Brain Damage Associated with Impaired Sentence Processing in Acute Aphasia

Sigfus Kristinsson[1], Helga Thors[1], Grigori Yourganov[1], Sigridur Magnusdottir[2], Haukur Hjaltason[2], Brielle C. Stark[3], Alexandra Basilakos[1], Dirk-Bart den Ouden[1], Leo Bonilha[4], Chris Rorden[1], Gregory Hickok[5], Argye Hillis[6], and Julius Fridriksson[1]

Abstract

■ Left-hemisphere brain damage commonly affects patients' abilities to produce and comprehend syntactic structures, a condition typically referred to as "agrammatism." The neural correlates of agrammatism remain disputed in the literature, and distributed areas have been implicated as important predictors of performance, for example, Broca's area, anterior temporal areas, and temporo-parietal areas. We examined the association between damage to specific language-related ROIs and impaired syntactic processing in acute aphasia. We hypothesized that damage to the posterior middle temporal gyrus, and not Broca's area, would predict syntactic processing abilities. One hundred four individuals with acute aphasia (<20 days poststroke) were included in the study. Structural MRI scans were obtained, and all participants completed a 45-item sentence–picture matching task. We performed an ROI-based stepwise regression analyses to examine the relation between cortical brain damage and impaired comprehension of canonical and noncanonical sentences. Damage to the posterior middle temporal gyrus was the strongest predictor for overall task performance and performance on noncanonical sentences. Damage to the angular gyrus was the strongest predictor for performance on canonical sentences, and damage to the posterior superior temporal gyrus predicted noncanonical scores when performance on canonical sentences was included as a cofactor. Overall, our models showed that damage to temporo-parietal and posterior temporal areas was associated with impaired syntactic comprehension. Our results indicate that the temporo-parietal area is crucially implicated in complex syntactic processing, whereas the role of Broca's area may be complementary. ■

INTRODUCTION

Left-hemisphere brain damage causes language deficits in 20–40% of stroke survivors (Engelter et al., 2006). The nature of the deficit depends on the extent and localization of damage (Caplan, Michaud, Hufford, & Makris, 2016; Ochfeld et al., 2010; Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Ferro, Mariano, & Madureira, 1999). Individuals who present with agrammatism after stroke have particular problems with the production and comprehension of functional categories and complex syntactic structures (Thompson, Bonakdarpour, & Fix, 2010; Goodglass, 1997). Despite being widely studied, the brain areas crucial for successful sentence processing are still disputed in the literature (Rogalsky et al., 2018; Hagoort & Indefrey, 2014). Functional brain imaging and lesion–symptom mapping studies have implicated distributed cortical and subcortical areas as important for sentence processing, including inferior frontal areas

(Grodzinsky, 2000; Caramazza & Zurif, 1976), anterior (Magnusdottir et al., 2013) and posterior (Den Ouden et al., 2019; Rogalsky et al., 2018) temporal areas, and temporo-parietal areas (Dronkers et al., 2004). Although functional divisions between these areas have been proposed by previous studies, these remain controversial (Indefrey, 2012). The current study aimed to investigate the lesion patterns predictive of receptive agrammatism in individuals in the acute phase of stroke by implementing a hypothesis-driven approach in a larger sample of patients with aphasia than previous studies.

Broca's area—pars opercularis and pars triangularis (roughly corresponding to BA 44 and 45)—is classically considered important for speech production (Broca, 1861). Furthermore, since Caramazza and Zurif's (1976) seminal paper, Broca's area has also been considered important for syntactic processing. Caramazza and Zurif (1976) compared processing of different sentence types in 15 patients with aphasia and found that patients with Broca's aphasia performed at chance level on sentences with noncanonical word order, that is, other than subject–verb–object, despite retaining relatively preserved lexical–semantic comprehension. They concluded that patients with Broca's aphasia are impaired in their

[1]University of South Carolina, [2]University of Iceland, [3]Indiana University, [4]Medical University of South Carolina, [5]University of California, Irvine, [6]Johns Hopkins University School of Medicine

capacity to use syntactic-like algorithmic processes necessary for decoding complex sentences. Consequently, later studies assumed a common neuroanatomical origin of syntax for language comprehension and production (e.g., Grodzinsky, 2000; Swinney & Zurif, 1995; but see Caramazza, Capitani, Rey, & Berndt, 2001; Caramazza & Hillis, 1989, for an alternative view).

The involvement of Broca's area in syntactic processing has been supported by numerous functional neuroimaging studies showing activation in the inferior frontal gyrus (IFG) for processing complex sentences in healthy adults (e.g., Santi & Grodzinsky, 2007; Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003; Caplan, Alpert, & Waters, 1998). On the basis of this line of evidence, Broca's area has been suggested to support core computations necessary for processing complex sentences, for example, syntactic movement (Grodzinsky & Santi, 2008; Ben-Shachar et al., 2003), propositional meaning (i.e., who is doing what to whom; Caplan, Alpert, & Waters, 1999; Caplan et al., 1998), and hierarchical structure building (Friederici, 2009). Others have claimed that Broca's area is involved in syntactic processing through working memory (Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; Kaan & Swaab, 2002), cognitive control (Novick, Trueswell, & Thompson-Schill, 2005), or articulatory rehearsal (Rogalsky, Almeida, Sprouse, & Hickok, 2015; Rogalsky, Matchin, & Hickok, 2008).

Aside from the specific role of Broca's area in sentence comprehension, the extent to which Broca's area is crucial for processing complex syntax is debated. First, the assumption made by Caramazza and Zurif (1976) that the behavioral profile of patients with Broca's aphasia (i.e., telegraphic speech) coherently aligns with a syntactic processing impairment in the same patients and that this be taken as evidence for a shared syntactic processing network located in Broca's area appears to be incorrect. Behaviorally diagnosed Broca's aphasia can be caused by lesions sparing Broca's area (Fridriksson, Bonilha, & Rorden, 2007), and damage to Broca's area alone does not necessarily cause Broca's aphasia (Mohr et al., 1978). Second, and more importantly, damage to Broca's area does not invariably result in syntactic comprehension deficits. Studying a group of patients with confirmed Broca's area damage, Caramazza, Capasso, Capitani, and Miceli (2005) found that most of their 38 participants performed better than predicted if Broca's area were critical for processing of noncanonical sentences. Similar results have since been reported elsewhere (Rogalsky et al., 2018; Thothathiri, Kimberg, & Schwartz, 2012). Third, the notion of a central syntactic deficit was challenged by studies showing unexpected sensitivity to syntactic structure in speakers with agrammatic aphasia (e.g., Linebarger, Schwartz, & Saffran, 1983). These findings call into question the specificity of Broca's area in syntactic processing.

Studies mapping the association between localized brain lesions and syntactic processing have revealed somewhat contrasting results. Earlier studies failed to localize syntactic processing, inferring that different brain regions were required for syntactic processing in different individuals (Caplan, Waters, & DeDe, 2007; Caplan, Hildebrandt, & Makris, 1996) or that syntax relied on a distributed left perisylvian network (Wilson & Saygin, 2004).

More recently, voxel-based lesion–symptom mapping (VLSM) studies in patients with chronic stroke-induced aphasia and primary progressive aphasia indicate that spared syntactic processing is modulated by a preserved functional connection between the left IFG and posterior temporal regions (Wilson et al., 2014; Tyler et al., 2011; Tyler, Wright, Randall, Marslen-Wilson, & Stamatakis, 2010). Specifically, studies have consistently found that lesions to posterior temporal and inferior parietal regions are predictive of poor syntactic processing (Den Ouden et al., 2019; Fridriksson et al., 2018; Rogalsky et al., 2018; Pillay, Binder, Humphries, Gross, & Book, 2017; Magnusdottir et al., 2013; Thothathiri et al., 2012; Dronkers et al., 2004). Dronkers et al. (2004) examined the association between lesioned brain areas and sentence comprehension in 64 patients with chronic left-hemisphere stroke. Participants were evaluated on the Curtiss–Yamada Comprehensive Language Evaluation-Receptive and required to select a line drawing from an array of three to four images. The results suggest that lesions to the posterior middle temporal gyrus (pMTG), anterior superior temporal gyrus (STG), STS, angular gyrus (AG), and two frontal areas including BAs 46 and 47 affect comprehension, with lesions involving the middle temporal gyrus (MTG) resulting in the worst overall sentence comprehension scores. In a subsequent study, Thothathiri et al. (2012) explored the neural basis of reversible sentence comprehension and found that a large cluster in the temporo-parietal region, spanning MTG, STG, AG, and supramarginal gyrus (SMG; BAs 21, 22, 39, and 40), was reliably associated with overall sentence comprehension. Looking at the effect of canonicity, the authors reported significant results for canonical and noncanonical sentences in the AG and SMG, whereas STG also contributed significantly only to the performance on noncanonical sentences. Consistent with these results, our group compared performance on canonical and noncanonical sentences in 50 Icelandic patients with acute stroke and concluded that poor performance on canonical sentences implicated lesions to the TPJ, whereas impaired performance on noncanonical sentences was predicted by lesion to anterior STG and MTG and the temporal pole (Magnusdottir et al., 2013).

Drawing from these previous studies, Pillay et al. (2017) tested the hypothesis that temporal lobe regions inferior to Wernicke's area contain a critical site for combinatorial processing of phrase-level and sentence-level language in 51 patients with chronic stroke. The study controlled for diverse language and domain-general

processes not specific to spoken language comprehension by incorporating a picture naming task as a covariate in the main analysis. Using this methodology, they found that sentence comprehension was correlated with damage in the middle MTG and pMTG, extending into the STS. Pillay et al. concluded that "this region appears to be critically necessary for integration of multiword combinations during spoken language comprehension."

Finally, Fridriksson et al. (2018), as well as Den Ouden et al. (2019), found lesions to posterior temporal areas to be most highly predictive of performance on sentence comprehension tasks, including tasks contrasting performance on noncanonical and canonical sentences, using both a univariate and a multivariate region-wise lesion–symptom mapping. Fridriksson et al. (2018) also implicated anterior temporal areas in predicting performance on noncanonical relative to canonical sentences, whereas Den Ouden et al. (2019) noted a dissociation between morphosyntactic production, supported by inferior frontal cortex, and syntactic comprehension, supported by posterior temporal areas.

A few studies have specifically explored hypotheses about the role of Broca's area in sentence comprehension. These studies have largely failed to support a specific role for Broca's area in receptive syntactic processing in favor of the TPJ (Newhart et al., 2012; Race, Ochfeld, Leigh, & Hillis, 2012), while identifying an association between damage to Broca's area with STM and working memory components (Newhart et al., 2012). More recently, Rogalsky et al. (2018) investigated the conflicting results in the literature regarding the roles of Broca's area and the anterior temporal lobe (ATL) in sentence comprehension, finding that neither patients with Broca's area damage nor ATL damage exhibited the expected agrammatic comprehension pattern (i.e., relatively better performance on canonical vs. noncanonical sentences). Maximal lesion overlap in patients who presented with agrammatic comprehension patterns was located in the left pSTG and pMTG. In addition, Rogalsky et al. also reported that ATL lesion contributed to overall lower performance on canonical and noncanonical sentences, similar to the findings of Magnusdottir et al. (2013).

Notably, none of the cited VLSM studies has found a predictive relationship between localized damage to Broca's area and complex syntactic processing, contradicting the classically suggested specific role of the area in sentence comprehension. Furthermore, recent large-scale meta-analyses of neuroimaging studies have identified consistent and robust activation in the left STG and MTG alongside activation in Broca's area in tasks requiring syntactic processing (Hagoort & Indefrey, 2014; Indefrey, 2012). Caplan et al. (2016) came to the same conclusion after performing an extensive literature review. A network analysis by Den Ouden et al. (2012) also shows that inferior frontal and posterior superior to middle temporal regions interact strongly during the processing of noncanonical sentences in unimpaired speakers.

Therefore, the dispute over which of these areas are crucial for successful syntactic processing, and to what extent, still remains.

In this study, we utilized a framework for the cortical organization of speech processing proposed in the dual-stream model of speech/language processing (DSM; Hickok & Poeppel, 2007). The model involves a ventral stream responsible for processing speech signals for comprehension and a dorsal stream that maps acoustic speech signals to frontal lobe articulatory networks. Building on cumulative empirical evidence obtained throughout the history of speech and language research, this model outlines an extensive neural organization of speech perception and is widely established as a leading model of speech and language processing (currently cited by 3,109 articles according to Google Scholar database). A recent large-scale, data-driven study of lesion data and behavioral testing from stroke survivors by Fridriksson et al. (2016) furthermore supported Hickok and Poeppel's model by revealing the suggested dual streams of speech processing, characterizing the dorsal stream as a form-to-articulation pathway and the ventral stream as a form-to-meaning pathway. Most recently, Matchin and Hickok (2019) built on these findings and extended the DSM to include a detailed neuroanatomical framework of syntax, wherein posterior temporal regions are postulated to perform hierarchical structuring functions and inferior frontal regions are sometimes co-opted for sentence comprehension as a form of syntactic working memory and syntactic prediction.

More specifically, we examined the association between proportional damage to predefined language-related ROIs within the DSM based on our prior work (Fridriksson et al., 2016) and syntactic processing abilities. We hypothesized that damage to the pMTG would predict overall sentence processing scores over and above other left-hemisphere language areas, and more specifically, we hypothesized that damage to the pMTG would be relatively more strongly associated with performance on noncanonical versus canonical syntactic structures. Furthermore, we hypothesized that proportional damage to Broca's area would not contribute significantly to performance on noncanonical relative to canonical sentences.

METHODS

Participants

One hundred four stroke patients with acute left-hemisphere injury participated in the study to determine brain damage associated with comprehension of spoken sentences matched with pictures.[1] Fifty-four patients admitted to the neurology ward at Landspitali University Hospital in Reykjavik were recruited for the Icelandic sample (previously reported on in Magnusdottir et al. [2013]), and 50 patients admitted to the Palmetto Health Richland

Hospital in Columbia, South Carolina, were recruited for the U.S. sample. All participants had incurred a single, unilateral stroke to the left hemisphere and gave informed consent for study participation. For both samples, participants were included if they (1) were in the acute phase of stroke, (2) had their stroke confirmed by a CT/MRI scan, (3) had no history of major psychiatric illness or other neurologic impairment affecting the brain, and (4) were native speakers of Icelandic in the Icelandic sample and native speakers of English in the U.S. sample. All study procedures were approved by the institutional review boards at each test site.

There were 25 female and 29 male participants with a mean age of 67.1 years (SD = 11.1 years) in the Icelandic sample. In the U.S. sample, there were 22 female and 28 male participants with a mean age of 63.1 years (SD = 8.8 years). In the combined sample, there were 47 female and 57 male participants with a mean age of 65.2 years (SD = 10.0 years). Mean lesion volume in the Icelandic sample was 4514 cc, compared to 2848 cc in the U.S. sample. Although lesion volume was not significantly different between groups, $t(102)$ = 1.66, p = .10, we do include lesion volume as a regressor in all analyses to account for the effects of lesion size on the outcome measures. Participant characteristics are presented in Table 1.

In most cases, MRI and behavioral examinations were conducted within 3 days of hospital admission, although a handful of patients underwent examination up to 20 days after stroke because of complications secondary to the stroke. MRI and behavioral testing were always completed within a day of each other.

Language impairment was assessed with the Bedside Evaluation Screening Test–Second Edition in the Icelandic sample (BEST-2; Fitch-West, Sands, & Ross-Swain, 1998) and the Western Aphasia Battery–Bedside (WAB-B; Kertesz, 2007) in the U.S. sample. Both test batteries provide an overall assessment of aphasia severity and include subtests that measure spontaneous speech abilities, auditory comprehension, speech repetition, and object naming. Participants' mean score was 108 of 135 possible points (80% correct) on the BEST-2 and 81.5 of 100 possible points (82% correct) on the WAB-B (see Table 1). The proportional score on BEST-2 and WAB was not statistically different between groups, $t(102)$ = −.36, p = .72.

Agrammatism Testing

All participants completed a sentence–picture matching test in which the experimenter produces target sentences of varying degrees of syntactic complexity and the participant is asked to point to a corresponding picture from a field of three horizontally aligned line drawings (Magnusdottir, 2005). Each set of line drawings includes a depiction of the target sentence, a version in which the roles in the target sentence are reversed, and a lexical foil. This syntactic processing test was developed simultaneously in Icelandic and English with the intent of

Table 1. Participant Characteristics and Aphasia Scores[a]

	Icelandic (n = 54)	U.S. Sample (n = 50)
Mean age in years (SD)	67.1 (11.1)	63.1 (8.8)
Sex (F/M)	25/29	22/28
Race (AA/Cauc)	0/54	21/29
Lesion volume	4.514 cc	2.848 cc
WAB-B Aphasia Quotient	−	81.5/100
Content	−	7.8/10
Fluency	−	8.0/10
Y/N questions	−	9.1/10
Sequential commands	−	8.1/10
Repetition	−	8.1/10
Object naming	−	7.8/10
BEST-2 Aphasia Quotient	108/135	−
Conversational expression	24/30	−
Naming objects	24/30	−
Describing objects	24/30	−
Repeating sentences	24/30	−
Pointing to objects	27/30	−
Point to parts of a picture	27/30	−
Reading	21/30	−

AA = African American; Cauc = Caucasian; F = female; M = male.

[a]Language impairment was evaluated with the BEST-2 (Fitch-West et al., 1998) in the Icelandic sample and the WAB-B (Kertesz, 2007) in the U.S. sample.

creating comparable sentences in terms of syntactic structure and complexity in both languages (Magnusdottir, 2000).

The test consists of 45 sentences grouped into nine sentence types, with each type containing five sentences (see Table 2). As described in Magnusdottir et al. (2013), it provides (a) the Icelandic version of the sample sentence and (b) the proposed syntactic structure to indicate the complexity involved. Here, "t_i" and "t_j" represent the position of the relevant "gaps" (i.e., the site of moved or deleted constituents) to be associated with the preceding coindexed element (see Chomsky, 1981, 1995). When a finite verb has supposedly moved from its basic position inside a verb phrase, this movement is indicated by a "t_v" in the basic position(s). The Icelandic sentences used are followed by an English word by word gloss and an idiomatic translation. Case marking of the Icelandic noun phrases and agreement features are also indicated in the translation, using standard linguistic

Table 2. Examples for Each of the Nine Sentence Types Presented in the Task in Icelandic and English

Sentence Type	Icelandic	English
1. Active declarative sentences	*Stelpan málar strákinn*	*The girl paints the boy*
	Stelpan$_j$ málar$_v$ [$_{VP}$ t$_i$ t$_v$ strákinn]	The girl(N) paints the boy(A)
2. Passive sentences	*Strákurinn er málaður af stelpunni*	*The boy is painted by the girl*
	Strákurinn$_i$ er [$_{VP}$ málaður t$_i$] af stelpunni	The boy(N) is painted(m,sg) by the girl(D)
3. Truncated passive sentences	*Strákurinn er málaður*	*The boy is painted*
	Strákurinn$_i$ er [$_{VP}$ málaður t$_i$]	The boy(N) is painted(m,sg)
4. Cleft sentences with subject gaps	*Það er strákurinn sem málar stelpuna*	*It is the boy that paints the girl*
	Það er strákurinn$_i$ [$_{CP}$ sem [$_{IP}$ t$_j$ málar$_v$ [$_{VP}$ t$_i$ t$_v$ stelpuna]]]	It is the boy(N) that paints the girl(A)
5. Cleft sentences with object gaps	*Það er stelpan sem strákurinn málar*	*It is the girl that the boy paints*
	Það er stelpan$_j$ [$_{CP}$ sem [$_{IP}$ strákurinn$_i$ málar$_v$ [$_{VP}$ t$_i$ t$_v$ t$_j$]]]	It is the girl(N) that the boy(N) paints
6. Sentences with topicalized object and a main verb	*Stelpuna málar strákurinn*	*The girl, the boy paints*
	Stelpuna$_j$ málar$_v$ [$_{IP}$ strákurinn$_i$ [$_{VP}$ t$_i$ t$_v$ t$_j$]]	The girl(A) paints the boy(N)
7. Sentences with topicalized object and an auxiliary verb	*Stelpuna er strákurinn að mála*	*The girl, the boy is painting*
	Stelpuna$_j$ er [$_{IP}$ strákurinn$_i$ [$_{VP}$ t$_i$ að mála t$_j$]]	The girl(A) is the boy(N) to paint
8. Referential *wh*-questions with subject gap and main verbs	*Hvaða strákur málar stelpuna?*	*Which boy paints the girl?*
	Hvaða strákur$_i$ málar$_v$ [$_{VP}$ t$_i$ t$_v$ stelpuna]	Which boy(N) paints the girl(A)?
9. Referential *wh*-questions with object gap and main verbs	*Hvaða strák málar stelpan?*	*Which boy is the girl painting?*
	Hvaða strák$_j$ málar$_v$ [$_{IP}$ stelpan$_i$ t$_v$ [$_{VP}$ t$_i$ t$_v$ t$_j$]]?	Which boy(A) paints the girl(N)?

In all sentences, knowing who did what to whom involves associating the gaps with the appropriate antecedent.

abbreviations (e.g., N for *nominative*, A for *accusative*, m for *masculine*, sg for *singular*). Knowing who did what to whom involves associating the gaps with the appropriate antecedents in all sentence types.

Separate analyses were performed on sentences with canonical word order, that is, subject–verb–object (CAN: Sentence Types 1, 4, and 8), and noncanonical word order (non-CAN: 2, 5, and 9). In this comparison, truncated sentences of Type 3 were excluded because these do not include an overt object (only a theme, but no agent), as well as the topicalized Sentence Types 6 and 7, as their noncanonical structure cannot be directly compared with a canonical counterpart. Total scores and scores on CAN and non-CAN sentence types were compared between the Icelandic and U.S. samples before the samples were combined to examine whether performance was comparable. An independent samples t test found no significant difference between the samples on total scores (Icelandic vs. US: 36.39 vs. 35.08, $t(102) = 0.725$, $p = .47$), CAN scores (12.76 vs. 13.12, $t(102) = −0.614$, $p = .541$), or non-CAN scores (11.56 vs. 11.08, $t(102) = 0.667$, $p = .506$). We also conducted a between-participant repeated-measures ANOVA to further examine possible effects of language

on task performance. Lesion size was included as a covariate in the analysis. We found no effect of Language, $F(1, 101) = 2.457$, $p = .120$. Thus, because no difference in performance was detected, we concluded that the task reflected the same construct in both groups, which justified merging the two groups.

Neuroimaging Data

For the lesion–symptom mapping, we used a parcellated brain atlas described by Faria et al. (2012) as a framework for anatomical boundaries. We examined the association between proportional damage in language-related ROIs and total score (number of correct items out of 45 total), CAN score, and non-CAN score (number of correct items out of 15 for each). To preserve a conservative ratio between the number of subjects and independent variables (roughly 10:1), we confined the analysis to 10 ROIs emphasized in previous studies (e.g., Pillay et al., 2017; Magnusdottir et al., 2013; Thothathiri et al., 2012), including Broca's area, temporo-parietal areas, and the temporal poles: IFG pars opercularis (IFGop) and pars triangularis (IFGtri), SMG, AG, STG, pole of STG (STGpole), MTG, pole of MTG

(MTGpole), posterior STG (pSTG), and pMTG. The anatomical ROIs used in the analysis are areas shown to be involved in either the dorsal stream (form-to-articulation pathway) or the ventral stream (form-to-meaning pathway) of speech processing obtained by prior work (Fridriksson et al., 2016). The areas and their anatomical boundaries are presented in Table 3.

We used clinical scans obtained as part of routine clinical care. In the Icelandic sample, MRI data were acquired on a 1.5-T MRI Siemens scanner. We obtained T1-weighted, diffusion-weighted imaging (DWI) and fluid-attenuated inversion recovery scans. A trained neurologist with extensive experience with lesion–symptom mapping demarcated brain lesions on DWI images, using fluid-attenuated inversion recovery and T1 images to help guide lesion boundaries. Spatial processing was conducted using SPM5 (Wellcome Institute of Imaging Neuroscience). DWI images were coregistered to the individual's T1 scan. This transform was applied to the lesion map. Then, the T1 image was warped to standard space (MNI152 template for older adults [mean age = 65 years]) using SPM5's unified segmentation and normalization algorithm. Lesion masks were smoothed at 8-mm FWHM and used as a cost-function mask to decrease the effect of abnormal tissue in the computation of normalization parameters (Andersen, Rapcsak, & Beeson, 2010; Brett, Leff, Rorden, & Ashburner, 2001). The transforms were applied to both the T1 scans and lesion maps for each participant, with the resulting images resliced to an isotropic 2 mm in standard Montreal Neurological Institute space to allow for voxelwise statistical analysis across participants.

In the U.S. sample, MRI data (T2 and DWI) were acquired on one of the following scanners: 1.5 General Electric scanner, 1.5-T Siemens scanner, or 3-T Siemens scanner. Imaging parameters were chosen by clinicians at the time of scanning and varied from one session to another. Lesions were demarcated by a neurologist and an assistant trained by the neurologist on MRI images in which the lesion was clearly visible. In most cases, lesions were drawn on T2-weighted structural scans using the MRIcro software (Rorden & Brett, 2000; www.mricro.com). In cases where T2-weighted images were not available or the lesion was not visible on these scans, the lesion was drawn on DWI images using MRIcro software. Spatial processing was conducted using the Clinical Toolbox (Rorden, Bonilha, Fridriksson, Bender, & Karnath, 2012) with MATLAB 2014b and SPM12. In most cases, we leveraged the Clinical Toolbox's "MR segment-normalize" function and specified the patient's T1, T2, and lesion files as inputs (Rorden et al., 2012). Normalization of the resulting binary lesion map was conducted using enantiomorphic normalization (Nachev, Coulthard, Jäger, Kennard, & Husain, 2008), and the resulting image was resliced to $1 \times 1 \times 1$ mm isotropic resolution. Lesions from patients without T1 scans, but with CT scans, were normalized using the Clinical Toolbox's "CT normalize" feature. This procedure adjusts the brightness of the CT scan to match the MNI152 template described above, invoked SPM's standard normalization function (ensuring that the process is not affected by damaged tissue by including a dilated lesion mask), and then converted the scan back to Hounsfield units. This process generated a normalized binary lesion mask along with a normalized CT image.

Table 3. Language-Related ROIs (Fridriksson et al., 2016) and the Number of Participants with Damage to Each Region

|	Area	Participants with Damage to this Area	Mean (SD) Proportional Damage	Range
Dorsal stream				
	IFGop	38	.2818 (.3081)	.0008–.9942
	IFGtri	33	.2083 (.2766)	.0000–.8617
	SMG	50	.1742 (.2641)	.0000–.9373
Ventral stream				
	AG	34	.1940 (.2110)	.0015–.7343
	STG	40	.1925 (.2866)	.0000–.9941
	STGpole	25	.1731 (.2832)	.0001–.9563
	MTG	18	.2225 (.3082)	.0000–.8622
	MTGpole	9	.4242 (.3413)	.0005–.8808
	pSTG	39	.2058 (.2892)	.0001–.9952
	pMTG	31	.1652 (.2394)	.0003–.9839
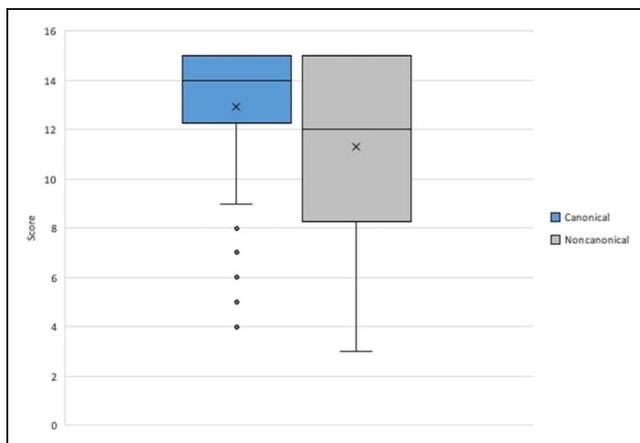
Figure 1. Boxplot showing the mean scores on CAN and non-CAN sentence structures. Lines represent median score, x is the mean score, boxes include range for scores in Quartiles 2 and 3, lines represent marginal quartiles, and dots indicate values that deviate from the overall distribution of scores.

Data Analysis

To test the hypothesis that damage to the pMTG was predictive of syntactic processing, our main analysis relied on a stepwise regression analysis using proportional damage in 10 predefined language ROIs (Fridriksson et al., 2016; Hickok & Poeppel, 2007; Table 3) as independent variables and total score, CAN scores, and non-CAN scores as dependent variables in three separate analyses. Furthermore, the same analyses were run again for non-CAN scores with CAN scores as a covariate. This analysis was conducted to isolate damage associated with impaired performance on non-CAN sentence structures, when factoring out the performance on CAN sentence structures. In addition, analyzing the non-CAN residual scores aims to alleviate lexical–semantic and working memory influences inherent in our sentence comprehension task (a similar method has been used in previous studies, e.g., Magnusdottir et al. [2013]; Pillay et al.

[2017], and Rogalsky et al. [2018]). For stepwise method criteria, we used probability of F with an entry alpha of .05 and a removal alpha of .10. Finally, we replicated all analyses using the enter procedure separately with frontal areas (pars triangularis and pars opercularis) and posterior temporal areas (pMTG and pSTG) to further explore variance in performance accounted for by each area. This procedure includes all independent variables in the regression equation at the same time, thus enabling an objective assessment of each predictor when holding other predictors constant. Therefore, this analysis allows us to examine variance in performance accounted for by posterior and frontal areas independently. Total lesion volume was included as a covariate in all analyses. Analyses were performed in SPSS 25 (International Business Machines, 2017), and an alpha level of .05 was used for all statistical tests.

RESULTS

Of the 104 study participants, 13 achieved a perfect score (45/45) on the syntactic processing test. Participants' mean total score was 35.76 ($SD = 9.18$) points with a range of 12–45 points. Scores on CAN sentences ranged from 4 to 15 points (out of 15 points) with a mean of 12.93 points ($SD = 2.99$), and scores on non-CAN sentences ranged from 3 to 15 points (out of 15 points) with a mean of 11.33 points ($SD = 3.62$). Patients scored significantly higher on CAN sentence types than non-CAN on average (mean difference of 1.61 [$SD = 2.27$], $t(103) = 7.21$, $p < .001$; see Figure 1). Using a χ^2 test, we found that ceiling effects occurred significantly more frequently on CAN sentence types than non-CAN sentence types (47 vs. 29, $\chi^2 = 12.03$, $p < .001$).

The lesion overlay map presented in Figure 2 shows that the lesions covered frontoparietal–temporal areas with the greatest overlap in the insula ($n = 29$). Areas

Figure 2. Lesion overlay map ($n = 104$; maximum overlap: $n = 29$) showing the lesion distribution of the participants. Warmer colors indicate greater overlap.
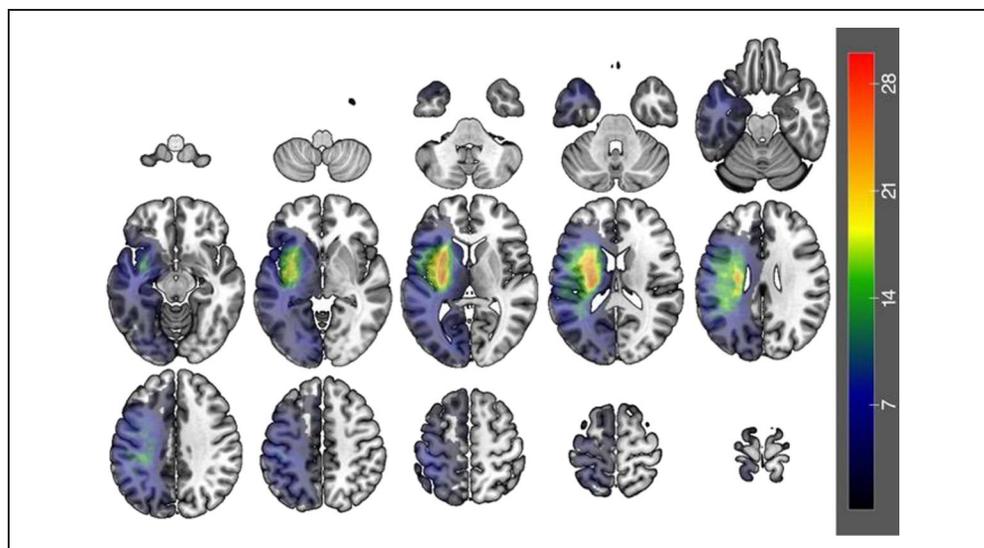
implicated in sentence processing, as described in the Introduction, including Broca's area (IFGop: $n = 38$; IFGtri: $n = 33$), pMTG ($n = 31$), and ATLs (MTGpole: $n = 9$; STGpole: $n = 25$), are all lesioned in a number of participants (see Table 3).

The stepwise regression procedure yielded significant prediction models in all analyses. Total score was best predicted by a model including proportional damage to pMTG, AG, and IFGtri, respectively, $F(3, 103) = 30.9$, adjusted (adj.) $R^2 = .466$, $p < .001$. All beta values were negative, indicating that damage to these areas leads to poorer performance on the sentence processing task. Damage to the pMTG alone accounted for 33.7% of the variance in performance. Damage to AG accounted for an additional 12.2%; and that to IFGtri, for 2.3%. CAN scores were best predicted by damage to AG, MTG, MTGpole, SMG, and STG, $F(5, 103) = 16.4$, adj. $R^2 = .427$, $p <$ .001. Beta values for AG and MTG were negative, and damage to these areas accounted for 24.8% and 9.5% of the variance in performance, respectively. Similarly, the beta value for SMG was negative, and the area accounted for an additional 2.5% of the variance. MTGpole (+6.2%) and STG (+2.6%) had positive beta values, which indicates that worse performance was associated with intactness in these areas, when accounting for damage to AG, MTG, and SMG. Non-CAN score was best predicted by damage to pMTG, AG, IFGop, STG, and MTGpole, $F(5, 103) = 23.7$, adj. $R^2 = .524$, $p < .001$. The pMTG alone accounted for 37.1% of variance in performance. Stepwise inclusion of AG, IFGop, STG, and MTG accounted for additional 6.9%, 4.6%, 3.2%, and 3.1%, respectively. All beta values were negative, except for STG where poor performance was associated with intactness of the region Figure 3.

Table 4. Stepwise Regression Models for Performance on the Sentence Processing Task

Scores (n = 104)	ROIs	R^2 Change	B	SE B	β	t	F Value	Adj. R^2	p Value
Total score							30.9	.466	<.001
	pMTG	.337	−17.24	2.70	−.50	−6.38			<.001
	AG	.122	−15.86	3.50	−.35	−4.54			<.001
	IFGtri	.023	−5.94	2.84	−.16	−2.09			.039
CAN score							16.4	.427	<.001
	AG	.248	−6.51	1.35	−.40	−4.81			<.001
	MTG	.095	−10.15	1.78	−.81	−5.69			<.001
	MTGpole	.062	4.21	2.08	.35	2.03			.045
	SMG	.025	−2.84	1.03	−.26	−2.76			.007
	STG	.026	3.48	1.62	.34	2.15			.034
Non-CAN score							23.7	.524	<.001
	pMTG	.371	−10.44	1.75	−.79	−5.98			<.001
	AG	.069	−4.24	1.26	−.24	−3.35			.001
	IFGop	.046	−3.70	0.89	−.32	−4.18			<.001
	STG	.032	6.73	1.83	.61	3.67			<.001
	MTG	.031	−4.28	1.66	−.34	−2.57			.012
Non-CAN residuals							13.6	.329	<.001
	pSTG	.192	−4.16	0.73	−.55	−5.73			<.001
	AG	.040							
	IFGop	.038	−5.90	1.43	−.74	−4.13			<.001
	SMG	.053	3.49	0.84	.45	4.13			<.001
	AG[a]	−.017							
	IFGtri	.049	4.60	1.68	.46	2.73			.008

Independent variables included pMTG, MTG, IFGop, IFGtri, SMG, AG, STG, STGpole, MTGpole, and pSTG. B = unstandardized beta value; $SE\ B$ = standard error of beta; β = standardized beta.

[a]Variable excluded.

Figure 3. Resulting stepwise regression models predicting total scores, canonical scores, noncanonical scores, and noncanonical residual scores. Warmer colors indicate a stronger predictive value (total score = pMTG, AG, IFGtri; CAN = AG, MTG, MTGpole, SMG, STG; non-CAN = pMTG, AG, IFGop, STG, MTG; non-CAN residuals = pSTG, IFGop, SMG, IFGtri).
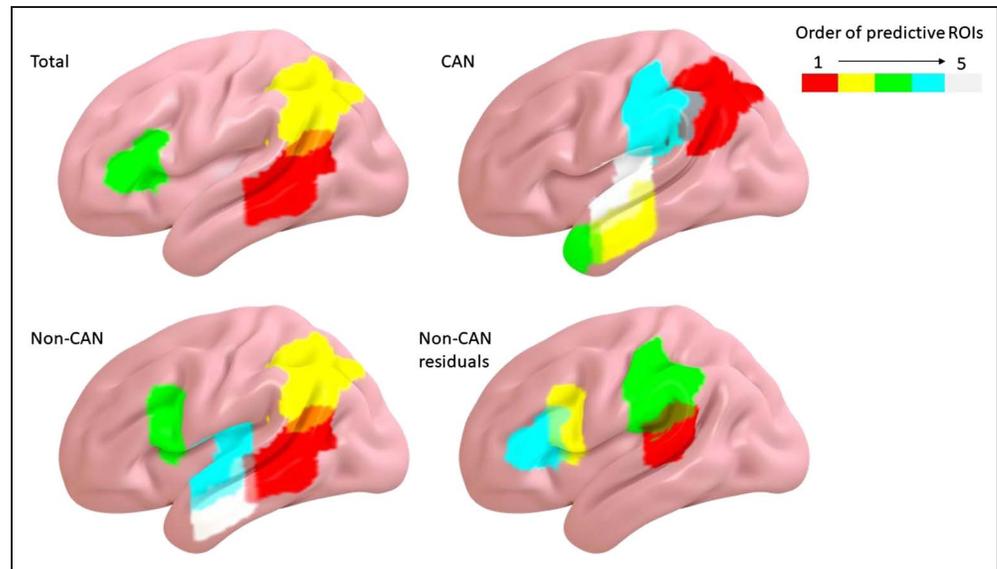
When variance accounted for by performance on CAN sentence structures was regressed out, the model predictive of non-CAN residual score included pSTG, IFGop, SMG, and IFGtri, $F(4, 103) = 13.6$, adj. $R^2 = .329$, $p < .001$. The pSTG accounted for 19.2% of the variance. Although AG was included as a significant predictor in the second step of building the model, this region was disregarded in the last step of model building as it did not contribute to the prediction over and above areas added to the model in later steps. IFGop accounted for an additional 3.8% of the variance; SMG, for 5.3%; and IFGtri, for 4.9%. Beta values for pSTG and IFGop were negative, but beta values for SMG and IFGtri were positive. This suggests the lesion profile predictive of complex syntactic processing relies on damage to pSTG and IFGop, accounting for the intactness of SMG and IFGtri. Full models are presented in Table 4.

In an effort to directly compare the relative contribution of damage to posterior temporal areas and frontal areas (Broca's area) in predicting performance on our task, we created two types of models representing the areas: a posterior model including pMTG and pSTG as independent variables and a frontal model including IFGop and IFGtri. The posterior model accounted for 32.6% of the variance in total score, compared with 1.5% accounted for by the frontal model. Beta values for pMTG and pSTG were negative, and pMTG was the only significant predictor ($p < .001$). The beta value was negative for IFGop and positive for IFGtri; however, neither was significant. For CAN, the posterior model accounted for 13.0% of the variance; pMTG had a negative beta value and was a significant predictor ($p = .005$). On the contrary, the frontal model yielded no significant predictors and essentially did not account for any variance. The posterior model accounted for 36.1% of the variance in non-CAN performance. Both predictors had negative beta values, but only pMTG was significant ($p = .001$).

In comparison, the frontal model accounted for 5.2% of the variance; IFGop was significant with a negative beta value, whereas IFGtri was not significant and had a positive beta value. Finally, 17.9% of the variance in non-CAN residual scores was accounted for by the posterior model; both predictors were negative, but only pSTG was significant. The frontal model accounted for 10.4% of the variance; both predictors were significant, but only IFGop had a negative beta value. Full posterior and frontal models are presented in Table 5.

An alternative way to analyze our data is to run a logistic regression with proportional task accuracy as a dependent variable. The motivation for implementing this approach comes from data suggesting that regression analysis of forced-choice variables can yield spurious results (e.g., spurious null results and/or spurious significances beyond the normal chance of Type 1 and Type 2 errors; Jaeger, 2008). We therefore constructed four stepwise logit models for proportional accuracy on canonical and noncanonical sentence types (participants' score divided by 15) and total score (divided by 45), in addition to analyzing noncanonical residual scores. To convert residual scores to meaningful proportions, we included proportional accuracy on canonical sentences as a covariate in a model with proportional accuracy on noncanonical sentences as a dependent factor. Statistical analyses were carried out in R statistical software using the "stepAIC()" function. This function bases model selection on Akaike's information criterion (AIC), as opposed to probability of F in our main analysis. Model parameters and independent variables were the same as before.

All four models are presented in Table 6. It should be emphasized that the stepwise procedure implemented in the logit models bases model selection on AIC, not p value for independent factors. As AIC estimates the trade-off between goodness of fit and model simplicity

Table 5. Enter Regression Models for Task Performance

Scores (n = 104)	ROIs	β	t	F Value	Adj. R²	p Value
Total score						
Posterior				25.9	.326	<.001
	pMTG	−.52	−3.37			.001
	pSTG	−.07	−0.47			.639
Frontal				1.8	.015	.171
	IFGop	−.31	−1.71			.090
	IFGtri	.19	1.01			.315
CAN						
Posterior				8.7	.130	<.001
	pMTG	−.51	−2.89			.005
	pSTG	.15	0.87			.384
Frontal				0.2	−.016	.823
	IFGop	−.02	−0.12			.907
	IFGtri	−.04	−0.23			.819
Non-CAN						
Posterior				30.1	.361	<.001
	pMTG	−.52	−3.47			.001
	pSTG	−.11	−.70			.484
Frontal				3.80	.052	.026
	IFGop	−.46	−2.08			.013
	IFGtri	.28	1.57			.120
Non-CAN residuals						
Posterior				12.2	.179	<.001
	pMTG	−.10	−0.59			.557
	pSTG	−.35	−2.08			.040
Frontal				7.0	.104	.001
	IFGop	−.64	−3.65			<.001
	IFGtri	.47	2.68			.009

Posterior models included pMTG and pSTG, and frontal models included IFGop and IFGtri.

(i.e., minimizes risk of overfitting or underfitting of the data), no factor in the model is required to reach significance at a given alpha level. In keeping with this, the logit models are largely consistent with our previous models. A model including lesion volume and damage to pMTG most accurately predicted within-sample total scores. Both coefficients were negative, suggesting that greater lesion volume and higher proportional damage are associated with lower scores. Canonical scores were best predicted by a model containing lesion volume (negative coefficient) and STG (positive coefficient). Noncanonical scores were best predicted by proportional damage to pMTG and IFGop. Both coefficients were negative, and proportional damage to pMTG was significantly predictive of scores when proportional damage to IFGop was included in the model. Unsurprisingly, proportional

Table 6. Stepwise Logistic Regression for Task Performance Accuracy

Accuracy Scores (n = 104)	Factors	B	SE B	z Value	p Value	AIC	Steps
Total score						79.08	8
	Lesion volume	$-7.20 \times e^{-5}$	$5.54 \times e^{-5}$	-1.30	.193		
	pMTG	-1.82	1.92	-0.95	.344		
CAN score						58.23	8
	Lesion volume	$-1.48 \times e^{-4}$	$7.07 \times e^{-5}$	-2.10	.036		
	STG	1.21	1.87	0.65	.519		
Non-CAN score						98.92	8
	pMTG	-3.83	1.77	-2.17	.030		
	IFGop	-1.42	0.93	-1.53	.126		
Non-CAN residuals						70.59	9
	CAN accuracy	4.27	1.28	3.35	<.001		
	pMTG	-1.64	1.79	-0.92	.358		

Independent variables included pMTG, MTG, IFGop, IFGtri, SMG, AG, STG, STGpole, MTGpole, and pSTG.

accuracy on canonical sentences was the strongest predictor in the noncanonical residual model, followed by proportional damage to pMTG. The interpretation of this last model can be rephrased as follows: When controlling for variance accounted for by performance on canonical sentences, the pMTG is the strongest predictor for proportional accuracy on noncanonical sentences.

We furthermore performed a post hoc analysis to explore the comparability of the Icelandic and U.S. samples. This analysis revealed that the lesion distribution between samples was largely similar, although the Icelandic sample had greater coverage in the temporal poles (STGpole: $n = 17$ and MTGpole: $n = 6$, compared with $n = 8$ and $n = 3$ in the U.S. sample, respectively) and a greater overlap in frontal regions (IFGop: $n = 25$ and IFGtri: $n = 21$, compared with $n = 13$ and $n = 12$; Figures S1 and S2). A separate regression analysis for each sample showed that the temporo-parietal area was strongly involved in sentence processing in both samples, and frontal areas also contributed in both samples (Tables S1 and S2). The main differences between the samples were greater implication of the temporal poles in the Icelandic sample (similar to Magnusdottir et al.'s [2013] results) compared with the U.S. sample, which may be driven by the greater lesion load in the temporal poles in the Icelandic sample. It should be noted that this analysis should be regarded as a validation for pooling the samples together. This analysis is not reliable in estimating the association between brain damage and syntactic processing as the ratio of partici-

pants to independent variables is heavily reduced, that is, roughly five participants per each independent variable (5:1).

DISCUSSION

This study examined the association between localized brain damage in predefined language ROIs and impaired receptive syntactic processing in acute stroke survivors. Participants completed a sentence–picture matching test and underwent neuroimaging to identify focal brain lesions. Using a hypothesis-driven approach, we found that lesions to pMTG were associated with overall syntactic processing scores and were strongly associated with performance on complex sentence structures. Furthermore, damage to the pSTG was found to be the factor most strongly associated with scores on noncanonical sentence structures when performance on canonical sentences was factored out. Therefore, consistent with previous lesion–symptom mapping studies (Den Ouden et al., 2019; Rogalsky et al., 2018; Pillay et al., 2017; Magnusdottir et al., 2013; Thothathiri et al., 2012; Dronkers et al., 2004) and supporting Matchin and Hickok's (2019) recent neuroanatomical framework for syntax, the current study adds to a growing evidence base indicating the importance of posterior temporal areas for syntactic processing over and above other cortical language areas.

Although most previous large-scale VLSM studies have included patients with chronic stroke (>6 months poststroke; e.g., Den Ouden et al., 2019; Rogalsky et al., 2018;

Pillay et al., 2017; Thothathiri et al., 2012; Dronkers et al., 2004), only a few studies have included participants in the acute phase of aphasia (Magnusdottir et al., 2013; Newhart et al., 2012; Race et al., 2012). Examining chronic patients can help identify functional modules where poststroke damage cannot be compensated for via plasticity. On the other hand, the main benefit of exploring the neural correlates of syntactic processing in participants in the acute phase of aphasia as opposed to chronic aphasia is that stroke-induced neural reorganization has not taken place, and more focal brain areas may be implicated in a given task (Ochfeld et al., 2010). Therefore, the population influences the implications that can be inferred from a study.

In our discussion, we will focus on results pertaining to non-CAN and non-CAN residual scores, for these represent the syntactic processing component of interest. Namely, it is difficult to distinguish syntactic processing abilities from word-level auditory comprehension in tasks that require passively listening to canonical sentences. Furthermore, almost half of our participants scored at ceiling ($n = 47$) on CAN sentences, which limits the variability in scores that can be associated with damage to particular ROIs.

In terms of our first hypothesis, the analysis revealed that a lesion to the pMTG is the single strongest predictive factor for total score, explaining 33.7% of the variance in performance. The pMTG is also the strongest predictor for performance on noncanonical sentences, explaining 37.1% of the variance in performance. In addition, proportional damage to the AG increased the explanatory power of the regression models for both total score and noncanonical sentence types (by 12.2% and 6.9%, respectively). We investigated our hypothesis further by exploring which language ROIs were predictive of noncanonical scores when the variance explained by performance on canonical scores was factored out (non-CAN > CAN). This analysis sought to account for the processing of word order represented by performance on canonical sentence structures. Noncanonical sentences are derived by the movement of syntactic constituents from their canonical position to a different position. As the movement requires more task-related resources, for example, holding syntactic information active in working memory to determine thematic roles, noncanonical sentences are presumably harder to process than canonical structures (Thothathiri et al., 2012; Grodzinsky & Santi, 2008). Accordingly, part of the agrammatic profile is a pattern in which patients show a disproportionate discrepancy in performance on the two sentence structures. Performance of participants in the current study was consistent with this pattern. Our primary analysis found that noncanonical residual scores were best predicted by proportional damage to the pSTG, which explained 19.2% of the variance in performance, whereas the full model implicated a distributed area including temporo-parietal areas and frontal areas.

Importantly, our main findings need to be evaluated in context of the direction of the beta weights in the best-fitting regression models. The prediction models for both non-CAN and non-CAN > CAN scores include negative and positive beta values. Negative beta values associate damage in a particular area with poorer performance, whereas positive beta values associate damage with better performance. From a biological perspective, it seems unlikely that damage in a given area enhances task performance. A more plausible explanation is that positive betas are observed in ROIs surrounding areas where damage critically impairs task performance or in areas left unaffected by the lesion. As an example, a lesion to posterior temporal areas may leave the frontal lobe intact, thus resulting in damage to frontal areas relating inversely with task performance. Therefore, an interpretation of the relative influence of damage to a given ROI on task performance in our study requires a simultaneous examination of the relationship between these ROIs and reference to existing theoretical background (see Geller, Thye, & Mirman, 2019, for a similar discussion of positive beta weights in interpreting task performance in patients with stroke).

As for non-CAN scores, STG had a positive beta weight in the full model. Unsurprisingly, proportional damage to STG correlates highly with damage to MTG and pMTG ($r = .84$, $p < .001$ and $r = .77$, $p < .001$, respectively). Considering that damage to the STG correlates negatively with non-CAN scores ($r = -.309$, $p < .001$), the positive beta weight in this case should not be taken to indicate that damage to STG improves syntactic processing. Rather, the model indicates that, whereas damage to pMTG, AG, IFGop, and MTG is held constant, damage to the STG does not contribute in the sense expected, but intactness in this ROI is associated with poor performance within the constraints of the model. A similar scenario seems to apply for the non-CAN residuals model. Although damage to all ROIs individually correlates negatively with the residual scores, beta weights for SMG and IFGtri are positive. Damage to SMG correlates highly with pSTG damage ($r = .516$, $p < .001$), and damage to IFGtri similarly correlates highly with damage to IFGop ($r = .843$, $p < .001$). Thus, while damage to pSTG, which most strongly impacts performance as represented by the amount of variance explained (19.2%), and IFGop is held constant, the lesion profile predictive of performance requires intactness of SMG and IFGtri.

The lesion-to-symptom profile is further established in our analysis of the contribution of damage to posterior versus frontal ROIs (Table 5) and the logistic regression for proportional task accuracy (Table 6). Focusing on non-CAN and non-CAN residual scores, our results clearly show that damage to posterior areas impacts performance to a much greater extent than damage to frontal areas. In comparing the contribution of posterior versus frontal ROIs, damage to pMTG and pSTG alone accounts for 36.1% of the variance in non-CAN scores and 17.9% of

variance in non-CAN residual scores, compared with 5.2% and 10.4% accounted for by frontal areas, respectively. Interestingly, only one of the posterior areas reaches significance in both models: pMTG for non-CAN scores ($p = .001$) and pSTG for non-CAN residual scores ($p = .04$). A high correlation between damage to pMTG and pSTG ($r = .818$, $p < .001$) is a likely cause, but the shift from pMTG as the strongest predictor for non-CAN scores to pSTG as the strongest predictor for non-CAN residual scores is nonetheless notable. The stepwise logistic regression, although based on different criteria from the other analyses (AIC), similarly suggests that damage to posterior temporal areas (in this case, pMTG) is associated with poorer task performance.

It should be noted that our results differ from results by Magnusdottir et al. (2013), although 54 of the participants in the current study were included in Magnusdottir et al.'s study. This difference is most likely because of an increase in power with the addition of 50 participants. However, our finding is consistent with Rogalsky et al.'s (2018) VLSM study on non-CAN > CAN performance in 66 individuals with focal lesions, where the pSTG was similarly strongly associated with performance. This may suggest that the posterior temporal and temporo-parietal areas provide the additional cognitive resources required for parsing the syntactic movement in noncanonical sentence structures, thus establishing its crucial role in complex syntactic processing, although not exclusively, as other areas also contribute to the overall performance (as discussed below).

Contrasting performance on canonical and noncanonical sentence structures is not without limitations. The main limitation may be that this analysis does not entail a fine-grained examination of particular sentence structures. To investigate more meticulously the areas involved in complex syntactic processing, we explored mean performance on each sentence type separately. We found the greatest discrepancy in performance was in subject versus object extracted sentences (Types 4 and 8 [$M = 86.5\%$ and 87.9% correct] vs. Types 5 and 9 [$M = 71.2\%$ and 74.4% correct], respectively). Thus, we performed a post hoc analysis contrasting performance on cleft sentences with subject gaps and referential *wh*-questions with subject gap and main verbs (canonical sentences) to the same sentence structures but with object gaps rather than subject gaps (noncanonical sentences). Although providing a more detailed account of sentence processing, this analysis supported our main analysis. For non-CAN residual scores, the analysis revealed a similar model explaining 41.1% of the variance in scores. The pSTG was identified as the strongest predictor, accounting for 26.2%. Other ROIs included in the model were SMG (+4.1%), IFGop (+5.6%), and STG (+6.7%). Model parameters are presented in Table S3.

The posterior perisylvian area has been implicated in various facets of language comprehension since Wernicke (1874) identified that comprehension impairment resulted from damage to the left posterior temporal cortex. Recent studies emphasize the importance of the posterior peri-sylvian area (particularly the pSTG and SMG regions) for phonological working memory, sensorimotor processing, and acoustic–phonetic processing (Dronkers et al., 2004; Hickok & Poeppel, 2004; Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003). Hickok and Poeppel (2004) concluded that these processes support comprehension by keeping auditory, phonological, and lexical representations temporarily active. Our results indicate that the posterior temporal (including the pMTG) and temporo-parietal areas additionally support complex syntactic processing over and above other language-related regions, lending support to an emerging evidence base of findings from lesion-based studies suggesting that the posterior perisylvian area, including pMTG and pSTG, may be crucial for complex sentence comprehension (Den Ouden et al., 2019; Matchin & Hickok, 2019; Fridriksson et al., 2018; Rogalsky et al., 2018; Pillay et al., 2017; Thothathiri et al., 2012; Dronkers et al., 2004). Importantly, these results do not contradict results from functional neuroimaging studies, as neuroimaging studies in healthy participants, including studies implicating activation in Broca's area in sentence processing tasks, also reliably show activation of the left pSTG and pMTG (Hagoort & Indefrey, 2014; Den Ouden et al., 2012; Indefrey, 2012).

Consistent with our second hypothesis, damage to Broca's area was not found to be strongly associated with impaired syntactic processing. This finding contradicts some previous findings that have implicated Broca's area as critical for syntactic processing (e.g., Caramazza & Zurif, 1976) or further as a common neuroanatomical origin of syntax for both comprehension and production (e.g., Grodzinsky, 2000; Swinney & Zurif, 1995). Our findings may be more consistent with claims that Broca's area is involved in sentence comprehension through other cognitive processes. Notably, in later steps of the model, both IFGop and IFGtri were identified as predictors in the model predicting noncanonical residuals (Table 4). The involvement of these areas in complex syntactic processing may suggest that Broca's area plays a complementary role to temporo-parietal areas in sentence comprehension (for a recent discussion, see Matchin & Hickok, 2019). Although most of the evidence for the contribution of Broca's area to sentence comprehension has come from functional neuroimaging studies (e.g., Santi & Grodzinsky, 2007; Bornkessel et al., 2005; Ben-Shachar et al., 2003; Caplan et al., 1998), an ongoing debate revolves around the specific role of the area in language comprehension. This debate is reflected in opposing lines of evidence arguing that activation in Broca's area during sentence comprehension tasks is driven by task-related working memory load (Fiebach et al., 2005; Kaan & Swaab, 2002), cognitive control (Novick et al., 2005), or articulatory rehearsal (Rogalsky et al., 2008, 2015), to name a few examples. Notably, Hagoort and

Indefrey's (2014) meta-analysis of 151 hemodynamic studies on sentence processing found that Broca's area and posterior temporal areas were reliably activated in processing demanding sentences. The study concluded that syntactic processing relies on the coactivation of neuronal populations in a network of posterior frontal and temporal regions. However, as previously noted, lesion–symptom mapping studies have consistently found that damage to the posterior temporal lobe, and not Broca's area, is associated with poor performance on complex syntactic processing tasks (Den Ouden et al., 2019; Rogalsky et al., 2018; Fridriksson et al., 2018; Pillay et al., 2017; Magnusdottir et al., 2013; Thothathiri et al., 2012; Dronkers et al., 2004).

In an attempt to settle the dispute and determine the role of Broca's area in sentence comprehension, Thothathiri et al. (2012) performed a post hoc analysis on performance on two-proposition minus one-proposition sentence scores and found the highest t value for this analysis in pars opercularis of Broca's area. The authors argue that this region contributes to sentence comprehension through additional task-related resources (e.g., working memory) required for two-proposition compared with one-proposition sentences. Similarly, Pillay et al. (2017) included a picture–naming task as a covariate in their analysis to control for language components not specific to sentence comprehension, for example, single-word semantic, phonological, executive, and articulatory processing, finding that multiword integration processes necessary for sentence comprehension were not impaired by damage to Broca's area. Rogalsky et al. (2018) specifically explored agrammatic performance in participants and found no association between damage in Broca's area and agrammatic profile; nonetheless, the study did implicate Broca's area in the cognitive task demands of their sentence–picture matching task. Accordingly, our results may be taken to indicate that the involvement of Broca's area, and pars opercularis in particular, in predicting performance on non-CAN residuals in our sample is likely secondary to sentence comprehension through working memory or articulatory rehearsal, rather than being a crucial locus for receptive syntactic processing.

There are a few potential limitations to this study that should be taken into consideration when interpreting our results. Most notably, although we had a strong theoretical motivation for using an ROI-based approach in our analysis, this approach has some inherent limitations. Each ROI treated as an independent variable in our models is represented by a single number (proportional damage), regardless of the size of the ROI. Consequently, spatial resolution is compromised. The issue could be mediated by using a multivariate voxel-based approach, which retains greater spatial accuracy for associating brain damage with task performance. For full disclosure, we did utilize a multivariate approach (Support Vector Regression) on our data and found that no significant voxels survived correction for multiple comparisons. We believe the reason for this relates to the representation of

syntactic processing in the brain. Various brain areas have been implicated as being either "crucial" or "important" for syntactic processing—although no firm consensus has been reached within our field. A likely intermediate conclusion is that syntactic processing relies on a distributed network of brain areas. Given the nature of syntax and the complexity implicit in hierarchical structure building in sentence comprehension, this does not seem like an unlikely explanation. On the basis of this assumption, slight individual variability in neural organization of syntax combined with potential uncertainty inherent in the template registration process may make the coarser ROI-based analysis more robust than a voxel-based analysis. Considering our strong theoretical motivation within the framework of the DSM (Matchin & Hickok, 2019; Fridriksson et al., 2016; Hickok & Poeppel, 2004) alongside these assumptions, we believe our approach has the merit to allow us to robustly examine the relationship between sentence comprehension and brain damage.

A couple of other limitations require attention. First, we included both native speakers of Icelandic and English. Icelandic is morphologically more complex than other Germanic languages, and meaning is more dependent on independent morphological structures (Thráinsson, 2007). The difference between noncanonical and canonical sentence types often requires morphosyntactic analysis of case marking in addition to word order analysis alone in Icelandic, for example, subject clefts versus object clefts (*stelpan* vs. *stelpuna* for "girl," nominative vs. accusative; Magnusdottir et al., 2013). However, we considered the languages similar enough to justify pooling the samples together, as represented by a comparable distribution of scores between the groups (Icelandic vs. US; total score = 36.39 vs. 35.08, $p = .470$; CAN = 12.76 vs. 13.12, $p = .541$; non-CAN = 11.56 vs. 11.08, $p = .506$), similar lesion distribution (Figures S1 and S2), and the involvement of the temporo-parietal area in sentence processing identified in each sample separately (Tables S1 and S2). Second, this study only included a single measure of sentence comprehension. Although this procedure allows for an examination of an agrammatic behavioral pattern, more extensive measures of syntactic processing would allow for stronger conclusions to be made. Finally, behavioral performance was only examined in relation to structural damage in cortical brain areas, although the inclusion of cerebral perfusion imaging could provide additional information (Fridriksson, Richardson, Fillmore, & Cai, 2012; Hillis et al., 2001, 2002, 2004). Perfusion analysis, for example, would enable investigation of dysfunctional tissue beyond the infarct that might contribute to the deficit in some patients with acute stroke.

Taking into consideration the limitations of the current study and the disputed modularity of syntax in the human brain, we have shown that the temporo-parietal area is crucially implicated in complex syntactic processing. Furthermore, although our study does not resolve the

dispute over Broca's area in sentence processing, the results indicate that Broca's area plays a complementary role to the temporo-parietal area, rather than being crucial for complex syntactic processing. We therefore conclude that damage to the temporo-parietal area, including the pMTG, results in impaired sentence comprehension.

Reprint requests should be sent to Sigfus Kristinsson, Department of Communication Sciences & Disorders, Center for the Study of Aphasia Recovery, University of South Carolina, 915 Greene St., Columbia, SC 29208, or via e-mail: sigfushelgik@gmail.com.

Note

1. Supplementary material for this paper can be retrieved from https://cstar.sc.edu/sentence-processing-in-aphasia/.

REFERENCES

Andersen, S. M., Rapcsak, S. Z., & Beeson, P. M. (2010). Cost function masking during normalization of brains with focal lesions: Still a necessity? *Neuroimage, 53*, 78–84.

Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., & Saxe, R. (2011). Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences, U.S.A., 108*, 4429–4434.

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*, 433–440.

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *Neuroimage, 26*, 221–233.

Brett, M., Leff, A. P., Rorden, C., & Ashburner, J. (2001). Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage, 14*, 486–500.

Broca, P. (1861). Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole) [Notes on the Seat of the Faculty of Articulated Language, Following an Observation of Aphemia (Loss of Speech)]. In *Bulletins de la Societe anatomique (Paris) 2e series 6* (pp. 330–357).

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience, 10*, 541–552.

Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *Neuroimage, 9*, 343–351.

Caplan, D., Hildebrandt, N., & Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain, 119*, 933–949.

Caplan, D., Michaud, J., Hufford, R., & Makris, N. (2016). Deficit–lesion correlations in syntactic comprehension in aphasia. *Brain and Language, 152*, 14–27.

Caplan, D., Waters, G. S., & DeDe, G. (2007). Specialized verbal working memory for language comprehension. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory*. Oxford: Oxford University Press.

Caramazza, A., Capasso, R., Capitani, E., & Miceli, G. (2005). Patterns of comprehension performance in agrammatic Broca's aphasia: A test of the trace deletion hypothesis. *Brain and Language, 94*, 43–53.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R. S. (2001). Agrammatic Broca's aphasia is not associated with a single pattern of comprehension performance. *Brain and Language, 76*, 158–184.

Caramazza, A., & Hillis, A. E. (1989). The disruption of sentence production: Some dissociations. *Brain and Language, 36*, 625–650.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language, 3*, 572–582.

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

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

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126*, 1193–1201.

Den Ouden, D. B., Malyutina, S., Basilakos, A., Bonilha, L., Gleichgerrcht, E., Yourganov, G., et al. (2019). Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Human Brain Mapping, 40*, 2513–2173.

Den Ouden, D. B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., et al. (2012). Network modulation during complex syntactic processing. *Neuroimage, 59*, 815–823.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition, 92*, 145–177.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., et al. (2006). Epidemiology of aphasia attributable to first ischemic stroke. *Stroke, 37*, 1379–1384.

Faria, A. V., Joel, S. E., Zhang, Y., Oishi, K., van Zjil, P. C., Miller, M. I., et al. (2012). Atlas-based analysis of resting-state functional connectivity: Evaluation for reproducibility and multi-modal anatomy–function correlation studies. *Neuroimage, 61*, 613–621.

Ferro, J. M., Mariano, G., & Madureira, S. (1999). Recovery from aphasia and neglect. *Cerebrovascular Diseases, 9*(Suppl. 5), 6–22.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping, 24*, 79–91.

Fitch-West, J., Sands, E. S., & Ross-Swain, D. (1998). *Bedside Evaluation Screening Test–Second Edition (BEST-2)*. Austin, TX: Pro-Ed.

Fridriksson, J., Bonilha, L., & Rorden, C. (2007). Severe Broca's aphasia without Broca's area damage. *Behavioural Neurology, 18*, 237–238.

Fridriksson, J., den Ouden, D. B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., et al. (2018). Anatomy of aphasia revisited. *Brain, 141*, 848–862.

Fridriksson, J., Richardson, J. D., Fillmore, P., & Cai, B. (2012). Left hemisphere plasticity and aphasia recovery. *Neuroimage, 60*, 854–863.

Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., den Ouden, D. B., & Rorden, C. (2016). Revealing the dual streams of speech processing. *Proceedings of the National Academy of Sciences, U.S.A., 113*, 15108–15113.

Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences, 13*, 175–181.

Geller, J., Thye, M., & Mirman, D. (2019). Estimating effects of graded white matter damage and binary tract disconnection on post-stroke language impairment. *Neuroimage, 189*, 248–257.

Goodglass, H. (1997). Agrammatism in aphasiology. *Clinical Neuroscience*, *4*, 51–56.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, *23*, 1–21.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*, 474–480.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, *37*, 347–362.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67–99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–401.

Hillis, A. E., Wityk, R. J., Barker, P. B., Beauchamp, N. J., Gailloud, P., Murphy, K., et al. (2002). Subcortical aphasia and neglect in acute stroke: The role of cortical hypoperfusion. *Brain*, *125*, 1094–1104.

Hillis, A. E., Wityk, R. J., Tuffiash, E., Beauchamp, N. J., Jacobs, M. A., Barker, P. B., et al. (2001). Hypoperfusion of Wernicke's area predicts severity of semantic deficit in acute stroke. *Annals of Neurology*, *50*, 561–566.

Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Reexamining the brain regions crucial for orchestrating speech articulation. *Brain*, *127*, 1479–1487.

Indefrey, P. (2012). Hemodynamic studies of syntactic processing. In M. Faust (Ed.), *Handbook of the neuropsychology of language* (pp. 209–228). Malden, MA: Blackwell.

International Business Machines. (2017). *IBM SPSS Statistics for Windows (Version 25.0)*. Armonk, NY: Author.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, *6*, 350–356.

Kertesz, A. (2007). *Western Aphasia Battery–Revised (WAB-R)*. New York: Grune & Stratton.

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

Magnusdottir, S. (2000). *On grammatical knowledge in agrammatism: Evidence from Icelandic* (Unpublished doctoral dissertation). Boston University, Boston, MA.

Magnusdottir, S. (2005). *Setningafreadiprof (Test of syntax)*. Reykjavik, Iceland: Landspitali University Hospital.

Magnusdottir, S., Fillmore, P., den Ouden, D. B., Hjaltason, H., Rorden, C., Kjartansson, O., et al. (2013). Damage to left anterior temporal cortex predicts impairment of complex syntactic processing: A lesion–symptom mapping study. *Human Brain Mapping*, *34*, 2715–2723.

Matchin, W., & Hickok, G. (2019). The cortical organization of syntax. https://doi.org/10.31234/osf.io/6394f.

Mohr, J. P., Pessin, M. S., Finkelstein, S., Funkenstein, H. H., Duncan, G. W., & Davis, K. R. (1978). Broca aphasia: Pathologic and clinical. *Neurology*, *28*, 311–324.

Nachev, P., Coulthard, E., Jäger, H. R., Kennard, C., & Husain, M. (2008). Enantiomorphic normalization of focally lesioned brains. *Neuroimage*, *39*, 1215–1226.

Newhart, M., Trupe, L. A., Gomez, Y., Cloutman, L., Molitoris, J. J., Davis, C., et al. (2012). Asyntactic comprehension, working memory, and acute ischemia in Broca's area versus angular gyrus. *Cortex*, *48*, 1288–1297.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience*, *5*, 263–281.

Ochfeld, E., Newhart, M., Molitoris, J., Leigh, R., Cloutman, L., Davis, C., et al. (2010). Ischemia in Broca's area is associated with Broca's aphasia more reliably in acute than chronic stroke. *Stroke*, *41*, 325–330.

Pillay, S. B., Binder, J. R., Humphries, C., Gross, W. L., & Book, D. S. (2017). Lesion localization of speech comprehension deficits in chronic aphasia. *Neurology*, *88*, 970–975.

Race, D. S., Ochfeld, E., Leigh, R., & Hillis, A. E. (2012). Lesion analysis of cortical regions associated with the comprehension of nonreversible and reversible yes/no questions. *Neuropsychologia*, *50*, 1946–1953.

Rogalsky, C., Almeida, D., Sprouse, J., & Hickok, G. (2015). Sentence processing selectivity in Broca's area: Evident for structure but not syntactic movement. *Language, Cognition, & Neuroscience*, *30*, 1326–1338.

Rogalsky, C., LaCroix, A. N., Chen, K. H., Anderson, S. W., Damasio, H., Love, T., et al. (2018). The neurobiology of agrammatic sentence comprehension: A lesion study. *Journal of Cognitive Neuroscience*, *30*, 234–255.

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*, 14.

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H. O. (2012). Age-specific CT and MRI templates for spatial normalization. *Neuroimage*, *61*, 957–965.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioral Neuroscience*, *12*, 191–200.

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *Neuroimage*, *37*, 8–17.

Swinney, D., & Zurif, E. (1995). Syntactic processing in aphasia. *Brain and Language*, *50*, 225–239.

Thompson, C. K., Bonakdarpour, B., & Fix, S. F. (2010). Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. *Journal of Cognitive Neuroscience*, *22*, 1993–2011.

Thothathiri, M., Kimberg, D. Y., & Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion–symptom mapping in aphasia. *Journal of Cognitive Neuroscience*, *24*, 212–222.

Thráinsson, H. (2007). *The syntax of Icelandic*. Cambridge: Cambridge University Press.

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., et al. (2011). Left inferior frontal cortex and syntax: Function, structure and behaviour in patients with left hemisphere damage. *Brain*, *134*, 415–431.

Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Reorganization of syntactic processing following left-hemisphere brain damage: Does right-hemisphere activity preserve function? *Brain*, *133*, 3396–3408.

Wernicke, C. (1874). *Der aphasische symptomenkomplex*. Breslau, Poland: Cohn & Weigert.

Wilson, S. M., DeMarco, A. T., Henry, M. L., Gesierich, B., Babiak, M., Mandelli, M. L., et al. (2014). What role does the anterior temporal lobe play in sentence-level processing? Neural correlates of syntactic processing in semantic variant primary progressive aphasia. *Journal of Cognitive Neuroscience*, *26*, 970–985.

Wilson, S. M., & Saygin, A. P. (2004). Grammaticality judgment in aphasia: Deficits are not specific to syntactic structures, aphasic syndromes, or lesion sites. *Journal of Cognitive Neuroscience*, *16*, 238–252.