated with domain-general processes in recovery of language. Two studies by Kiran and co-workers (Kiran et al., 2015; Sandberg, Bohland, & Kiran, 2015) found increased connectivity – following a semantic featured-based and a word retrieval treatment for anomia – between the IFG and the middle frontal gyrus (MFG), a region implicated in cognitive control (Fedorenko, Duncan, & Kanwisher, 2013), working memory (Curtis & D'Esposito, 2003) and attentional processes (Corbetta, Patel, & Shulman, 2008), in both hemispheres. Changes in activation in bilateral MFG and superior parietal lobule (SPL), regions that are part of the dorsal attention network (DAN, Corbetta, Patel, & Shulman, 2008), were also observed following sentence processing treatment in the study by Thompson and den Ouden et al. (2010) and bilateral upregulation of the SPL was found following a verb-argument structure treatment for production of canonical sentences (Thompson et al., 2013), possibly reflecting increased engagement of top-down attentional-executive mechanisms, such as self-monitoring and response inhibition (Geranmayeh, Brownsett, & Wise, 2014; Kurland, Baldwin, & Tauer, 2010). Attesting to this, Brownsett et al. (2013) tested neurotypical and aphasic individuals using a listen-and-repeat functional magnetic resonance imaging (fMRI) task in which, in order to simulate the difficulties in speech perception experienced by individuals with aphasia, healthy participants were exposed to three-channel noise-vocoded speech. Results showed that two domain-general networks, i.e., the salience network (encompassing the anterior cingulate and the operculum) and the executive control network (encompassing the dorsolateral prefrontal and parietal cortices) in both hemispheres, were equally active during speech perception in both groups. Brownsett, et al. (2013), however, did not find changes in

activation of these networks following language intervention. Therefore, the role of domain-general processes in language recovery is not completely clear. Further, no studies to our knowledge have addressed the relation between the extent to which regions within domain-general networks (i.e., in the left hemisphere) are damaged by stroke and recruitment of domain-general systems to support recovery.

The type of treatment as well as the neuroimaging tasks used to evaluate neural change may also affect outcomes, likely due to differences in the linguistic processes exploited (see Gainotti, 2015; Kiran & Thompson, 2019, for reviews). Most neuroimaging studies of aphasia treatment have focused on the effects of naming treatment, using, for example, cued retrieval therapy (Bonilha et al., 2016; Fridriksson et al., 2012, 2009) or semantic-feature based treatment (Kiran et al., 2015), with mixed findings. Notably, RH (more so than LH) recruitment is associated with semantically-based treatment, likely because semantic processing engages bilateral neural tissue in healthy people (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Kielar, Deschamps, Jokel, & Meltzer, 2016; Wright, Stamakakis, & Tyler, 2012). This is in line with the idea that the RH contributes to processing highly-imageable and highly-frequent words (see Gainotti, 2013 for a review) and provides a ‘coarse’ interpretation of word meanings, due to weak and broadly distributed activation of semantic features, including secondary and/or irrelevant ones (Jung-Beeman, 2005). Moreover, aphasic individuals with semantic deficits rely more on RH activation following anomia treatment than patients with phonological deficits (Abel, Weiller, Huber, & Willmes, 2014), consistent with behavioral studies demonstrating that the RH has poor phonological and syntactic competence (Gainotti, 2013).

In the domain of syntactic processing, only a few studies have investigated neural plasticity following treatment (Mohr et al., 2014; Thompson and den Ouden et al., 2010; Wierenga et al., 2006). These studies have found both RH and LH activation shifts associated with intervention. In healthy people, syntactic processing is primarily dependent on a LH fronto-temporal (dorsal stream) network, with some (but not all) studies also finding RH posterior perisylvian activation (Ben-Shachar, Palti, & Grodzinsky, 2004; Caplan, Chen, & Waters, 2008; Europa, Gitelman, Kiran, & Thompson, 2019; Grodzinsky & Friederici, 2006; Mack, Meltzer-Asscher, Barbieri, & Thompson, 2013; see Walenski, Europa, Caplan, & Thompson, 2019, for a recent meta-analysis). In one study, Thompson and den Ouden et al. (2010) found a LH network associated with complex noncanonical sentence processing (i.e., Wh-movement structures: object relatives, *Pete saw the groom who the bride carried*) in healthy listeners, using an auditory sentence-picture verification task (also see Den Ouden et al., 2012). Using the same task, individuals with aphasia were tested prior to and following provision of Treatment of Underlying Forms (TUF, Thompson & Shapiro, 2005, 2007), a linguistically-based approach focused on the lexical properties of verbs and syntactic mapping, to train comprehension and production of object-relative structures. Results showed that participants acquired the trained structures and showed generalized comprehension and production to simpler untrained Wh-movement structures (e.g., object Wh-questions) in keeping with the Complexity Account of Treatment Efficacy (CATE; Thompson, Shapiro, Kiran, & Sobecks, 2003). Analysis of changes in neural activation (as indexed by fMRI) from pre- to post-treatment showed upregulation of undamaged regions within the LH frontotemporal network as seen in healthy controls, including the middle

temporal gyrus (MTG), the angular gyrus (AG) and the SPL. In addition, RH regions, specifically the supramarginal gyrus (SMG) and the SPL showed upregulation following treatment. These findings suggest that people with aphasia recruit available tissue within domain-specific networks to support recovery.

Notably, TUF also has been shown to impact online processing strategies associated with complex sentence comprehension. In a recent study, using a sentence-picture matching task, we monitored the eye movements of unimpaired and aphasic participants as they listened to sentences and viewed two pictures depicting a single action with semantically reversed participants (e.g., *man shaving boy*; *boy shaving man*) (Mack & Thompson, 2017). The eye movements of unimpaired older adults showed evidence of *thematic prediction* followed by *thematic integration* for both active and passive sentences. For both sentence types, eye movements indicated an ‘Agent-first’ strategy, reflecting thematic prediction: when they heard the first noun phrase (NP; e.g., *the boy*) looks to the picture showing the boy as the Agent were dominant. For passive sentences, when the verb form (i.e., *was shaved*) was encountered, eye movements shifted to the picture depicting *the boy* as Theme, reflecting accurate thematic integration. Prior to treatment, the aphasic listeners failed to show either pattern for passive sentences (cf. Mack, Wei, Gutierrez, & Thompson, 2016; Meyer, Mack, & Thompson, 2012), however, at post-treatment their eye movements reflected emergence of normal-like processes (for both thematic prediction and integration), which was significantly associated with pre- to post-treatment improvements in comprehension accuracy. These findings indicate that treatment not only improves offline sentence comprehension, but also that it results in shifts to more normal sentence processing strategies. No studies to date, however, have evaluated the relation between changes in online sentence processing and the neural mechanisms that support it.

### 1.1. The present study

In this study we extended our previous work examining the behavioral and neurocognitive effects of TUF focused on training Wh-movement structures, to sentences associated with NP-movement. As in Wh-movement, in NP-movement structures like passives [e.g., *The woman<sub>i</sub> was kissed (t<sub>i</sub>) by the man*], the Theme argument is moved to the subject position, leaving behind a trace (t) or copy of movement in its original post-verbal position (Merge; Chomsky, 1995). Passive sentences are, therefore, noncanonical as the Theme precedes the Agent argument. NP-movement also is involved (on some accounts) in sentences with unaccusative verbs (Perlmutter, 1978), which select for a Theme in the object position that moves to the grammatical subject position [e.g., *The cat<sub>i</sub> was disappearing (t<sub>i</sub>) in the bushes*]. In this study we recruited two groups of participants with aphasia, with one group trained to comprehend and produce passive structures and the other serving as a control group, receiving no treatment. At baseline and three-months following, we examined comprehension and production of trained passive sentences and syntactically related, NP-movement (i.e., untrained passives and unaccusatives) as well as unrelated, Wh-movement (object clefts) structures in both participant groups. To ascertain the neurocognitive effects of treatment, we used online eyetracking and administered fMRI to evaluate passive sentence processing at both time points and examined the relation between treatment-induced changes in both offline and online sentence processing ability and neural shifts in

regions within sentence processing and domain-general networks in both the right and left hemisphere.

As in previous studies, we expected improved production and comprehension of trained and untrained, linguistically related structures for the treated, but not the control, group. We also expected treated participants to exhibit significantly greater shifts in neural activation and show changes in eye movements during passive sentence processing as compared to controls. Based on previous studies with healthy participants showing that processing of passive sentences is mostly left-lateralized (Hirotsu, Makuuchi, Rüschemeyer, & Friederici, 2011; Kinno, Kawamura, Shioda, & Sakai, 2008; Mack et al., 2013) and our previous findings of bilateral upregulation following sentence processing treatment (Thompson and den Ouden et al., 2010), we expected upregulation in regions within the sentence processing network (rather than in domain-general regions) in both hemispheres to positively correlate with treatment improvement as well as treatment-induced changes in eye movement patterns, reflecting plasticity of neurocognitive mechanisms, rather than maladaptive changes.

## 2. Method

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. Limited data are available without restrictions at <https://osf.io/u2qxa/>. The conditions of our ethical approval do not permit public archiving of raw MRI data. Researchers who wish to get access to this data should contact the Center for the Neurobiology of Language Recovery (cnlr@northwestern.edu) and sign a collaboration agreement.

### 2.1. Participants

Nineteen monolingual, English-speaking individuals [7 females, mean age = 49.7 yrs. (SD = 11.2; range = 22–73)] with chronic, stroke-induced aphasia [time post-stroke: mean = 49.1 months (SD = 33; range: 13–107)] were included in the study (see Table 1). Participants were randomly assigned to the treatment (n = 14; P1 – P14; treatment group) or control (natural history) group (n = 5; C1 – C5). Participants in the two groups were matched for age, education, and months post-onset based on non-parametric Mann–Whitney tests (all  $p$ -values >.1). All had suffered a single thromboembolic or hemorrhagic stroke (with the exception of P9<sup>1</sup>) in the left hemisphere at least one year prior to the study, and had no other impairments that impacted the ability to complete the experimental tasks (e.g., vision and hearing was within normal limits). Neuroimaging data from 23 cognitively healthy individuals (12 females) also were acquired for task validation. The healthy participants, ranging in age from 24 to 64 years (M = 37.1 yrs, SD = 13.1), were slightly younger than those with aphasia, based on Mann–Whitney non-parametric tests ( $p$  = .012). All participants met the criteria for MRI safety and provided written informed consent according to North-western University Institutional Review Board policies. The study was conducted

---

<sup>1</sup>P9 had two subsequent strokes on the same day, both within the left hemisphere.

in accordance with the rules established by the Declaration of Helsinki for experiments with human subjects.

The diagnosis of aphasia was made following extensive assessment of language functions, which included administration of the revised version of the *Western Aphasia Battery* (WAB-R; Kertesz, 2007), the *Northwestern Assessment of Verbs and Sentences* (NAVS; Thompson, 2011), the *Northwestern Naming Battery* (NNB; Thompson & Weintraub, 2014), as well as analysis of spontaneous speech (Cinderella story) using the Northwestern Narrative Language Analysis protocol (NNLA; Hsu & Thompson, 2018). Tests indicated language deficits consistent with nonfluent aphasia and agrammatism, with WAB Aphasia Quotients (WAB-AQs) ranging from 52.8 to 91.7, better production and comprehension of canonical than non-canonical sentences, and largely preserved word comprehension (Table 1). Analysis of narrative language samples revealed either (a) reduced speech rate [i.e., words per minutes (WPM)], and/or (b) decreased production of grammatically correct sentences, and/or (c) fewer productions of verbs (compared to nouns; i.e., noun:verb ratio) across participants. Non-parametric Mann-Whitney tests indicated no significant differences between treatment and natural history group on any of these measures (all  $p$ -values > .1).

## 2.2. Treatment procedures

**2.2.1. Sentence structures and experimental stimuli**—Sentence types selected for the study included syntactically related NP-movement structures [passive sentences and actives with unaccusative verbs (1)–(3)], semantically reversible active transitive sentences (4), and a linguistically unrelated Wh-movement structure [object clefts (5)]. Full passive sentences with locative adjuncts (1) were trained following TUF protocols (Thompson & Shapiro, 2005, 2007), while generalization to all other structures was tested.

1. The boy was shaved by the man in the barbershop.
2.
  - a. The man was shaved by the boy.
  - b. The man was shaved in the barbershop.
3. The man was arriving at the hospital.
4. The man was shaving the boy in the barbershop.
5. It was the boy who the man was shaving in the barbershop.

For full passives with adjuncts [FP; (1)], untrained simpler passives [UP; (2a,b)], active transitives [ACT; (4)], and object clefts [OC; (5)], twenty sentence/picture pairs (black and white line drawings) with semantically reversible participants were developed (see Fig. 1). Verbs in sentences used to test these structures were all transitive, selected animate (Agent and Theme) arguments, and had regular past participle inflection (-ed). For active unaccusatives [UAC; (3)], ten sentences with non-alternating unaccusative verbs and target/foil picture pairs, displaying the same action performed by participants of the opposite sex, were developed (see Appendices A and B). Unaccusative verbs all selected for animate (Theme) arguments and were matched to verbs used in FP/UP/ACT/OC structures for length in syllables, frequency of usage as a verb based on the Corpus of Contemporary American English (COCA, Davies, 2008), and imageability based on the Medical Research Council

(MRC, Coltheart, 1981) database. The same six animate nouns (see Appendices A and B) were used across all structures. Nouns used as adjuncts in UAC structures were different from adjuncts used in FP/UP/ACT/OC structures, but were matched for length in syllables and frequency of usage as a noun according to COCA. Treatment materials were developed for half of the FP structures and included word cards for the Action (verb; in both the active and passive form), Agent, Theme, and Location, as well as two sentence templates (one for the active and one for the passive form of the sentence).

**2.2.2. Experimental design**—A combined between-subjects and multiple baseline design across participants was used to evaluate the effects of treatment (Thompson, 2006). Upon entering the study, all participants were first administered a full probe task that tested production and comprehension of all sentence structures: FP (n = 20), UP (n = 40), ACT (n = 20), UAC (n = 10) and OC (n = 20) structures, for a total of 110 sentences. Next, participants in the treatment group received up to three baseline probes, which tested all FP and UAC structures only (one probe: P1, P2, P7, P10, P13; two probes: P3, P5, P8, P11, P12, P14; three probes: P4, P6, P9). These probes were administered within one to four weeks prior to treatment for all participants. Following baseline testing, participants in the treatment group received treatment twice a week (90 min each) for 12 weeks. When participants reached at least 80% accuracy on production and comprehension of FP structures in the weekly probe task, treatment sessions were reduced to one session per week, with the exception of P1 who showed rapid improvement across structures within 6 weeks, hence treatment was completed at this point. Participants who were assigned to the natural history group did not receive treatment, however, 12 weeks following initial testing, all were again tested using the full probe set for both production and comprehension.

**2.2.3. Treatment protocols**—Comprehension and production of passive sentences were trained using TUF (Thompson & Shapiro, 2005, 2007), a treatment approach that emphasizes the thematic roles of verb arguments and syntactic mapping from canonical (i.e., active) to noncanonical (i.e., passive) sentence forms using a set of metalinguistics steps. For each trial, comprehension was trained by presenting an action picture (e.g., a man shaving a boy at the barbershop) and by asking participants to point to word cards corresponding to the pictured verb (“action”), Agent of the action (i.e., “doer”) and the Theme (i.e., “recipient”). Next, the examiner built an active sentence by placing the word cards into the active sentence template and then demonstrated active to passive sentence formation using the passive sentence template. This involved replacing the active verb form with the past participle (*shaved*), moving the Recipient from the post-verbal position in the active structure to the subject position of the passive structure, and moving the Doer from the subject position of the active structure to the post-verbal adjunct position of the passive structure, with the addition of “by”. Production training followed similar steps, although participants were required – following the examiner’s demonstration – to build a passive structure using the provided word cards, and were instructed to read aloud and/or repeat after the examiner (in case of incorrect production) the targets (i.e., verbs, thematic roles and target sentences) at each training step. Treatment was provided by trained research assistants and monitored for fidelity, with an independent observer scoring half of the treatment sessions for adherence to the treatment protocol.



**2.2.4. Sentence comprehension and production probes**—Sentence production and comprehension were tested using a sentence production priming task and a picture verification task, respectively, both administered on a Lenovo computer running Super Lab 5.0. (Cedrus Corporation, [www.superlab.com](http://www.superlab.com)). For both modalities, full probes tested all structures, with the order of structures pseudorandomized across participants. The same shortened version of the probe task, including trained and untrained FP structures ( $n = 20$ ) and unaccusatives ( $n = 10$ ), that was used for baseline testing was administered weekly (every other treatment session) to monitor treatment progress. For both full and weekly probes, in the sentence production priming task, participants were shown two black and white drawings, side by side, each depicting the same action and the same characters, but with reversed thematic roles (i.e., a man shaving a boy and a boy shaving a man). Participants were instructed to listen to a ‘prime’ sentence describing one picture, and then produce a sentence just like it for the other picture (maximum response time: 15 sec). Responses were recorded using Audacity 1.2.5 and transcribed and scored by two trained lab members to ensure inter-rater reliability. Responses were considered correct if they included a verb, noun phrase(s), and locative adjunct (when required), with all produced in the correct order. Passive sentences (PA; i.e., FP and UP) also required at least two out of three passive markers (i.e., auxiliary, past participle, or *by*). In addition, OC structures required production of the main clause introduced by *it was*. Phonological paraphasias, semantic substitutions for nouns (e.g., *woman* for *girl*) and verb substitutions (with the same argument structure as the target, e.g., *watch* for *examine*), omission/substitutions of determiners, auxiliaries or prepositions within locative adjuncts, and omission of the relative pronoun *who/that* in OC sentences, were not counted as errors. The picture verification task required participants to listen to a sentence (recorded at 44100 Hz, with speech rate ranging between 3.3 and 3.5 syllables/second) while looking at an action picture appearing on the screen, and then press a key (max. time allowed to respond: 5500 msec) with their left hand to indicate a match (“F” on the keyboard) or a mismatch (“D” on the keyboard).

### 2.3. Eyetracking

All participants also performed a sentence-picture matching task as their eye movements were monitored to examine their use of online sentence comprehension strategies (see Mack & Thompson, 2017, for methodological details). Briefly, participants listened to semantically reversible active ( $n = 48$ ; e.g., *The woman was lifting the man*) or passive ( $n = 48$ ; e.g., *The man was lifted by the woman*) sentences while viewing a target picture that matched the sentence (e.g., a woman lifting the man) and a foil picture with reversed thematic roles (e.g., a man lifting the woman). Across items, the location (left or right) of the target picture was equally distributed, as was the location of the Agent and Theme within each picture. Thirty-two intransitive sentences were also included as fillers. There was no overlap between the verbs/sentences used in training and the eyetracking stimuli.

### 2.4. Neuroimaging

**2.4.1. Anatomical images**—Images were acquired on a Siemens 3T TIM Trio scanner or a Siemens 3T Prisma scanner. A standard T1-weighted 3D MPRAGE (TR = 2300 msec; TE = 2.91 msec; flip angle = 9°, FOV = 256 mm; voxel size = 1 × 1 × 1mm) anatomical scan was acquired in the sagittal plane using a 32- or a 64-channel head coil. Lesion masks

were developed for each participant using a semi-automated procedure: first, up to 5 lesion masks per participant were generated by a quality assurance anatomical pipeline available within the Northwestern University Neuroimaging Data Archive (NUNDA; Alpert, Kogan, Parrish, Marcus, & Wang, 2016), which employed a machine-learning algorithm to identify stroke-induced lesions based on signal intensity (Wang, Wang, Wang, Katsaggelos, & Parrish, 2019). Next, the best automatically-generated lesion map was independently selected by two members of the research team and any disagreement was discussed and resolved with the help of a third member; finally, lesion maps were manually modified on each axial slice and then visually inspected in all three planes by a member of the research team, using MRICron (Rorden & Brett, 2000). Using the pre-processing pipeline available on NUNDA (Song, Wang, Alpert, Wang, & Parrish, 2015), modified lesion masks (smoothed using a  $4 \times 4 \times 4$  mm filter) and anatomical images were co-registered and normalized to the VBM/DARTEL template (provided by Christian Gaser's VBM toolbox, <http://www.neuro.uni-jena.de/vbm/download/>) using enantiomorphic normalization, a non-linear registration method that derives normalization parameters from the undamaged hemisphere, thereby reducing normalization errors in the presence of focal lesions (Nachev, Coulthard, Jäger, Kennard, & Husain, 2008).

## 2.4.2. Functional imaging: acquisition and preprocessing

**2.4.2.1. fMRI TASK.:** All participants with aphasia performed an fMRI sentence comprehension task (presented using E-Prime version 2.0; Psychology Software Tools, Pittsburgh, PA) at two time points – baseline and 12-weeks later. The task employed a picture verification paradigm similar to the probe task described used to evaluate sentence comprehension ability and used a block design task (3 trials/block), where blocks of the experimental conditions (duration: 25.5 sec/block) were alternated with blocks of a control condition (duration: 19.5 sec/block, Fig. 2).

For the experimental conditions, 96 sentences ( $n = 48$  actives,  $n = 48$  passives) were developed using 24 transitive verbs, with each used to develop two active (e.g., *The boy was shaving the man/The man was shaving the boy*) and two passive structures (e.g., *The man was shaved by the boy/The boy was shaved by the man*). The same 20 transitive verbs used in FP/UP/ACT/OC structures of the full probe task were included, together with 4 additional verbs (i.e., *lick, lift, pinch, and push*) that were selected from the stimuli used in the eyetracking task and met the same criteria as all other verbs. In all 96 sentences, the same six nouns from the full probe task were used as Agents and Themes. Each experimental trial started with a black and white line drawing appearing on the screen for 500 msec (msec), followed by a sentence<sup>2</sup> that played through headphones (duration: 2100–2900 msec). Participants were instructed to indicate via button-press – within a 4000 msec response window – whether the sentence and picture matched (left index finger) or mismatched (left middle finger). Half of the sentences elicited a YES, and half elicited a NO, response. Each trial ended with a fixation cross (duration: 1100–1900 msec), to ensure that duration of each trial was exactly 8500 msec.

<sup>2</sup>All sentences were recorded using the same speech rate as for the behavioral version of the task.

For the control condition, the auditory stimuli consisted of eight time-reversed audio files (four randomly selected from each experimental condition). The visual stimuli were derived from eight randomly selected pictures that were partitioned into 8 × 8 grids, scrambled and rotated by 180° (50% of the instances) to become non-identifiable. In the control task, the auditory and visual stimuli were presented simultaneously (total trial duration: 6500 msec) and participants were instructed to press any button within a 2000 msec response window. Experimental (n = 96) and control (n = 24) trials were pseudorandomized across two runs (duration: 9m38sec/each) that were administered 1–7 days apart at both test points.

**2.4.2.2. FUNCTIONAL IMAGES.:** Blood oxygen level dependent (BOLD) functional images (TR = 2400 msec; TE = 20 msec; flip angle = 90°; FOV = 220 mm; voxel size = 1.7 × 1.7 × 3mm) were obtained using gradient echo-planar sequences. Task-dependent functional scans were pre-processed using the Robust fMRI pipeline provided by NUNDA (Song, et al., 2015), which combines pre-processing routines included in SPM (<https://www.fil.ion.ucl.ac.uk/spm>), AFNI (Cox, 1996), FSL (<https://www.fmrib.ox.ac.uk/fsl>) and FreeSurfer (<https://surfer.nmr.mgh.harvard.edu/>). Functional scans were first despiked to reduce the contribution of large spike signals on volume registration (Jo et al., 2013), then re-oriented to the radiological orientation and co-registered to the middle TR. Motion artifacts were identified and regressed out using a framewise displacement threshold of .5 mm (see Power, Barnes, Snyder, Schlaggar, & Petersen, 2012). Next, the BOLD signal was converted to percent signal change by scaling each voxel's mean value to 1000, and functional volumes were re-sliced at a resolution of 2 × 2 × 2 mm. Finally, functional images were co-registered to the corresponding anatomical image, warped to MNI space using a non-linear registration and smoothed using AFNI's 3dmerge filter with a 6 × 6 × 6 mm kernel.

## 2.5. Data analysis

**2.5.1. Treatment**—Performance on the weekly probe tasks (for both comprehension and production) was plotted over time to show acquisition curves for the trained FP items. For each individual in the treatment group, effect sizes for trained items were computed using Cohen's (1988) for comparison of proportions of correct responses. At baseline this corresponded to the average of the multiple baseline probes for all participants except for those who were tested only once at baseline; the proportion of correct responses on the final probe was derived from performance on the post-test. Effect sizes were computed only for participants showing an increase in performance from baseline to post-testing.

To investigate generalization to untrained items and structures, pre- and post-testing sentence comprehension and production accuracy on the full probes were calculated for each individual and averaged to obtain group means. Statistical analysis of the group data was undertaken using mixed-effects logistic regression (Jaeger, 2008) with *time point* (baseline, post-testing) and *sentence type* (FP untrained items, UP, UAC, ACT, OC) as fixed effects, and *item* and *participant* as random effects. For the regression analyses, in the presence of a significant interaction effect, planned comparisons were run for each structure, and FDR (false discovery rate, Benjamini & Hochberg, 1995) correction was applied to the uncorrected *p*-values.

**2.5.2. Eyetracking**—*Thematic prediction and thematic integration* values were calculated to quantify eye movement patterns derived from baseline and post-testing of passive sentences. Prediction scores were defined as the proportion of target fixations within the temporal region in which the first noun (N1) and verb (V) were presented with at least 200 msec total fixation time across both regions. Integration scores were defined as the difference in proportion of target fixations between the V and the sentence end (S.End) region with at least 100 msec total fixation time in each region. Both scores were computed only for correct trials and were compared from baseline to post-testing in both treatment and natural history groups using the Wilcoxon signed rank test. Difference scores were entered as predictor variables for the neuroimaging data analyses (see below).

### 2.5.3. fMRI

**2.5.3.1. GENERAL LINEAR MODEL (GLM) ANALYSES.:** First-level analyses were carried out using SPM8. For the two aphasic group participants, time and dispersion derivatives were modeled to account for possible effects of stroke on the hemodynamic response function (HRF, see Bonakdarpour, Parrish, & Thompson, 2007). For healthy participants, data were modeled using a standard HRF. Second-level (group-level) analyses were first conducted in SPM8, where group T-maps for all participant groups were thresholded at  $p < .001$  voxel-level. Next, the residuals derived from group maps were run through AFNI (Cox, 1996) to determine the appropriate cluster size threshold (see Eklund, Nichols, & Knutsson, 2016), based on an estimate of image smoothness (using the 3dFWHM and the 3dClustSim functions). Anatomical labels for regions above this threshold were obtained from the Harvarde–Oxford atlas (Desikan et al., 2006).

**2.5.3.2. SENTENCE-PROCESSING AND DOMAIN-GENERAL ROIS.:** Thresholded maps derived from the first-level analysis were binarized and intersected with the cortical regions of the Harvarde–Oxford atlas to extract regions-of-interest (ROI) suprathreshold activation for our contrast of interest, i.e., the difference in activation between passive and control blocks at baseline (BL) and post-testing (POST) (Passive > Control, POST > BL). The resulting map was intersected with a grey matter mask to extract ROI-based activation within the grey matter. Analyses were run within two sets of ROIs: one constituting the ‘sentence processing network (SPN)’ and one forming the ‘dorsal attention network (DAN)’, a domain-general network. The SPN, identified in a recent meta-analysis conducted on neuroimaging studies of sentence comprehension (Walenski et al., 2019), included the following regions: the pars triangularis of the IFG (IFGtri), the temporal- occipital portion of the MTG (MTGtpo), the posterior part of superior temporal gyrus (STGp) and the angular gyrus (AG). Although this network is almost completely left-lateralized in healthy individuals, analyses were conducted using ROIs from both hemispheres, with the aim to address the contribution of homologous regions in the contralesional hemisphere to language recovery. Notably, the areas included in the SPN also coincide with regions identified in Thompson and Meltzer-Asscher’s (2014) model of sentence processing as important for phrase structure building (IFG), access to verb arguments and thematic roles (AG) and integration of argument/thematic information within the syntax (posterior superior temporal and middle temporal gyri, pSTG/pMTG). The DAN included the following (bilateral) ROIs: the middle frontal gyrus (MFG), precentral gyrus (PCG), superior parietal lobule (SPL) and

the superior portion of the lateral occipital cortex (sLOC) (Corbetta et al., 2008; Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). This network was first identified in monkeys, as a set of regions encompassing the frontal eye fields (FEF) and the intraparietal sulcus (IPS). In humans, the FEF are located at the intersection of the MFG and the PCG (Vernet, Quentin, Chanes, Mitsumasa, & Valero-Cabré, 2014) and the IPS encompasses four regions (Swisher, Halko, Merabet, McMains, & Somers, 2007) whose MNI coordinates fall within the SPL and the sLOC ROIs of the Harvard–Oxford atlas. Notably, research demonstrates that the DAN plays a fundamental role in top-down allocation of attentional resources to visual features (Corbetta & Shulman, 2002; Vossel, Thiel, & Fink, 2006; also see; Muggleton, Kalla, Juan, & Walsh, 2011).

**2.5.3.3. REGRESSION ANALYSES.:** Binomial general linear mixed-effects models (GLMM) analyses were conducted using R 3.3.3. (R Core Team, 2017) to investigate 1) changes in activation between time points (e.g., POST > BL) in the two participant groups (treatment, natural history), and 2) the relation between changes in activation and changes in language scores across individuals. For offline comprehension, changes in language scores from baseline to post-testing were determined by computing z-scores for trained and untrained FP items (derived from the full probes) at each time point for each participant, summing these scores to derive a composite z-score, and computing the difference between time points. Treatment effects in online comprehension were calculated as the change between time points in prediction and integration scores, following z-transformation. Activation within each ROI ( $k > 5$ ) for the contrast of interest (Passive > Control, POST > BL) was the dependent variable in all regression models and was computed as the number of active voxels (per ROI) above the determined  $p < .001$  threshold, divided by the total number of intact voxels (per ROI). To account for differences in ROI volume, the total number of (intact) voxels in each ROI was introduced in each regression model using the ‘weight’ function available within the ‘lme4’ package (Bates, Maechler, Bolker, & Walker, 2015). The contribution of the following fixed effects was evaluated: *group* (treatment, natural history), *hemisphere* (left, right), and *language change* (for both offline and online comprehension). Follow-up analyses were conducted in the presence of significant interactions. Random effects for participant were introduced in all regression models and the proportion of intact gray matter within each ROI was entered as a covariate in all analyses to account for differences in lesion volume. Furthermore, to better understand the relation between lesion volume and treatment outcome, mixed-effects regression analyses were run using *lesion volume* as the dependent variable, and *group* and *language change* as fixed effects.

### 3. Results

#### 3.1. Offline accuracy

**3.1.1. Acquisition of trained FP structures—**Acquisition curves for trained FP items are shown in Fig. 3, reflecting the proportion of correct responses at baseline and on weekly probes administered throughout the treatment phase. For production (Figs. 3a), 13/14 participants showed significant improvement, with an increase of at least 30% correct production over baseline (group:  $z = 7.558$ ,  $p < .0001$ ). Effect sizes (Table 2) supported these

numerical acquisition trends, ranging from 1.16 to 2.5 (all large ESs). Only one participant (P12) showed no change in production of trained items (i.e., 0% correct on all probes). Comprehension was more variable (Fig. 3b), with 7/14 individuals showing a significant response to treatment (Cohen's  $H > .40$ , group analysis:  $z = 4.687$ ,  $p < .0001$ ) and 6 showing little to no change (% increase  $< 20\%$ ) from baseline to post-testing. Notably, baseline accuracy on trained items was high for one participant (P1), which may have prevented the ability to detect changes.

**3.1.2. Generalization to untrained items and structures**—Group results reflecting generalization to untrained structures (i.e., baseline and post-testing scores) are shown in Fig. 4 (production) and 5 (comprehension) for both the treatment (a) and the natural history (b) groups. Mixed-effects regression analyses indicated no significant between group differences in baseline accuracy on any of the structures for production (all  $p > .07$ ) or comprehension (FP.u:  $z = .752$ , nsec; UP:  $z = -1.705$ , nsec; UAC:  $z = -1.2$ , nsec; ACT:  $z = -.831$ , nsec; OC:  $z = -.464$ , nsec). In production, both groups showed higher accuracy for UAC than for PA (treatment group; FP.u:  $z = 6.341$ ,  $p < .0001$ ; UP:  $z = 7.172$ ,  $p < .0001$ ; natural history group; FP.u:  $z = 5.317$ ,  $p < .0001$ ; UP:  $z = 5.240$ ,  $p < .0001$ ) and accuracy was higher for ACT compared to PA for the treatment group, but not the natural history group (treatment group; FP.u:  $z = 5.061$ ,  $p < .0001$ ; UP:  $z = 5.732$ ,  $p < .0001$ ; natural history group; FP.u:  $z = .525$ , nsec; UP:  $z = -1.526$ , nsec). Analyses comparing accuracy at baseline and post-testing showed a significant *Group\*Phase* interaction, indicating improvement for the treatment group ( $z = 19.805$ ,  $p < .0001$ ), and no differences between time points for the natural history group ( $z = 1.043$ , nsec). The three-way *Group\*Phase\*Structure* interaction also was significant ( $p < .0001$ ) and separate analyses by structure indicated that participants in the treatment group (but not the natural history group) improved (from baseline to post-testing) in production of FP untrained items ( $z = 10.505$ ,  $p < .0001$ ), as well as of UP ( $z = 17.935$ ,  $p < .0001$ ), UAC ( $z = 6.247$ ,  $p < .0001$ ) and ACT structures ( $z = 2.761$ ,  $p = .009$ ). Although model estimates for OC structures could not be computed due to issues with convergence,<sup>3</sup> treated participants did not produce any correct OC sentences at any time point, while participants in the natural history group showed a very small change in accuracy (from 0% to 10% correct).

Across all participants and structures, comprehension was superior to production during the baseline phase, but poorer for all PA structures compared to UAC for both groups (treatment group; FP.u:  $z = -5.606$ ,  $p < .0001$ ; UP:  $z = -6.514$ ,  $p < .0001$ ; natural history group; FP:  $z = -3.576$ ,  $p = .0012$ ; UP:  $z = -3.701$ ,  $p = .0011$ ) and compared to ACT for the treatment group, but not the natural history group (treatment group; FP.u:  $z = 2.806$ ,  $p = .0072$ ; UP:  $z = 3.693$ ,  $p = .0004$ ; natural history group; FP.u:  $z = 1.569$ , nsec; UP:  $z = 1.527$ , nsec). However, no between group differences were found for any of the structures (all  $p$ -values = nsec), including baseline comprehension of PA and OC structures, indicating impaired comprehension of both non-canonical forms in both groups.

<sup>3</sup>Regression models for OC structures failed to converge because Mean and SD at BL were equal to zero in both participant groups. For this reason, estimates for the *Group\*Phase\*Structure* interaction were derived from models containing only FP, UP and UAC structures.

Analyses that compared accuracy at baseline and post-testing performance revealed a marginally significant *Group\*Phase* interaction ( $p = .094$ ), indicating significantly improved comprehension for the treatment group ( $z = 4.085$ ,  $p < .0001$ ) and no differences in accuracy across test point for the natural history group ( $z = .464$ , nsec). Although the *Group\*Phase\*Structure* interaction was not significant, separate analyses by structure (see Fig. 5) revealed significant improvement following treatment for FP.u ( $z = 2.463$ ,  $p = .0276$ ) and UP structures ( $z = -4.463$ ,  $p < .0001$ ), but not for ACT ( $z = .441$ , nsec), UAC ( $z = .622$ , nsec) or OC structures ( $z = 1.124$ , nsec).

### 3.2. Online (eyetracking) effects

The proportion of fixations to the target picture during correct passive trials at baseline and post-testing is shown in Fig. 6 for participants in the treatment and natural history groups. Eye tracking data for two participants in the treatment group (P7, P14) and one in the natural history group (C4) were not available (one due to equipment difficulties and two participants' data were not analyzable). Baseline eye-movement patterns did not significantly differ between the two groups at baseline (Wilcoxon test,  $p < .1$ ); neither group showed evidence of an Agent-first strategy (i.e., thematic prediction) as seen for neurotypical listeners (see Mack & Thompson, 2017), nor did they show evidence of thematic integration with a downstream shift in eye movements to the correct picture. However, at post-testing, eye movements consistent with an Agent-first strategy were noted for the treated group, with greater early looks to the incorrect picture, followed by looks to the correct picture in the sentence end (S.End) region. *Thematic Prediction (TP)* and *Thematic Integration (TI)* scores derived from eyetracking are summarized in Table 3. In line with the patterns shown in Fig. 6, within the treatment group, both prediction and integration scores significantly improved from baseline to post-testing. The mean TP score improved from .57 at baseline to .47 (Wilcoxon signed rank test,  $p = .024$ ) at post-testing, indicating increased fixations to the distractor picture (i.e., evidence of an Agent-first strategy). Notably, nine of the 12 treated participants showed this pattern. The TI scores also improved from baseline to post-testing, indicating a significant shift from distractor to target picture fixations (Wilcoxon signed rank test,  $p = .024$ ), with this pattern noted for 10/12 participants. All participants in the natural history group also showed improved TP scores at post-testing, with mean change at  $-.13$  for the group, and one participant improved in TI although the group mean indicated no improvements (group mean =  $-.03$ ). Further, statistical analyses of baseline to post-testing showed no significant changes in either score for the natural history group (Wilcoxon signed-rank test,  $p > .1$ ).

### 3.3. Neuroimaging results

**3.3.1. Healthy participants**—The *Passive > Control* contrast (see Fig. 7 and Table 4) revealed LH and RH clusters in posterior brain regions, with greater activation in the LH.

In the left, peak activation was found in the temporo-occipital portion of the inferior temporal gyrus (ITGtpo), extending to the inferior portion of the lateral occipital cortex (iLOC), posterior portion of the supramarginal gyrus (SMGp), STGp, fusiform gyrus (FG), and occipital pole, with smaller clusters in the sLOC, and paracingulate gyrus. In the right hemisphere, peak activation was found in the iLOC, extending to the AG, ITGtpo, MTGtpo,

and FG, and occipital pole. Bilateral activation also was found in the SPL. Activation in anterior cortical tissue was also greater in the LH compared to the RH, in the L pars opercularis and triangularis of the IFG (IFGop, IFGtri) and in two adjacent clusters peaking in the PCG and extending to the MFG and superior frontal gyrus (SFG). Activation in the RH frontal region was found only in SFG. For the opposite contrast (*Control > Passive*) RH activation dominated, with the largest cluster of peak activation in the planum temporale, extending to the anterior portion of the superior temporal gyrus (STGa), followed by clusters centered within the lingual gyrus and the sLOC, extending to the SMGp.

Smaller clusters also were found in posterior portions of the right cingulate gyrus and MTG, as well as in the MFG. Only one LH cluster, peaking in the central operculum, extending to the planum temporale and Heschl's gyrus was found. Bilateral activation was also found in the frontal pole.

**3.3.2. Patients (treatment and natural history groups)**—Data reported in this section do not include results from P14 because he developed a haematoma between baseline and post-testing assessments. Although no behavioral effects of it were noted, this event potentially altered neural processing.

**3.3.2.1. LESION PATTERNS.:** Fig. 8 illustrates lesion overlap in the left hemisphere for participants in the treatment (8a) and natural history (8b) groups. Regions with lesioned tissue are identified in Table 5, showing those with the most to the least overlap across participants in left anterior and posterior language regions as well as the proportion of lesioned tissue within region. Participants in the treatment and natural history group showed similar proportion of lesioned tissue in both left anterior ( $t = -.255$ , nsec) and left posterior regions ( $t = 1.076$ , nsec). Lesion overlap also was observed in both participant groups in the left primary auditory cortex and planum temporale (17/18 participants, with an average of 85% lesioned tissue within these regions), and left pre and post-central gyri (12/18 participants, with 24% of tissue within these regions lesioned).

Table 6 summarizes lesion overlap in regions within the left SPN and DAN for both participant groups. Lesion volume within these networks did not significantly differ between the treatment and the natural history group (SPN:  $t = -.827$ , nsec; DAN:  $t = -.052$ , nsec).

**3.3.2.2. BASELINE TO POST-TEST ACTIVATION: GROUP-LEVEL WHOLE BRAIN ANALYSES.:** Results of a one-sample t-test for the treatment group are displayed in Fig. 8 and Table 7, for the (*Passive > Control, POST > BL*) contrast. Clusters of upregulation were found almost exclusively in the RH, except for one cluster peaking in the L cuneus and extending anteriorly to the precuneus. In the RH, the largest cluster peaked in the SPL and extended to the AG, SMGp, and sLOC; upregulation was also found in the anterior portion of the SMG and the postcentral gyrus. One cluster of upregulation also was found in the anterior regions of the RH, with activation peaking in the MFG and extending to the PCG. No significant clusters of downregulation were found. Group-level analyses conducted on the natural history group yielded no significant changes in activation at post-testing.



**3.3.2.3. ACTIVATION SHIFTS WITHIN SPN AND DAN.:** Network/ROI analyses were conducted by extracting the proportion of upregulated tissue within SPN and DAN ROIs, bilaterally. Only ROIs with at least 10% intact tissue and at least 5 upregulated voxels were included. Analysis of changes in activation (i.e., POST > BL) in regions within SPN and DAN showed no significant *Group\*Hemisphere* interaction. However, for SPN, follow-up comparisons indicated a trend toward an effect of Group in the RH ( $z = 1.902, p = .057$ ), with greater upregulation of activation for participants in the treatment (vs natural history) group. For DAN, the results showed a significant effect of Group in both hemispheres (LH:  $z = 2.447, p = .014$ ; RH:  $z = 2.583, p = .01$ ), where upregulation from pre-to post-testing was greater for participants in the treatment (vs natural history) group.

**3.3.2.4. RELATION BETWEEN TREATMENT GAINS AND SPN AND DAN ACTIVATION.:** To further examine the relation between baseline to post-test changes in sentence processing and shifts in activation, change scores (z-score) in offline comprehension accuracy (comprehension full probe scores, FP structures only) were used. For online eyetracking, changes in TP and TI scores were entered into the analyses. For these analyses, data from both the treatment and natural history group were included.

For offline comprehension, a significant *Change\*Hemisphere* interaction was found in both networks (SPN:  $z = 9.788, p < .0001$ ; DAN:  $z = 26.588, p < .0001$ ). Follow-up analyses revealed, for SPN, that upregulation in RH ROIs was positively associated with changes in offline comprehension ( $z = 2.155, p = .031$ ), indicating that participants with larger (vs smaller) offline changes showed greater upregulation in RH ROIs. No significant association was found in LH ROIs within the SPN (see Fig. 9a, showing a trend in the opposite direction). For DAN, upregulation in the RH ROIs also was positively associated with offline change scores ( $z = 2.18, p = .029$ ), indicating that upregulation was greater in participants with larger (vs smaller) baseline to post-test changes; again, no significant relation was found between activation in the LH ROIs within the DAN and treatment-induced improvement (Fig. 9b).

Fig. 10 shows the results of changes in online comprehension conducted for the TP and TI scores separately. A significant *Prediction\*Hemisphere* interaction was found in both networks (SPN:  $z = -4.462, p < .0001$ ; DAN:  $z = -28.032, p < .0001$ ), with follow-up comparisons indicating that a decrease (from baseline to post-testing) in target fixations within the N1 region (i.e., a shift toward more normal-like eye movements, reflecting an Agent-first strategy) was associated with upregulation of RH ROIs within the SPN ( $z = -2.996, p = .003$ , see Fig. 10a); whereas, no significant relation was found for LH ROIs ( $z = -1.662, nsec$ ), although a trend in the same direction can be seen in Fig. 10a. Within the DAN, none of the follow-up comparisons reached significance (RH:  $z = -1.342, nsec$ ; LH:  $z = -.016, nsec$ , Fig. 10b). Analysis of TI scores also showed an *Integration\*Hemisphere* interaction in both networks (SPN:  $z = 5.471, p < .0001$ ; DAN:  $z = 4.444, p < .0001$ ). Follow-up comparisons did not reach significance for RH or LH ROIs in the SPN (RH:  $z = .995, nsec$ ; LH:  $z = -1.829, nsec$ ), although Fig. 10c shows a trend toward a negative association in the LH SPN ROIs. Conversely, in the DAN, an increase in target fixations from the V to the S.End region (i.e., a shift toward more normal-like processing) was

positively associated with upregulation in the RH ( $z = 2.068$ ,  $p = .038$ ), but not the LH, ROIs ( $z = -.606$ , nsec, see Fig. 10d).

Additional analyses examining the relation between LH lesion volume within SPN and DAN ROIs using mixed-effects regression analyses demonstrated that changes in offline comprehension accuracy were not predicted by lesion volume within either network: SPN, ( $z = .128$ , nsec; DAN,  $z = -1.171$ , nsec). Similarly, changes in measures of online sentence comprehension were not associated with lesion size within the SPN (TP:  $z = -.403$ , nsec; TI:  $z = -.582$ , nsec) or the DAN (TP:  $z = -.456$ , nsec; TI:  $z = -.664$ , nsec). However, lesion volume was a significant predictor of changes in activation in both networks: for the SPN, the relation between upregulation and lesion size was positive ( $z = 7.308$ ,  $p < .0001$ ), indicating greater upregulation for participants with larger (*vs* smaller) lesions; for the DAN, the relation was negative ( $z = -34.743$ ,  $p < .0001$ ), indicating more upregulation for participants with smaller (*vs* larger) lesions.

#### 4. Discussion

The present study investigated the impact of linguistically-based treatment of sentence production and comprehension (Treatment of Underlying Forms, TUF, Thompson & Shapiro, 2005, 2007) on neurocognitive mechanisms of sentence processing in individuals with chronic agrammatic aphasia. The aim was to shed light on patterns of neural re-organization, as indexed by fMRI activation, and cognitive processes, as revealed by eye movements, in treated and untreated patients, presenting with similar demographics, lesion and language deficit profiles. We also aimed to inform current theories of aphasia recovery, addressing the role of the right hemisphere (RH) during recovery as well as that of language domain-specific, sentence processing, and domain-general neural networks.

Results showed improved offline production of trained items, for all but one participant (P12), reflected in acquisition curves derived from weekly testing of trained structures, and corresponding large effect sizes. Notably, P12 presented with a severe pre-treatment language deficit, as demonstrated by low accuracy in production of nouns (44% correct) and verbs (56% correct) on the Northwestern Naming Battery (NNB; Thompson & Weintraub, 2014), as well as of canonical sentences on the Northwestern Assessment of Verbs and Sentences (NAVS; Thompson, 2011), Sentence Production Priming Test (SPPT; 7% correct). Further, his Western Aphasia Battery, Aphasia Quotient (WAB AQ) was 56. As noted by Ballard and Thompson (1999), patients with WAB AQ scores below 60 show poorer acquisition and generalization effects of TUF as compared to patients with WAB AQs above 60, suggesting that more severe language impairments may preclude substantial benefit from TUF. In line with this observation, P11, who showed the smallest change in production of trained items (30% pre-to post-treatment improvement) also had a WAB AQ under 60 (57.2). We note, however, that P4 and P7 also evinced low AQs (53.5 and 52.8, respectively) but, nevertheless, showed substantial treatment gains, indicating that lower language ability may not be the only variable related to successful response to sentence processing treatment.

Group-level analyses of the production data further indicated significant generalization to untrained structures for the treatment, but not the natural history group, including full



indicates that participants did not learn a non-linguistic, Theme-first strategy during treatment; if this were the case, a decrease in accuracy from baseline to post-treatment would have been expected. Finally, in line with CATE, the treatment group showed no improvement in comprehension of object-cleft structures, as predicted, in that these structures are linguistically unrelated to trained, passive structures.

Analysis of the online eye movement data indicated, in fact, that the treated participants did not learn a Theme-first strategy, rather they showed pre-to post-treatment emergence of an Agent-first strategy, reflecting normal thematic prediction. Briefly, following treatment, when listening to passive sentences, such as '*The man was shaved by the boy*', participants' eyes fixated on the incorrect picture in anticipation (predicting) that the first NP would be the Agent. Notably, participants did not use this strategy prior to treatment. Further downstream, evidence of improvement in thematic integration processes was also noted as fixations to the correct picture, depicting the Theme as sentential subject, increased and fixations to the incorrect picture decreased, reflecting re-assignment of the Theme to the first NP, assignment of the Agent to the second NP, and successful integration of the two. These results indicate a shift toward the eye movement patterns observed in healthy participants, supporting results reported by Mack and Thompson (2017; also see Mack, Nerantzini, & Thompson, 2017 for a similar finding using a visual world, eyetracking paradigm to monitor treatment-induced changes in online sentence production). Analyses of the eye movements of the natural history group did not show significant changes in prediction or integration scores, although the individual data showed a trend toward emergence of thematic prediction.

Participants who received treatment also showed shifts from pre-to post-treatment in neural activation, with the exception of P12, the only participant who also showed no treatment-induced improvement in sentence comprehension or production. Overall, summary activation maps indicated that the treatment, but not the natural history, group evinced significant upregulation in right hemisphere regions in both posterior and anterior brain regions. Notably, these regions showed overlap with those recruited by the healthy participants. For the *Passive* > *Control* contrast, the healthy participants showed bilateral activation in the superior and inferior parietal lobule (i.e., left supramarginal and right angular gyri) and the lateral occipital cortex and left hemisphere recruitment of the middle frontal and precentral gyri; whereas, the patients showed post-treatment > pre-treatment activation in these regions only in the right hemisphere. These findings indicate that the patients recruited right hemisphere regions activated by the healthy control participants as well as right hemisphere homologs of left-brain areas engaged by healthy listeners. Interestingly, upregulation in these areas was also found in a previous study by our group (Thompson and den Ouden et al., 2010) following training of object-relative sentences. However, in the present study, the healthy controls showed more widespread activation in both posterior and anterior regions, including the left inferior and superior temporal gyri, right middle and inferior temporal gyri, bilateral fusiform and the left pars opercularis and triangularis of the inferior frontal gyrus for passive sentence processing.

Placing these activation patterns within the context of psycholinguistic processes engaged for sentence comprehension, these findings support models implicating a frontal-parietal-

temporal network. The inferior frontal gyrus often is associated with initial syntactic parsing and phrase structure building operations (Friederici, 2012). The present data (and that of others) suggest that this process may also enlist the left middle frontal region in neurotypical adults and its right hemisphere homolog when the left hemisphere is damaged (Shetreet, Friedmann, & Hadar, 2010; Shetreet, Palti, Friedmann, & Hadar, 2007; Thompson, Bonakdarpour, & Fix, 2010; Thompson et al., 2007; also see Walenski et al., 2019, for meta-analysis). Consistent with our findings, several studies also have implicated tissue within the inferior parietal lobule, bilaterally, in both young and older unimpaired adults (Thompson et al., 2007, Thompson and Bonakdarpour et al., 2010) and unilaterally (in the right hemisphere) in aphasic patients (Thompson and den Ouden et al., 2010; see Thompson & Meltzer-Asscher, 2014, for review) for lexical-semantic processing, including verb argument mapping, during sentence comprehension. Finally, superior and middle temporal regions are associated with semantic-syntactic integration and are often active during processing of syntactically complex sentences, including passive structures (Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003; Bornkessel-Schlesewsky, Schlewewsky, & von Cramon, 2009; Crinion, Warburton, Lambon-Ralph, Howard, & Wise, 2006; Europa et al., 2019; Fiebach, Friederici, Muller, & von Cramon, 2002; Mack et al., 2013; Thompson and den Ouden et al., 2010; also see; Vigneau et al., 2011; Walenski et al., 2019). In the present study, the healthy controls showed activation in temporal lobe regions as seen in previous studies. However, the patients who received treatment did not show upregulation in either the superior or middle temporal gyri. Notably, however, both the healthy and patient groups showed activation in the lateral occipital cortex, which is adjacent to the middle temporal gyrus. The healthy controls showed bilateral activation in this region; whereas the patients recruited the right lateral occipital region. In their meta-analysis of neuroimaging studies of sentence processing in unimpaired participants, Walenski et al. (2019) also found bilateral temporo-occipital activation for both auditory and visual sentence comprehension.

The healthy controls and the patients also engaged the superior parietal lobule for passive sentence processing, again with bilateral activation in this region for the healthy controls and unilateral (right) hemisphere upregulation in the patients who underwent treatment. Although not typically associated with sentence processing per se, the superior parietal lobule is part of domain-general processing networks, including attention, visuospatial processing and working memory (Colby & Goldberg, 1999; Koenigs, Barby, Postle, & Grafman, 2009; Vandenberghe & Gillebert, 2009). This finding suggests that restoration of linguistic processes underlying comprehension of syntactically complex sentences is supported by right hemisphere domain-general processes. This is not surprising if we consider that TUF is explicitly focused on processes of thematic role assignment and syntactic mapping (supported by regions in the syntactic processing network), that rely on strategies that engage sustained attention and working memory resources (supported by regions in the dorsal attention network).

This latter conjecture was born out in the results of our ROI analyses that indicated significantly greater upregulation in the treatment than in the natural history group in regions within both the sentence processing and dorsal attention networks. In addition, composite z-scores reflecting offline behavioral gains resulting from treatment were positively related to upregulation in both networks. Crucially, however, activation shifts within these networks

were associated with unique aspects of improved sentence processing as measured by eye movements. Upregulation within regions of the sentence processing network, but not within the dorsal attention network, was associated with improved thematic prediction as shown by post-treatment, but not pre-treatment, use of an Agent-first strategy (i.e., fixations on the picture depicting the first NP heard in the sentence as Agent of the action) in early stages of online sentence processing. Conversely, upregulation in right hemisphere regions within the domain-general network was associated with improved thematic integration scores as indexed by eye movements to and fixation on the correct picture from the verb region (i.e., when verb morphology is processed) to the post-sentence region.

These findings provide compelling evidence that treatment impacts the processing strategies used to solve sentences, resulting in the use of more normal-like online thematic processes and concomitant recruitment of associated neural tissue that supports these processes. Successful outcomes following TUF are associated with improved ability to predict thematic roles early on during auditory sentence comprehension and to integrate thematic roles into the syntax downstream of the verb. The former engages brain structures that – in a healthy brain – are known to be engaged in sentence processing and support mechanisms of phrase structure building and thematic role assignment; whereas the latter, which arguably requires cognitive control mechanisms to maintain linguistic material (e.g., thematic roles) in auditory working memory, for ‘error monitoring’ and thematic/syntactic revision processes as necessary, as well as for spatial attention and execution of eye movements, are supported, at least in part, by the dorsal attention network – part of the Multiple-Demand System (Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Duncan & Owen, 2000; Fedorenko et al., 2013; Geranmayeh et al., 2014; King & Kutas, 1995; Kuperberg, Sitnikova, & Lakshmanan, 2008; Muggleton et al., 2011; Vernet et al., 2014; Vossel, Geng, & Fink, 2014).

We advance these interpretations with some caution in that participants in the natural history group showed improved thematic prediction from baseline to follow-up. However, minimal activation changes for untreated participants were found, particularly in the right hemisphere. Therefore, it is likely that the relation between improved thematic prediction and upregulation of right hemisphere homologs of regions within the normal sentence processing network reflected the effects of treatment, although the possibility that repeated exposure to the linguistic stimuli also impacted thematic prediction cannot be completely ruled out. In addition, with regard to thematic integration, it is possible that the positive relation between upregulation in the dorsal attention network and changes in sentence processing found in the treatment group may stem from the treatment provided, which required spatial manipulation of word cards to form target sentences. Thus, individuals may have learned a spatial strategy for arranging cards, thereby engaging (at post- but not pre-treatment) the dorsal attention network. This explanation, however, does not explain participants’ concomitant ability to comprehend and/or produce active transitive sentence structures.

A final issue that deserves discussion is the role of lesion volume in language recovery. As reported in the introduction, research on the predictors of recovery has often claimed that better recovery may be associated with smaller lesions, possibly due to preservation of a

greater amount of perilesional/ipsilesional tissue, making it available for recruitment into the language network (Ansaldo, Arguin, & Lecours, 2002; Fridriksson et al., 2012; Hillis, 2007). Results of our study indicate that this is likely an oversimplification of a complex interaction between lesion volume and location. We found that participants with larger (*vs* smaller) lesions within the left hemisphere sentence processing network showed more upregulation of neural activity within right hemisphere homologs of the sentence processing network, indicating a positive relation between lesion volume and recovery. Conversely, our finding that participants with smaller (*vs* larger) lesion volume within the left dorsal attention network showed greater upregulation of activation within this network in the right hemisphere indicates a negative relation between lesion volume and recovery. Although seemingly counterintuitive, these findings indicate that participants who show greater upregulation in both networks (and who most benefited from treatment) presented with large lesions within the sentence processing network, but relatively spared tissue within regions implicated in domain-general processes in the left hemisphere. Given that upregulation within these regions was observed in the right hemisphere, this result suggests that right hemisphere regions are more likely recruited when the sentence processing network in the left hemisphere is largely lesioned, but regions supporting domain-general processes are largely spared.

Overall, these findings indicate that regions within the right hemisphere region are viable for the support of language recovery in patients with left-hemisphere stroke and sentence-level deficits, in line with other studies focused on treatment-induced neuroplasticity (Breier et al., 2009; Cherney & Small, 2006; Kiran et al., 2015; Raboyeau et al., 2008; Thompson et al., 2013, Thompson and den Ouden et al., 2010; Wan et al., 2014; also see; Kiran & Thompson, 2019). Notably, the present results show that changes in right hemisphere activation following treatment, particularly for recovery of sentence processing, reflect its adaptive role in neurocognitive processing, rather than a maladaptive role (*i.e.*, possibly due to impaired inhibitory interhemispheric connections as suggested by Heiss & Thiel, 2006, and others).

## 5. Conclusion

The present findings show that linguistically-based treatment of sentence processing [*i.e.*, Treatment of Underlying Forms (TUF; Thompson & Shapiro, 2005, 2007)] improves both offline comprehension and production of complex and simple sentences, as found in previous studies, and results in more normal-like online sentence processing as reflected by the emergence of thematic prediction and thematic integration processes in patients with chronic agrammatic aphasia. Notably, both offline and online treatment-induced changes in sentence processing correlate positively with changes in neural activation (from baseline to post-testing) in right hemisphere regions homologous to left brain regions activated by healthy individuals for sentence processing, with improvements in thematic prediction and thematic integration associated with upregulation in regions within the neural network for sentence processing and domain-general processes, respectively. The study provides evidence for an adaptive role of right hemisphere regions in recovery from sentence processing deficits, and suggests that greater lesions within left hemisphere language-specific networks, but lesser domain-general network damage is associated with recruitment of right hemisphere networks that support recovery.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The work reported here is part of a larger, multi-site project examining the neurobiology of language recovery in people with aphasia (NIH-NIDCD Clinical Research Center grant P50DC012283, PI: Cynthia K. Thompson). No part of the study procedures or analyses was pre-registered prior to the research being undertaken.

We are grateful to Dr. Aya Meltzer-Asscher for her contributions to task development and pre-testing, to Sarah Chandler for assistance with data collection and language training, and to James Higgins for helpful discussions on the analysis of neuroimaging data. Finally, we thank Amanda Salman for her assistance with manuscript preparation.

Declaration of interest

CKT serves as an associate editor for *Cortex*. The remaining authors wish to confirm that there are no known conflicts of interest associated with this publication and that the funding source did not sustain any role in the design, management, analysis, and interpretation of the data, as well as in the preparation, review, or approval of the manuscript.

## References

- Abel S, Weiller C, Huber W, & Willmes K (2014). Neural underpinnings for model-oriented therapy of aphasic word production. *Neuropsychologia*, 57, 154–165. 10.1016/j.neuropsychologia.2014.03.010. [PubMed: 24686092]
- Abel S, Weiller C, Huber W, Willmes K, & Specht K (2015). Therapy-induced brain reorganization patterns in aphasia. *Brain*, 138(4), 1097–1112. 10.1093/brain/awv022. [PubMed: 25688082]
- Alpert K, Kogan A, Parrish T, Marcus D, & Wang L (2016). The northwestern university neuroimaging data archive (NUNDA). *Neuroimage*, 124, 1131–1136. 10.1016/j.neuroimage.2015.05.0600. [PubMed: 26032888]
- Ansaldo AI, Arguin M, & Lecours AR (2002). The contribution of the right cerebral hemisphere to the recovery from aphasia: A single longitudinal case study. *Brain and Language*, 82(2), 206–222. 10.1016/S0093-934X(02)00017-2. [PubMed: 12096877]
- Baciu M, Acher A, Kauffmann L, Cousin E, Boilley C, Hueber T, et al. (2016). Effect of visual feedback on speech recovery and language plasticity in patients with post-stroke non-fluent aphasia. Functional MRI assessment. *Annals of Physical and Rehabilitation Medicine*, 59, e75–e76. 10.1016/j.rehab.2016.07.176.
- Bakheit AMO, Shaw S, Carrington S, & Griffiths S (2007). The rate and extent of improvement with therapy from the different types of aphasia in the first year after stroke. *Clinical Rehabilitation*, 21(10), 941–949. 10.1177/02692155070784522. [PubMed: 17981853]
- Ballard KJ, & Thompson CK (1999). Treatment and generalization of complex sentence structures in agrammatism. *Journal of Speech, Language, and Hearing Research*, 42, 690–707. 10.1044/jslhr.4203.690.
- Barwood CH, Murdoch BE, Whelan BM, Lloyd D, Riek S, O’Sullivan J, et al. (2011). The effects of low frequency Repetitive Transcranial Magnetic Stimulation (rTMS) and sham condition rTMS on behavioural language in chronic non-fluent aphasia: Short term outcomes. *Neurorehabilitation*, 28(2), 113–128. 10.3233/NRE-2011-0640. [PubMed: 21447912]
- Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. arXiv:1406.5823 [stat.CO].
- Belin P, Zilbovicius M, Remy P, Francois C, Guillaume S, Chain F, et al. (1996). Recovery from nonfluent aphasia after melodic intonation therapy A PET study. *Neurology*, 47(6), 1504–1511. 10.1212/WNL.47.6.1504. [PubMed: 8960735]



- Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, & Grodzinsky Y (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science*, 14(5), 433–440. 10.1111/1467-9280.01459. [PubMed: 12930473]
- Ben-Shachar M, Palti D, & Grodzinsky Y (2004). Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *Neuroimage*, 21(4), 1320–1336. 10.1016/j.neuroimage.2003.11.027. [PubMed: 15050558]
- Benjamini Y, & Hochberg Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300. [www.jstor.org/stable/2346101](http://www.jstor.org/stable/2346101).
- Berwick R, & Weinberg A (1984). *The grammatical basis of linguistic performance: Language use and language acquisition*. Cambridge, MA: MIT Press.
- Binder JR, Westbury CF, McKiernan KA, Possing ET, & Medler DA (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, 17(6), 905–917. 10.1162/08989290540211102. [PubMed: 16021798]
- Blasi V, Young AC, Tansy AP, Petersen SE, Snyder AZ, & Corbetta M (2002). Word retrieval learning modulates right frontal cortex in patients with left frontal damage. *Neuron*, 36(1), 159–170. 10.1016/S0896-6273(02)00936-4. [PubMed: 12367514]
- Bonakdarpour B, Parrish TB, & Thompson CK (2007). Hemodynamic response function in patients with stroke-induced aphasia: Implications for fMRI data analysis. *Neuroimage*, 36(2), 322–331. 10.1016/j.neuroimage.2007.02.035. [PubMed: 17467297]
- Bonilha L, Gleichgerrcht E, Nesland T, Rorden C, & Fridriksson J (2016). Success of anomia treatment in aphasia is associated with preserved architecture of global and left temporal lobe structural networks. *Neurorehabilitation and Neural Repair*, 30(3), 266–279. 10.1177/15459683155938088. [PubMed: 26150147]
- Bornkessel-Schlesewsky I, Schlesewsky M, & von Cramon DY (2009). Word order and Broca's region: Evidence for a supra-syntactic perspective. *Brain and Language*, 111(3), 125–139. 10.1016/j.bandl.2009.09.004. [PubMed: 19853290]
- Breier JI, Juranek J, Maher LM, Schmadeke S, Men D, & Papanicolaou AC (2009). Behavioral and neurophysiologic response to therapy for chronic aphasia. *Archives of Physical Medicine and Rehabilitation*, 90(12), 2026–2033. 10.1016/j.apmr.2009.08.144. [PubMed: 19969164]
- Breier JI, Maher LM, Novak B, & Papanicolaou AC (2006). Functional imaging before and after constraint-induced language therapy for aphasia using magnetoencephalography. *Neurocase*, 12(6), 322–331. 10.1080/13554790601126054. [PubMed: 17182395]
- Brownsett SL, Warren JE, Geranmayeh F, Woodhead Z, Leech R, & Wise RJ (2013). Cognitive control and its impact on recovery from aphasic stroke. *Brain*, 137(1), 242–254. 10.1093/brain/awt289. [PubMed: 24163248]
- Caplan D, Chen E, & Waters G (2008). Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex*, 44(3), 257–275. 10.1016/j.cortex.2006.06.005. [PubMed: 18387556]
- Cherney LR, & Small SL (2006). Task-dependent changes in brain activation following therapy for nonfluent aphasia: Discussion of two individual cases. *Journal of the International Neuropsychological Society*, 12(6), 828–842. 10.1017/S1355617706061017. [PubMed: 17064446]
- Chomsky N (1995). Categories and transformations. *The Minimalist Program*, 219, 394.
- Cohen J (1988). *Statistical power analysis for the social sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Colby CL, & Goldberg ME (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22(1), 319–349. 10.1146/annurev.neuro.22.1.319.
- Coltheart M (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, 33(4), 497–505.
- Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, & Shulman GL (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3(3), 292. 10.1038/73009. [PubMed: 10700263]

- Corbetta M, Patel G, & Shulman GL (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58(3), 306–324. 10.1016/j.neuron.2008.04.017. [PubMed: 18466742]
- Corbetta M, & Shulman GL (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201–215. 10.1038/nrn755. [PubMed: 11994752]
- Cox RW (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29, 162–173. [PubMed: 8812068]
- Crinion JT, Warburton EA, Lambon-Ralph MA, Howard D, & Wise RJ (2006). Listening to narrative speech after aphasic stroke: The role of the left anterior temporal lobe. *Cerebral Cortex*, 16(8), 1116–1125. 10.1093/cercor/bhj053. [PubMed: 16251507]
- Crosson B, Moore AB, McGregor KM, Chang YL, Benjamin M, Gopinath K, et al. (2009). Regional changes in word-production laterality after a naming treatment designed to produce a rightward shift in frontal activity. *Brain and Language*, 111(2), 73–85. 10.1016/j.bandl.2009.08.001. [PubMed: 19811814]
- Curtis CE, & D’Esposito M (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423. 10.1016/S1364-6613(03)00197-9. [PubMed: 12963473]
- Davies M (2008). *The corpus of contemporary American English. BYE*. Brigham Young University.
- Demeurisse G, Demol O, Derouck M, De Beuckelaer R, Coekaerts MJ, & Capon A (1980). Quantitative study of the rate of recovery from aphasia due to ischemic stroke. *Stroke*, 11(5), 455–458. 10.1161/01.STR.11.5.455. [PubMed: 7423574]
- Den Ouden DB, Saur D, Mader W, Schelter B, Lukic S, Wali E, et al. (2012). Network modulation during complex syntactic processing. *Neuroimage*, 59(1), 815–823. 10.1016/j.neuroimage.2011.07.057. [PubMed: 21820518]
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, 31, 968–980. 10.1016/j.neuroimage.2006.01.021. [PubMed: 16530430]
- Duncan J, & Owen AM (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483. 10.1016/S0166-2236(00)01633-7. [PubMed: 11006464]
- Eklund A, Nichols TE, & Knutsson H (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113, 7900–7905. 10.1073/pnas.1602413113.
- Europa E, Gitelman D, Kiran S, & Thompson CK (2019). Neural connectivity in syntactic movement processing. *Frontiers in Human Neuroscience*, 13, 27. 10.3389/fnhum.2019.00027. [PubMed: 30814941]
- Fedorenko E, Duncan J, & Kanwisher N (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41), 16616–16621. 10.1073/pnas.1315235110.
- Fiebach CJ, Friederici AD, Müller K, & Cramon DYV (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, 14(1), 11–23. 10.1162/089892902317205285. [PubMed: 11798383]
- Forkel SJ, Thiebaut de Schotten M, Dell’Acqua F, Kalra L, Murphy DG, Williams SC, et al. (2014). Anatomical predictors of aphasia recovery: A tractography study of bilateral perisylvian language networks. *Brain*, 137(7), 2027–2039. 10.1093/brain/awu113. [PubMed: 24951631]
- Fridriksson J, Bonilha L, Baker JM, Moser D, & Rorden C (2009). Activity in preserved left hemisphere regions predicts anomia severity in aphasia. *Cerebral Cortex*, 20(5), 1013–1019. 10.1093/cercor/bhp160. [PubMed: 19687294]
- Fridriksson J, Richardson JD, Fillmore P, & Cai B (2012). Left hemisphere plasticity and aphasia recovery. *Neuroimage*, 60(2), 854–863. 10.1016/j.neuroimage.2011.12.057. [PubMed: 22227052]
- Fridriksson J, Rorden C, Elm J, Sen S, George MS, & Bonilha L (2018). Transcranial direct current stimulation vs sham stimulation to treat aphasia after stroke: A randomized clinical trial. *JAMA Neurology*, 75(12), 1470–1476. [PubMed: 30128538]

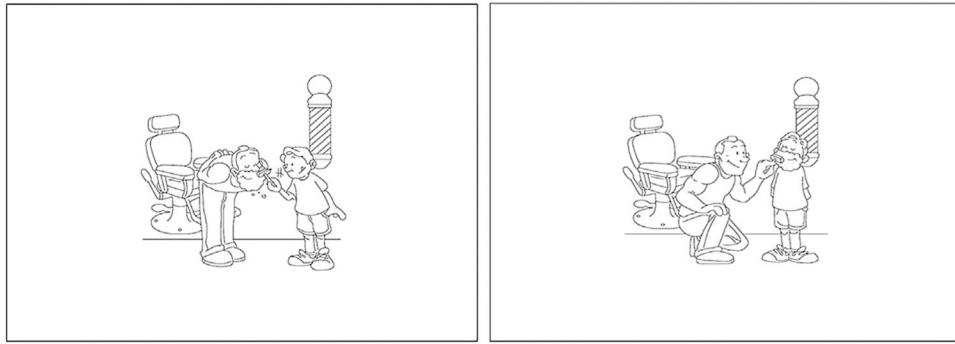
- Friederici AD (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268. 10.1016/j.tics.2012.04.001. [PubMed: 22516238]
- Gainotti G (2013). The contribution of Language to the right-hemisphere conceptual representations: A selective survey. *Journal of Clinical and Experimental Neuropsychology*, 35(6), 563–572. 10.1080/13803395.2013.798399. [PubMed: 23678989]
- Gainotti G (2015). Contrasting opinions on the role of the right hemisphere in the recovery of language. A critical survey. *Aphasiology*, 29(9), 1020–1037. 10.1080/02687038.2015.1027170.
- Geranmayeh F, Brownsett SL, & Wise RJ (2014). Task-induced brain activity in aphasic stroke patients: What is driving recovery? *Brain*, 137(10), 2632–2648. 10.1093/brain/awu163. [PubMed: 24974382]
- Grodzinsky Y, & Friederici AD (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology*, 16(2), 240–246. 10.1016/j.conb.2006.03.007. [PubMed: 16563739]
- Hartwigsen G, & Saur D (2019). Neuroimaging of stroke recovery from aphasia – Insights into plasticity of the human language network. *NeuroImage*, 190, 14–31. [PubMed: 29175498]
- Heiss WD, & Thiel A (2006). A proposed regional hierarchy in recovery of post-stroke aphasia. *Brain and Language*, 98(1), 118–123. 10.1016/j.bandl.2006.02.002. [PubMed: 16564566]
- Heiss WD, Thiel A, Kessler J, & Herholz K (2003). Disturbance and recovery of Language function: Correlates in PET activation studies. *Neuroimage*, 20, S42–S49. 10.1016/j.neuroimage.2003.09.005. [PubMed: 14597295]
- Hillis AE (2007). Aphasia progress in the last quarter of a century. *Neurology*, 69(2), 200–213. 10.1212/01.wnl.0000265600.69385.6f. [PubMed: 17620554]
- Hirotani M, Makuuchi M, Rüschemeyer SA, & Friederici AD (2011). Who was the agent? The neural correlates of reanalysis processes during sentence comprehension. *Human Brain Mapping*, 32(11), 1775–1787. 10.1002/hbm.21146. [PubMed: 21391256]
- Hope TM, Seghier ML, Leff AP, & Price CJ (2013). Predicting outcome and recovery after stroke with lesions extracted from MRI images. *Neuroimage: Clinical*, 2, 424–433. 10.1016/j.nicl.2013.03.005. [PubMed: 24179796]
- Hsu CJ, & Thompson CK (2018). Manual versus automated narrative analysis of agrammatic production patterns: The northwestern narrative language analysis and computerized Language Analysis. *Journal of Speech, Language, and Hearing Research*, 61(2), 373–385. 10.1044/2017\_JSLHR-L-17-0185.
- Jacobs B, & Thompson CK (2000). Cross-modal generalization effects of training noncanonical sentence comprehension and production in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research*, 43, 5–20. 10.1044/jslhr.4301.05.
- Jaeger TF (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446. 10.1016/j.jml.2007.11.007. [PubMed: 19884961]
- Jo HJ, Gotts SJ, Reynolds RC, Bandettini PA, Martin A, Cox RW, et al. (2013). Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state fMRI. *Journal of Applied Mathematics*, 2013. 10.1155/2013/935154.
- Jung-Beeman M (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9(11), 512–518. 10.1016/j.tics.2005.09.009. [PubMed: 16214387]
- Kertesz A (2007). *Western aphasia Battery*. San Antonio, TX: PsychCorp (Revised).
- Kielar A, Deschamps T, Jokel R, & Meltzer JA (2016). Functional reorganization of language networks for semantics and syntax in chronic stroke: Evidence from MEG. *Human Brain Mapping*, 37(8), 2869–2893. 10.1002/hbm.23212. [PubMed: 27091757]
- King JW, & Kutas M (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376–395. 10.1162/jocn.1995.7.3.376. [PubMed: 23961867]
- Kinno R, Kawamura M, Shioda S, & Sakai KL (2008). Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Human Brain Mapping*, 29(9), 1015–1027. 10.1002/hbm.20441. [PubMed: 17924553]

- Kiran S, Meier EL, Kapse KJ, & Glynn PA (2015). Changes in task-based effective connectivity in language networks following rehabilitation in post-stroke patients with aphasia. *Frontiers in Human Neuroscience*, 9 10.3389/fnhum.2015.00316.
- Kiran S, & Thompson CK (2019). Neuroplasticity of language networks in aphasia: Advances, updates and future challenges. *Frontiers in Neurology*, 10, 295 10.3389/fneur.2019.00295. [PubMed: 31001187]
- Koenigs M, Barbey AK, Postle BR, & Grafman J (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *Journal of Neuroscience*, 29(47), 14980–14986. 10.1523/JNEUROSCI.3706-09.2009. [PubMed: 19940193]
- Kuperberg GR, Sitnikova T, & Lakshmanan BM (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: Evidence from functional magnetic resonance imaging. *Neuroimage*, 40(1), 367–388. 10.1016/j.neuroimage.2007.10.009. [PubMed: 18248739]
- Kurland J, Baldwin K, & Tauer C (2010). Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic Wernicke’s aphasia. *Aphasiology*, 24(6–8), 737–751. 10.1080/02687030903524711.
- Laska AC, Hellblom A, Murray V, Kahan T, & Von Arbin M (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, 249(5), 413–422. 10.1046/j.1365-2796.2001.00812.x. [PubMed: 11350565]
- Lazar RM, Minzer B, Antonello D, Festa JR, Krakauer JW, & Marshall RS (2010). Improvement in aphasia scores after stroke is well predicted by initial severity. *Stroke*, 41(7), 1485–1488. 10.1161/STROKEAHA.109.577338. [PubMed: 20538700]
- Leger A, Demonet JF, Ruff S, Aithamon B, Touyeras B, Puel M, et al. (2002). Neural substrates of spoken language rehabilitation in an aphasic patient: An fMRI study. *Neuroimage*, 17(1), 174–183. 10.1006/nimg.2002.1238. [PubMed: 12482075]
- Maas MB, Lev MH, Ay H, Singhal AB, Greer DM, Smith WS, et al. (2012). The prognosis for aphasia in stroke. *Journal of Stroke and Cerebrovascular Diseases*, 21(5), 350–357. 10.1016/j.jstrokecerebrovasdis.2010.09.009. [PubMed: 21185744]
- Mack JE, Meltzer-Asscher A, Barbieri E, & Thompson CK (2013). Neural correlates of processing passive sentences. *Brain Sciences*, 3(3), 1198–1214. 10.3390/brainsci3031198. [PubMed: 24961525]
- Mack JE, Nerantzini M, & Thompson CK (2017). Recovery of sentence production processes following language treatment in aphasia: Evidence from eyetracking. *Frontiers in Human Neuroscience*, 11, 101 10.3389/fnhum.2017.00101. [PubMed: 28348524]
- Mack JE, & Thompson CK (2017). Recovery of online sentence processing in aphasia: Eye movement changes resulting from treatment of underlying forms. *Journal of Speech, Language, and Hearing Research*, 60(5), 1299–1315. 10.1044/2016\_JSLHR-L-16-0108.
- Mack JE, Wei AZS, Gutierrez S, & Thompson CK (2016). Tracking sentence comprehension: Test-retest reliability in people with aphasia and unimpaired adults. *Journal of Neurolinguistics*, 40, 98–111. 10.1016/j.jneuroling.2016.06.001. [PubMed: 27867260]
- Marcotte K, Adrover-Roig D, Damien B, de Preaumont M, Genereux S, Hubert M, et al. (2012). Therapy-induced neuroplasticity in chronic aphasia. *Neuropsychologia*, 50(8), 1776–1786. 10.1016/j.neuropsychologia.2012.04.001. [PubMed: 22564481]
- Martin PI, Naeser MA, Ho M, Doron KW, Kurland J, Kaplan J, et al. (2009). Overt naming fMRI pre- and post-TMS: Two nonfluent aphasia patients, with and without improved naming post-TMS. *Brain and Language*, 111(1), 20–35. 10.1016/j.bandl.2009.07.007. [PubMed: 19695692]
- Mattioli F, Ambrosi C, Mascaro L, Scarpazza C, Pasquali P, Frugoni M, et al. (2014). Early aphasia rehabilitation is associated with functional reactivation of the left inferior frontal gyrus. *Stroke*, 45(2), 545–552. 10.1161/STROKEAHA.113.003192. [PubMed: 24309584]
- McGlone J (1977). Sex differences in the cerebral organization of verbal functions in patients with unilateral brain lesions. *Brain*, 100(4), 775–793. 10.1093/brain/100.4.775. [PubMed: 608120]
- Meinzer M, Flaisch T, Breitenstein C, Wienbruch C, Elbert T, & Rockstroh B (2008). Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. *Neuroimage*, 39(4), 2038–2046. 10.1016/j.neuroimage.2007.10.008. [PubMed: 18096407]

- Meyer AM, Mack JE, & Thompson CK (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, 25(1), 31–43. 10.1016/j.jneuroling.2011.08.001. [PubMed: 22043134]
- Mohr B, Difrancesco S, Harrington K, Evans S, & Pulvermüller F (2014). Changes of right-hemispheric activation after constraint-induced, intensive language action therapy in chronic aphasia: fMRI evidence from auditory semantic processing. *Frontiers in Human Neuroscience*, 8, 919. 10.3389/fnhum.2014.00919. [PubMed: 25452721]
- Muggleton NG, Kalla R, Juan CH, & Walsh V (2011). Dissociating the contributions of human frontal eye fields and posterior parietal cortex to visual search. *Journal of Neurophysiology*, 105, 2891–2896. 10.1152/jn.01149.2009. [PubMed: 21490286]
- Nachev P, Coulthard E, Jäger HR, Kennard C, & Husain M (2008). Enantiomorphic normalization of focally lesioned brains. *Neuroimage*, 39(3), 1215–1226. 10.1016/j.neuroimage.2007.10.002. [PubMed: 18023365]
- Naeser MA, Helm-Estabrooks N, Haas G, Auerbach S, & Srinivasan M (1987). Relationship between lesion extent in ‘Wernicke’s area’ on computed tomographic scan and predicting recovery of comprehension in Wernicke’s aphasia. *Archives of Neurology*, 44(1), 73–82. 10.1001/archneur.1987.00520130057018. [PubMed: 3800725]
- Nardo D, Holland R, Leff AP, Price CJ, & Crinion JT (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, 140(11), 3039–3054. 10.1093/brain/awx234. [PubMed: 29053773]
- Norise C, & Hamilton RH (2017). Non-invasive brain stimulation in the treatment of post-stroke and neurodegenerative aphasia: Parallels, differences, and lessons learned. *Frontiers in Human Neuroscience*, 10, 675. 10.3389/fnhum.2016.00675. [PubMed: 28167904]
- Pedersen PM, Vinter K, & Olsen TS (2004). Aphasia after stroke: Type, severity and prognosis. *Cerebrovascular Diseases*, 17(1), 35–43. 10.1159/000073896.
- Perlmutter DM (1978). Impersonal passives and the unaccusative hypothesis. In *Annual meeting of the Berkeley linguistics society* (Vol. 4, pp. 157–190).
- Postman-Caucheteux WA, Birm RM, Pursley RH, Butman JA, Solomon JM, Picchioni D, et al. (2010). Single-trial fMRI shows contralesional activity linked to overt naming errors in chronic aphasic patients. *Journal of Cognitive Neuroscience*, 22(6), 1299–1318. 10.1162/jocn.2009.21261. [PubMed: 19413476]
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, & Petersen SE (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59, 2142–2154. 10.1016/j.neuroimage.2011.10.018. [PubMed: 22019881]
- R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing <https://www.R-project.org/>.
- Raboyeau G, De Boissezon X, Marie N, Balduyck S, Puel M, Bezy C, et al. (2008). Right hemisphere activation in recovery from aphasia lesion effect or function recruitment? *Neurology*, 70(4), 290–298. 10.1212/01.wnl.0000287115.85956.87. [PubMed: 18209203]
- Rorden C, & Brett M (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, 12(4), 191–200. 10.1155/2000/421719. [PubMed: 11568431]
- Sandberg CW, Bohland JW, & Kiran S (2015). Changes in functional connectivity related to direct training and generalization effects of a word finding treatment in chronic aphasia. *Brain and Language*, 150, 103–116. 10.1016/j.bandl.2015.09.002. [PubMed: 26398158]
- Schlaug G, Marchina S, & Norton A (2009). Evidence for plasticity in white-matter tracts of patients with chronic broca’s aphasia undergoing intense intonation-based speech therapy. *Annals of the New York Academy of Sciences*, 1169(1), 385–394. 10.1111/j.1749-6632.2009.04587.x. [PubMed: 19673813]
- Shetreet E, Friedmann N, & Hadar U (2010). The neural correlates of linguistic distinctions: Unaccusative and unergative verbs. *Journal of Cognitive Neuroscience*, 22, 2306–2315. 10.1162/jocn.2009.21371. [PubMed: 19925202]
- Shetreet E, Palti D, Friedmann N, & Hadar U (2007). Cortical representation of verb processing in sentence comprehension: Number of complements, subcategorization, and thematic frames. *Cerebral Cortex*, 17(8), 1958–1969. 10.1093/cercor/bhl105. [PubMed: 17101687]

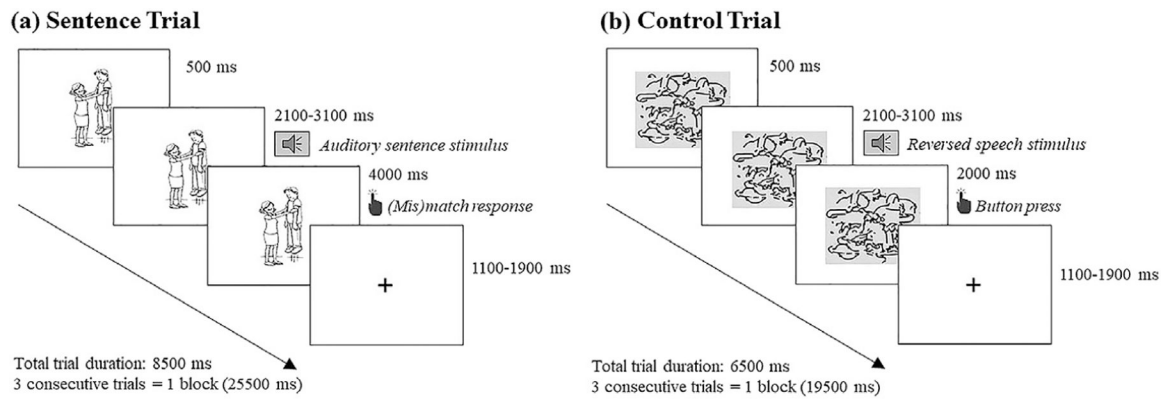
- Song X, Wang X, Alpert KA, Wang L, & Parrish TB (2015). Robust web-based parallel-optimized minimal preprocessing and analysis pipeline for fMRI big data In Paper presented at the 21st Meeting of the Organization for the Human Brain Mapping (OHBM). Honolulu, HI.
- Stadie N, Schröder A, Postler J, Lorenz A, Swoboda-Moll M, Burchert F, et al. (2008). Unambiguous generalization effects after treatment of non-canonical sentence production in German agrammatism. *Brain and Language*, 104(3), 211–229. 10.1016/j.bandl.2007.08.006. [PubMed: 17928044]
- Swisher JD, Halko MA, Merabet LB, McMains SA, & Somers DC (2007). Visual topography of human intraparietal sulcus. *Journal of Neuroscience*, 27, 5326–5337. 10.1523/JNEUROSCI.0991-07.2007. [PubMed: 17507555]
- Thiel A, Hartmann A, Rubi-Fessen I, Anglade C, Kracht L, Weiduschat N, et al. (2013). Effects of noninvasive brain stimulation on language networks and recovery in early poststroke aphasia. *Stroke*, 44(8), 2240–2246. 10.1161/STROKEAHA.111.000574. [PubMed: 23813984]
- Thompson CK (2006). Single subject controlled experiments in aphasia: The science and the state of the science. *Journal of Communication Disorders*, 39(4), 266–291. 10.1016/j.jcomdis.2006.02.003. [PubMed: 16635494]
- Thompson CK (2011). *Northwestern assessment of verbs and sentences (NAVS)*. Evanston, IL: Northwestern University.
- Thompson CK, Bonakdarpour B, & Fix SF (2010). Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. *Journal of Cognitive Neuroscience*, 22(9), 1993–2011. 10.1162/jocn.2009.21334. [PubMed: 19702460]
- Thompson CK, Bonakdarpour B, Fix SC, Blumenfeld HK, Parrish TB, Gitelman DR, et al. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19(11), 1753–1767. 10.1162/jocn.2007.19.11.1753. [PubMed: 17958479]
- Thompson CK, Cho S, Hsu CJ, Wieneke C, Rademaker A, Weitner BB, et al. (2012). Dissociations between fluency and agrammatism in primary progressive aphasia. *Aphasiology*, 26(1), 20–43. [PubMed: 22199417]
- Thompson CK, den Ouden DB, Bonakdarpour B, Garibaldi K, & Parrish TB (2010b). Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia*, 48(11), 3211–3227. 10.1016/j.neuropsychologia.2010.06.036. [PubMed: 20603138]
- Thompson CK, & Meltzer-Asscher A (2014). Neurocognitive mechanisms of verb argument structure processing In Bachrach A, Roy I, & Stockall L (Eds.), *Structuring the Argument: Multidisciplinary research on verb argument structure* (pp. 141–168). John Benjamins Publishing Company.
- Thompson CK, Riley EA, den Ouden DB, Meltzer-Asscher A, & Lukic S (2013). Training verb argument structure production in agrammatic aphasia: Behavioral and neural recovery patterns. *Cortex*, 49(9), 2358–2376. 10.1016/j.cortex.2013.02.003. [PubMed: 23514929]
- Thompson CK, & Shapiro L (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology*, 19(10–11), 1021–1036. 10.1080/02687030544000227. [PubMed: 17410280]
- Thompson CK, & Shapiro LP (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech and Language Pathology*, 16, 30–42.
- Thompson CK, Shapiro LP, Ballard K, Jacobs B, Schneider S, & Tait ME (1997). Training and generalized production of wh- and NP-movement structures in agrammatic speakers. *Journal of Speech, Language and Hearing Research*, 40, 228–244. 10.1044/jslhr.4002.228.
- Thompson CK, Shapiro LP, Kiran S, & Sobecks J (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research*, 46(3), 591–607. 10.1044/1092-4388(2003/047).
- Thompson CK, & Weintraub S (2014). *Northwestern naming Battery*. Evanston, IL: Northwestern University.
- Turkeltaub PE, Messing S, Norise C, & Hamilton RH (2011). Are networks for residual language function and recovery consistent across aphasic patients? *Neurology*, 76(20), 1726–1734. 10.1212/WNL.0b013-31821a44c1. [PubMed: 21576689]

- Vandenberghe R, & Gillebert CR (2009). Parcellation of parietal cortex: Convergence between lesion-symptom mapping and mapping of the intact functioning brain. *Behavioural Brain Research*, 199(2), 171–182. 10.1016/j.bbr.2008.12.005. [PubMed: 19118580]
- Vernet M, Quentin R, Chanes L, Mitsumasu A, & Valero-Cabré A (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, 8 10.3389/fnint.2014.00066.
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, et al. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, 54(1), 577–593. 10.1016/j.neuroimage.2010.07.036. [PubMed: 20656040]
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, & Buckner RL (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, 100, 3328–3342. 10.1152/jn.90355.2008. [PubMed: 18799601]
- Vossel S, Geng JJ, & Fink GR (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, 20(2), 150–159. 10.1177/1073858413494269. [PubMed: 23835449]
- Vossel S, Thiel CM, & Fink GR (2006). Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *Neuroimage*, 32, 1257–1264. 10.1016/j.neuroimage.2006.05.019. [PubMed: 16846742]
- Walenski M, Europa E, Caplan D, & Thompson CK (2019). An ALE-based meta-analysis of neuroimaging studies of sentence comprehension and production. *Human Brain Mapping*, 40, 2275–2304. [PubMed: 30689268]
- Wang B, Wang X, Wang Y, Katsaggelos AK, & Parrish TB (2019). An automated segmentation pipeline for chronic stroke lesion with new evaluation metrics. *Neuroimage: Clinical (under review)*.
- Wan CY, Zheng X, Marchina S, Norton A, & Schlaug G (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain and Language*, 136, 1–7. 10.1016/j.bandl.2014.03.011. [PubMed: 25041868]
- Weiduschat N, Thiel A, Rubi-Fessen I, Hartmann A, Kessler J, Merl P, et al. (2011). Effects of repetitive transcranial magnetic stimulation in aphasic stroke. *Stroke*, 42(2), 409–415. 10.1161/STROKEAHA.110.597864. [PubMed: 21164121]
- Wierenga CE, Maher LM, Moore AB, White KD, Mcgregor K, Soltysik DA, et al. (2006). Neural substrates of syntactic mapping treatment: An fMRI study of two cases. *Journal of the International Neuropsychological Society*, 12(1), 132–146. 10.1017/S135561770606019X. [PubMed: 16433953]
- Wright P, Stamatakis EA, & Tyler LK (2012). Differentiating hemispheric contributions to syntax and semantics in patients with left-hemisphere lesions. *Journal of Neuroscience*, 32(24), 8149–8157. 10.1523/JNEUROSCI.0485-12.2012. [PubMed: 22699896]

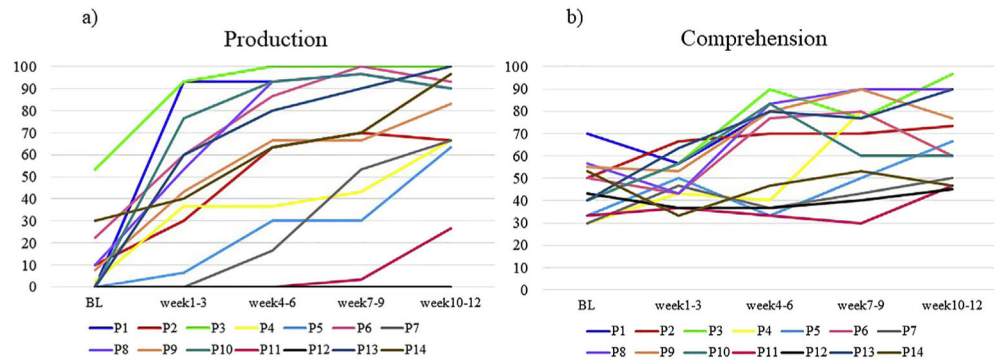


**Fig. 1 –.**  
Example pair of semantically reversible pictures used in treatment, probe tasks, and the comprehension fMRI task.

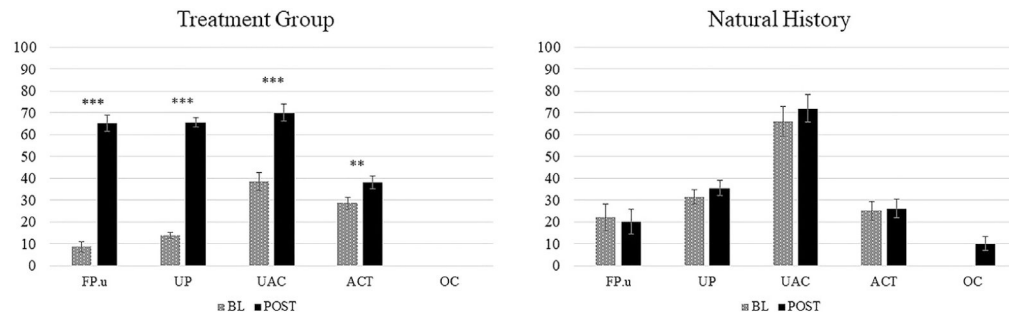




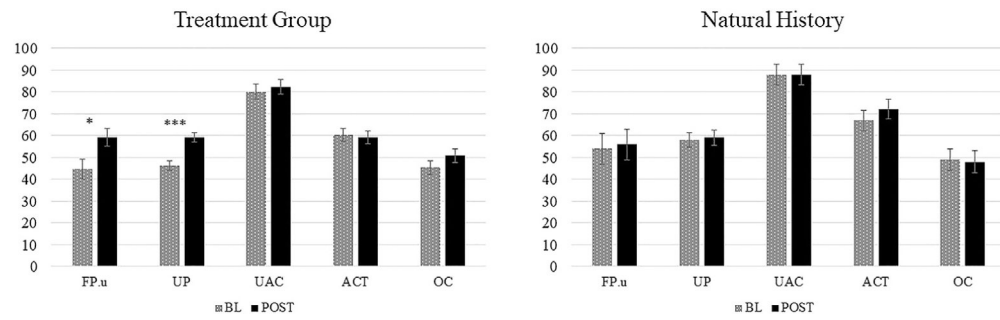
**Fig. 2 –.**  
Schematic representation of the fMRI protocol, for sentence (active, passive) trials (a) and control trials (b).



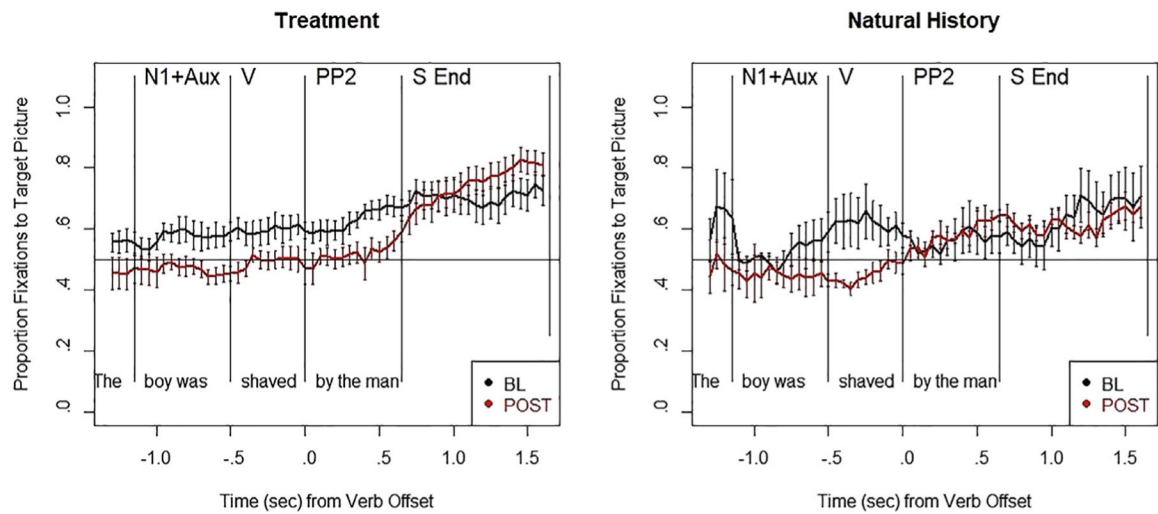
**Fig. 3 –.** Individual acquisition curves derived from the weekly probe tasks for a) production and b) comprehension of the trained items. Lines display the percentage of correct responses at BL (average of multiple baseline assessments) and at 4 time points over the course of treatment (i.e., every three weeks), for each participant.



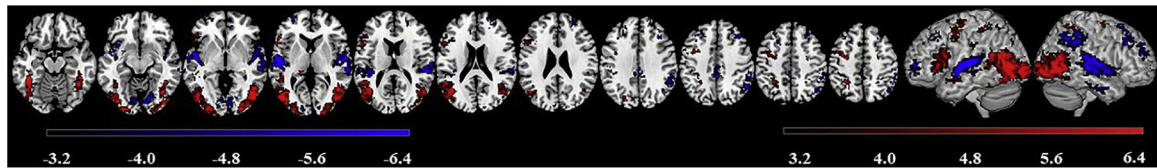
**Fig. 4** – Mean percent accurate responses by sentence type, at baseline (BL) and post-testing (POST), for the two participant groups on the sentence production priming task (production). FP.u = full passives with adjunct (untrained items); UP = untrained passives, UAC = active unaccusatives, ACT = active transitives; OC = object clefts. Bars indicate mean standard error.



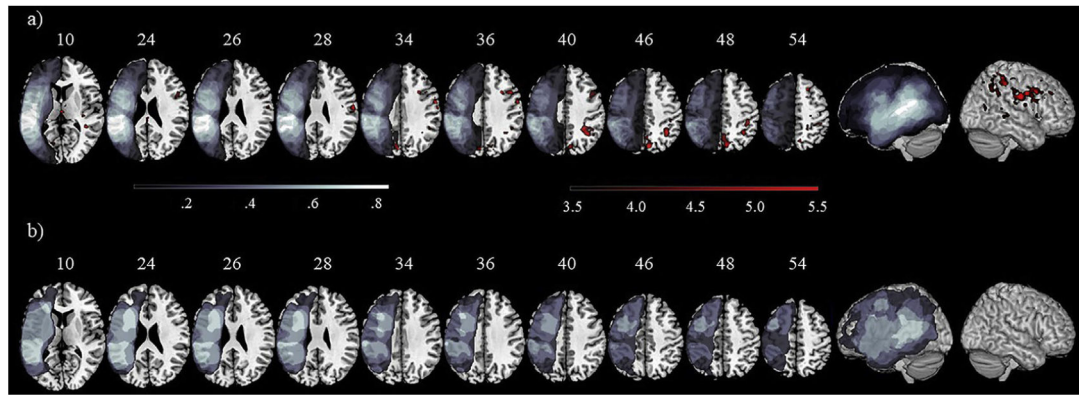
**Fig. 5** -. Mean percent correct responses by sentence type, at baseline (BL) and post-testing (POST), for the two participant groups on the full probe picture verification task (comprehension). FP.u = full passives with adjunct (untrained items); UP = untrained passives, UAC = active unaccusatives, ACT = active transitives; OC = object clefts. Bars indicate mean standard error.



**Fig. 6 –** Proportion of fixations to the target picture during online comprehension of correctly-answered passive sentences at baseline (BL) (black line) and post-testing (POST), for the two aphasic groups. Sentence regions: N1+Aux = subject noun + auxiliary; V = verb; N/PP2 = post-verbal noun or prepositional phrase; S End = sentence end.

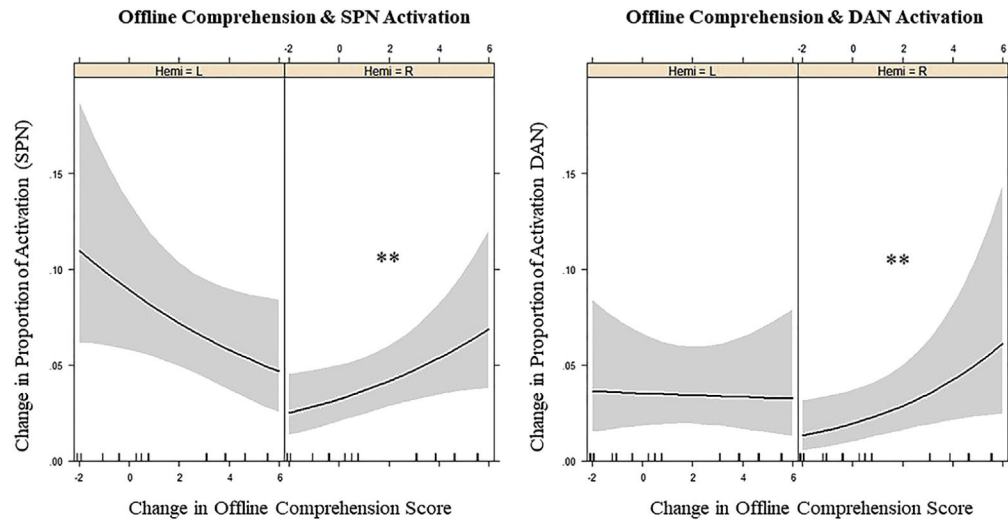


**Fig. 7 –.**  
Clusters of suprathreshold activation ( $p < .001$ , uncorrected,  $k > 46$ ) for the *Passive > Control* (red) and the *Control > Passive* (blue) contrasts, in the group of healthy participants. Color bars indicate activation intensity, with lighter shades indicating larger T-values (i.e., greater activation).



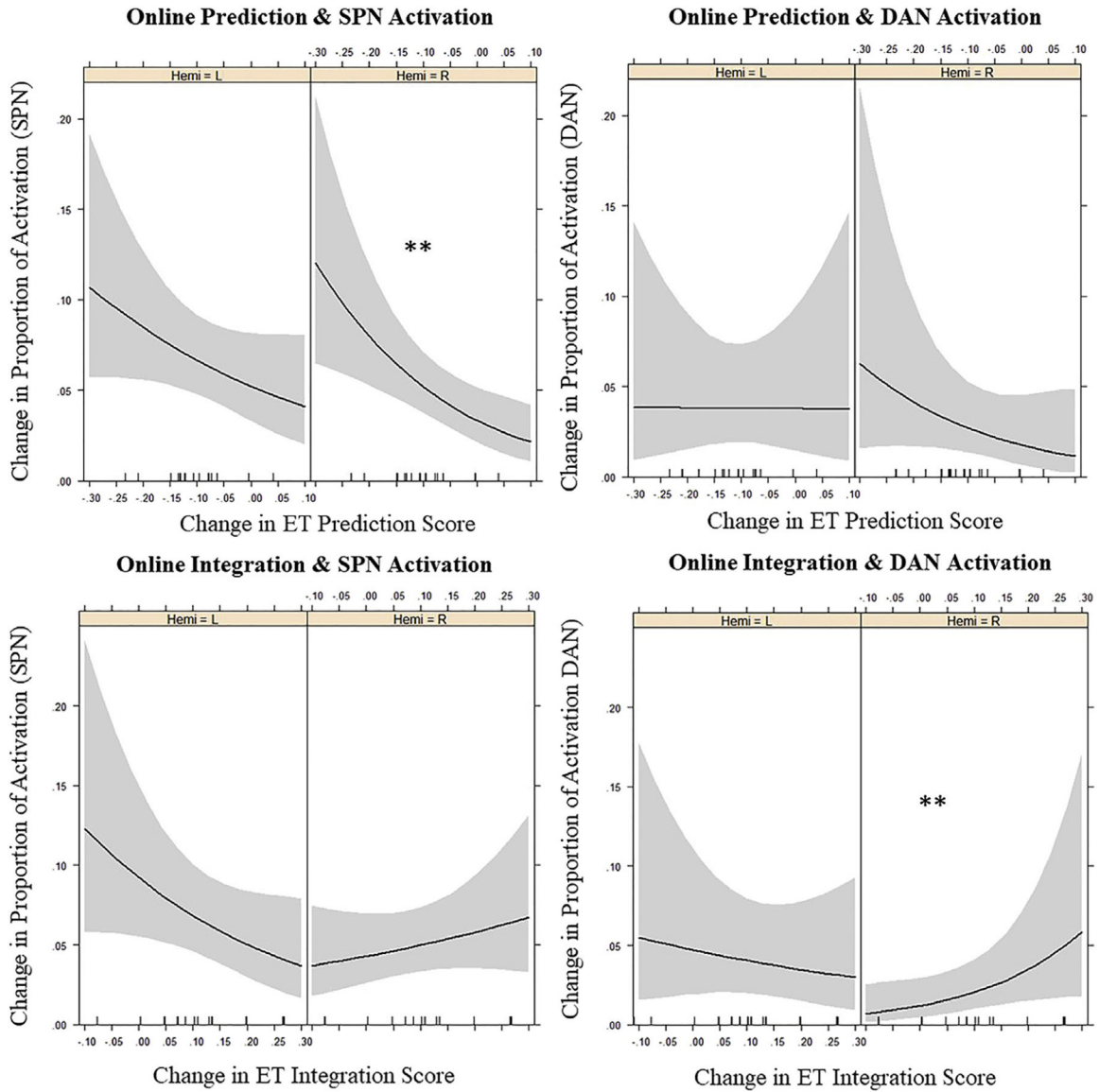
**Fig. 8 –.**

Areas of left hemisphere lesion overlap across participants in the treatment (a) and natural history (b) groups are shown in shades of gray, with lighter shades indicating maximal overlap. Clusters of significant suprathreshold activation ( $p < .001$ , uncorrected,  $k > 37$ ) for the (*Passive* > *Control*, *POST* > *BL*) contrast are shown in shades of red, with lighter shades corresponding to higher T-values, i.e., greater activation intensity. No significant activation was observed in the natural history group.



**Fig. 9** -. Relation between increased (POST > BL) activation in the SPN and DAN networks and changes in the offline comprehension z-score. Asterisks indicate statistically significant comparisons.





**Fig. 10 –.**  
 Relation between increased (POST > BL) activation in the SPN and DAN networks and changes in online thematic prediction (a, b) and online thematic integration (c, d). 9a shows a significant relation between upregulation in the right SPN and a decrease in the proportion of fixations to the target, reflecting a shift toward an Agent-first strategy during processing of passive sentences. 9d shows a significant relation between upregulation in the right DAN and an increase in target fixations from the verb region to the end of the sentence.





Table 2 –

Percent correct responses at baseline (BL, i.e., average of multiple baseline measurements) and at post-testing (POST) for production and comprehension of trained full passives with locative adjunct. Percent increase (i.e., difference in percent correct responses between BL and Post) and effect sizes (ESs), computed as Cohen's H index, are also provided by participant and modality.

Participant	Production			Comprehension				
	BL accuracy	POST accuracy	FP trained % increase	Cohen's H	BL accuracy	POST accuracy	FP trained % increase	Cohen's H
P1	.0	90.0	90.0	2.498	70.0	70.0	–	–
P2	10.0	70.0	60.0	1.339	50.0	80.0	30.0	.644
P3	53.3	100.0	46.7	1.505	40.0	100.0	60.0	1.772
P4	2.5	70.0	67.5	1.665	30.0	40.0	10.0	.210
P5	.0	80.0	80.0	2.214	33.3	70.0	36.7	.758
P6	22.5	100.0	77.5	2.153	50.0	70.0	20.0	.412
P7	.0	65.0	65.0	1.875	30.0	70.0	40.0	.823
P8	10.0	100.0	90.0	2.498	56.7	80.0	23.3	.509
P9	7.5	90.0	82.5	1.943	55.0	70.0	15.0	.311
P10	.0	90.0	90.0	2.498	40.0	40.0	–	–
P11	.0	30.0	30.0	1.159	33.3	50.0	16.7	.347
P12	.0	0.0	–	–	43.3	40.0	–	–
P13	.0	80.0	80.0	2.214	40.0	95.0	55.0	1.321
P14	30.0	90.0	60.0	1.339	53.3	60.0	6.7	.135
group average	9.7	75.4	70.7	1.9	44.6	66.8	28.5	.7

**Table 3 –**

Eyetracking results showing Thematic Prediction (TP) and Thematic Integration (TI) scores for each participant at baseline (BL) and post-testing (Post) phases of the study. The TP score is the proportion of target fixations within the region encompassing the sentence subject (NI) and the verb (V) region. The TI score is derived by subtracting the proportion of target fixations in the V region from the proportion of target fixations in the sentence end region. Both scores were computed on correctly-answered passive sentences only. NT = not tested.

Score	Participant	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Group Average
TP	BL	.52	.63	.57	.45	.73	.50	NT	.49	.62	.59	.42	.73	.57	NT	.57
	FU	.51	.52	.47	.61	.41	.37	NT	.41	.47	.52	.46	.55	.33	NT	.47
	Change	.00	-.11	-.10	.16	-.32	-.13		-.08	-.15	-.07	.04	-.18	-.23		-.10
TI	BL	.22	.11	-.16	-.03	.04	.32	NT	-.04	-.01	-.02	.10	-.17	.53	NT	.07
	FU	.30	.24	-.02	.17	.08	.24	NT	.31	.26	.24	.20	.15	.60	NT	.23
	Change	.09	.13	.14	.20	.05	-.08		.35	.27	.27	.10	.32	.07		.16
Score	Participant	C1	C2	C3	C4	C5										Group Average
TP	BL	.45	.69	.55	NT	.60										.57
	FU	.39	.48	.42	NT	.47										.44
	Change	-.06	-.21	-.12		-.14										-.13
TI	BL	.07	.05	.22	NT	-.02										.08
	FU	-.04	.16	.10	NT	-.02										.05
	Change	-.11	.11	-.12		.00										-.03

Clusters of suprathreshold ( $p < .001$  voxel-level uncorrected,  $k > 46$ ) activation for *Passive* > *Control* and *Control* > *Passive* for the group of healthy participants. Region labels are derived from the Harvard-Oxford (Desikan et al., 2006). For each cluster the first label (italicized) provides the region with the peak activation, for which MNI coordinates and corresponding t-values are also given.

Table 4 –

Contrast	cluster size (k)	Peak coordinates (xyz)	t-value	Region	L/R	Contrast	cluster size (k)	Peak coordinates (xyz)	t-value	Region	L/R
Passive > Control	3206	-44 -50 -18	13.84	<i>ITG temporo-occipital</i>	L	Control > Passive	1593	58 16 8	-9.07	<i>Planum Temporale</i>	R
				ITG posterior	L					SMG posterior	R
				STG posterior	L					STG anterior	R
				SMG posterior	L					Planum Polare	R
				LOC inferior	L		1397	-50 -12 8	-11.6	<i>Central Operculum</i>	L
				FG occipital	L					STG posterior	L
				Occipital Pole	L					PoCG	L
	2450	48 -72 -2	9.63	<i>LOC inferior</i>	R					Heschl's Gyrus	L
				MTG temporo-occipital	R					Planum Temporale	L
				ITG temporo-occipital	R		555	4 -82 -4	-6.68	<i>Lingual Gyrus</i>	R
				AG	R					Intracalcarine Cortex	R
				FG temporo-occipital	R					FG occipital	R
				FG occipital	R		457	52 -44 50	-6.95	<i>SMG posterior</i>	R
				Occipital Pole	R					SMG anterior	R
	324	-58 -16 4	6.6	<i>IFG opercularis</i>	L					LOC superior	R
				IFG triangularis	L		185	2 -40 40	-6.59	<i>CG posterior</i>	R
				MFG	L					PCG	R
	240	-28 -8 50	6.09	<i>PCG</i>	L		176	34 28 36	-5.76	<i>MFG</i>	R
				MFG	L					SFG	R
	153	-6 8 62	8.53	<i>n/a</i>	L					Frontal Pole	R
				Paracingulate	L		87	28 52 24	-5.02	<i>Frontal Pole</i>	R
				Gyrus							
	124	-30 -46 50	6.07	<i>SPL</i>	L		66	-30 58 4	-6.08	<i>Frontal Pole</i>	L
	93	26 -6 60	6.08	<i>n/a</i>	R		54	56 -8 -28	-5.68	<i>n/a</i>	R

Contrast	cluster size (k)	Peak coordinates (xyz)	t-value	Region	L/R	Contrast	cluster size (k)	Peak coordinates (xyz)	t-value	Region	L/R
				SFG	R					MTG posterior	R
	74	-26 -72 30	7.8	PCG	R					ITG posterior	R
	71	-48 0 50	4.91	<i>LOC superior</i>	L						
				<i>PCG</i>	L						
				MFG	L						
	69	28 -56 60	5.58	<i>SPL</i>	R						

AG = angular gyrus; FG = fusiform gyrus; IFG = inferior frontal gyrus; ITG = inferior temporal gyrus; LOC = lateral occipital cortex; MFG = middle frontal gyrus; MTG = middle temporal gyrus; PCG = precentral gyrus; SFG = superior frontal gyrus; SMG = supramarginal gyrus; SPL = superior parietal lobule; STG = superior temporal gyrus.

Summary of lesioned regions-of-interest (ROIs) in anterior and posterior regions of the brain across participants. ROIs are derived from the Harvard-Oxford (Desikan et al., 2006) atlas. For the treatment (Tx) and natural history (NH) groups separately, the number of participants showing at least 10% lesioned tissue in each ROI are reported together with the mean percent/range of lesioned tissue by region.

**Table 5 –**

ROI	Tx Group (N = 13)	NH Group (N = 5)	Total (N = 18)	Mean % lesioned tissue (range)
<i>Anterior Regions</i>				
Frontal Operculum	11	3	14	69 (4–100)
IFG opercularis	11	3	14	61.5 (1.2–99.8)
IFG triangularis	9	3	12	36.7 (1.1–100)
Orbitofrontal Cortex	8	3	11	33.5 (2–91)
Insula	12	4	16	70.8 (11.9–100)
MFG	8	4	12	33 (8–92.2)
SFG	2	2	4	13.4 (2–97.5)
Frontal Pole	3	2	5	11 (1–88.4)
<i>Posterior Regions</i>				
STG posterior	12	4	16	62.6 (33.2–99.3)
STG anterior	9	3	12	60.2 (6.7–100)
AG	12	4	16	58.5 (7.1–99)
SMG posterior	13	4	17	56 (18–94.4)
SMG anterior	11	4	15	52 (3.7–90.3)
MTG temporo-occipital	10	4	14	48.5 (1–98.7)
MTG anterior	9	2	11	48 (1.3–100)
MTG posterior	10	4	14	43.3 (3.1–98.2)
Temporal Pole	9	3	12	35 (5–84.3)
SPL	5	2	7	22.1 (6–91.2)
ITG anterior	6	1	7	21.1 (7.8–84.1)
LOC inferior	8	2	10	20.4 (3–71.7)
LOC superior	6	2	8	18.5 (2–86.3)
ITG posterior	4	1	5	13.4 (3.4–90.6)
ITG temporo-occipital	4	1	5	10.3 (7–80.2)

AG = angular gyrus; IFG = inferior frontal gyrus; ITG = inferior temporal gyrus; LOC = lateral occipital cortex; MFG = middle frontal gyrus; MTG = middle temporal gyrus; SFG = superior frontal gyrus; SMG = supramarginal gyrus; SPL = superior parietal lobule; STG = superior temporal gyrus.



**Table 6 –**

Summary of lesioned regions-of-interest (ROIs) within the left sentence-processing and dorsal attention networks across participants. ROIs are derived from the Harvard-Oxford (Desikan et al., 2006) atlas. The number of participants showing at least 10% lesioned tissue in each ROI are reported for each participant group. The mean percent/range of lesioned tissue across all participants is also displayed.

ROI	Tx Group (N = 13)	NH Group (N = 5)	Total (N = 18)	Mean % lesioned tissue (range)
Sentence Processing Network				
IFG triangularis	9	3	12	36.7 (1.1–100)
MTG temporo-occipital	10	4	14	48.5 (1–98.7)
STG posterior	12	4	16	62.6 (33.2–99.3)
AG	12	4	16	58.5 (7.1–99)
Dorsal Attention Network				
MFG	8	4	12	33 (.8–92.2)
PCG	9	3	12	24.6 (1.8–90.2)
SPL	5	2	7	22.1 (.6–91.2)
LOC superior	6	2	8	18.5 (.2–86.3)

AG = angular gyrus; IFG = inferior frontal gyrus; LOC = lateral occipital cortex; MFG = middle frontal gyrus; MTG = middle temporal gyrus; PCG = precentral gyrus; STG = superior temporal gyrus.

**Table 7 –**

Clusters of suprathreshold ( $p < .001$  voxel-level uncorrected,  $k > 37$ ) activation for the (*Passive > Control, POST > BL*) contrast in the treatment group. Region labels are derived from the Harvard–Oxford (Desikan et al., 2006). For each cluster the first label (italicized) provides the region with the peak activation, for which MNI coordinates and corresponding t-values are also given.

Contrast	cluster size (k)	Peak coordinates (xyz)	t-value	Region	L/R
Passive > Control, POST > BL	90	22 -48 58	7.130	<i>SPL</i>	R
	42	60 -10 30	6.885	PoCG	R
	58	44 -40 58	6.880	SMG anterior	R
				<i>SPL</i>	R
	141	8 -74 46	6.835	PoCG	R
				SMG anterior	R
	182	36 -52 46	6.338	SMG posterior	R
				<i>Precuneous</i>	R
	43	-4 -78 34	5.919	Cuneous	R
				LOC superior	R
	48	48 12 38	5.527	<i>SPL</i>	R
				SMG posterior	R
	48	48 12 38	5.527	AG	R
				LOC superior	R
	48	48 12 38	5.527	<i>Cuneous</i>	L
				Precuneous	L
	48	48 12 38	5.527	<i>MFG</i>	R
				PCG	R

AG = angular gyrus; LOC = lateral occipital cortex; MFG = middle frontal gyrus; PCG = precentral gyrus; PoCG = postcentral gyrus; SMG = supramarginal gyrus; SPL = superior parietal lobule.