doi:10.1093/brain/awaa027 BRAIN 2020: 143; 862–876 | 862

BRAIN
A JOURNAL OF NEUROLOGY

Dissociation between frontal and temporal-parietal contributions to connected speech in acute stroke

Junhua Ding,[1] Randi C. Martin,[2] A. Cris Hamilton[3] and Tatiana T. Schnur[1,4]

Humans are uniquely able to retrieve and combine words into syntactic structure to produce connected speech. Previous identification of focal brain regions necessary for production focused primarily on associations with the content produced by speakers with chronic stroke, where function may have shifted to other regions after reorganization occurred. Here, we relate patterns of brain damage with deficits to the content and structure of spontaneous connected speech in 52 speakers during the acute stage of a left hemisphere stroke. Multivariate lesion behaviour mapping demonstrated that damage to temporal-parietal regions impacted the ability to retrieve words and produce them within increasingly complex combinations. Damage primarily to inferior frontal cortex affected the production of syntactically accurate structure. In contrast to previous work, functional-anatomical dissociations did not depend on lesion size likely because acute lesions were smaller than typically found in chronic stroke. These results are consistent with predictions from theoretical models based primarily on evidence from language comprehension and highlight the importance of investigating individual differences in brain-language relationships in speakers with acute stroke.

1 Department of Neurosurgery, Baylor College of Medicine, Houston, Texas, USA
2 Department of Psychological Sciences, Rice University, Houston, Texas, USA
3 Department of Institution Reporting, Research and Information Systems, University of Texas at Austin, Austin, Texas, USA
4 Department of Neuroscience, Baylor College of Medicine, Houston, Texas, USA

Correspondence to: Tatiana T. Schnur
Department of Neurosurgery, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA
E-mail: tschnur@bcm.edu

Keywords: acute stroke; lesion behaviour mapping; language; syntax; connected speech
Abbreviation: PC = principal components analysis

Introduction

In the past century and a half since Pierre Paul Broca's work with patient 'Tan', the study of language deficits in speakers with focal brain damage from stroke has been used to understand which brain regions are necessary for language. Recent applications of MRI analysis techniques allow better quantification of how focal brain damage impacts language deficits via lesion behaviour mapping. The lesion behaviour mapping approach provides critical evidence to clarify whether multiple regions identified in functional MRI studies are epiphenomenal or necessary for function. Here, for the first time we identify the brain regions required for producing words and organizing them during spontaneous connected speech before brain-behaviour reorganization occurs in a large group of speakers with focal acute left hemisphere stroke.

Left hemisphere brain damage often impairs language production demonstrating that an extensive left frontal-temporal

Received October 29, 2019. Revised December 3, 2019. Accepted December 17, 2019. Advance access publication March 10, 2020

parietal network is necessary to produce multiple words in syntactically accurate structure (for reviews see Price, 2010; Fedorenko and Thompson-Schill, 2014; Hagoort and Indefrey, 2014). However, the processes for producing and organizing the content (i.e. words) into structure (lexical or syntactically accurate combinations) during connected speech localized within this left lateralized brain network are not well identified. This partly stems from a difficulty in eliciting spontaneous speech where deficits can be systematically quantified across individuals. As a result, almost all lesion behaviour mapping analyses of connected speech in speakers with stroke have been limited to analysing the content, i.e. the words elicited from picture descriptions or direct questions, or the fluency with which that content is produced (the number of words produced per minute or the number of words produced within utterances are often used as measures of the fluency of language production; Thompson *et al.*, 2013). In the first large lesion behaviour mapping study of connected speech impairments, Borovsky *et al.* (2007) asked 50 subjects with diagnosed aphasia after left hemisphere stroke a series of autobiographical questions. Damage to a wide range of regions within frontal, parietal, and temporal lobes was associated with a reduced number of words produced overall and within utterances. Damage specific to posterior loci in the middle and superior temporal gyri as well as the angular gyri was associated with deficits in producing a reduced variety of words. However, the contribution of stroke severity as measured by lesion volume was not analysed. Halai *et al.* (2017) examined similar aspects of connected speech during picture description using principal components analysis (PCA) in a smaller group of subjects ($n = 31$) with diagnosed aphasia after left hemisphere stroke (*cf.* Halai *et al.*, 2018). Here too, the ability to produce more words and more words quickly was associated with damage to frontal regions while reduced lexical diversity was associated with posterior regions including the middle and superior temporal gyri, and supramarginal gyrus. However, across these studies none of these regions survived correction for overall lesion volume. Using a multivariate lesion behaviour mapping model and large left hemisphere chronic aphasia stroke sample ($n = 90$), Yourganov *et al.* (2016) found the speech fluency scores estimated as part of the Western Aphasia Battery (Kertesz, 1982) were predicted by damage to both left anterior (middle frontal gyrus and the inferior frontal gyrus pars operculum) and posterior areas (supramarginal and posterior superior temporal gyri). However, here too lesion size predicted the fluency score and was highly correlated with the areas revealed. Significant findings controlling for the effect of lesion size were from a combined PCA and lesion behaviour mapping analysis which included not only the connected speech measures but other production and cognitive measures. In these subsequent analyses, fluency related factors were primarily associated with the frontal lobe (Halai *et al.*, 2017, 2018; Lacey *et al.*, 2017). Critically, however, the fluency-related PCA-derived factors were composed of additional variables with high loadings including naming, repetition, and reading among others, which complicates the relationship between fluency deficits and brain damage localization. When controlling for lesion size, this relationship disappeared when performing the PCA within only connected speech measures (Halai *et al.*, 2017). In summary, although there appears to be an anterior (frontal) to posterior (temporal-parietal) division in terms of the number of words produced (fluency) versus the ability to produce more diverse lexical content during spontaneous connected speech, the ability to produce content is best accounted for by stroke severity not the location of damage, i.e. the more brain damage incurred the fewer and less-diverse words produced, irrespective of where the damage occurred.

Relatively little is known about how individual differences in focal brain damage affect the ability to accurately produce syntactic structure independent from content. In the only lesion behaviour mapping study of content and structure deficits during spontaneous connected speech of which we are aware, Mirman *et al.* (2019) examined story-telling narratives from 46 speakers with diagnosed aphasia after chronic left hemisphere stroke. After controlling for the contribution of lesion size, the proportion of words produced in sentences (utterances with both a subject noun and verb) was associated with damage to left middle and inferior frontal gyri, postcentral gyrus, and inferior parietal lobe. However, no lesion pattern emerged to predict syntactic deficits measured by the number of closed-class words (e.g. determiners, prepositions) produced out of all words in the narrative. No other aspects of spontaneous connected speech were examined. It remains unknown whether damage to focal regions within the left hemisphere language network critical for producing content dissociates from damage to regions responsible for generating syntactic structure during spontaneous connected speech.

The current study

Here, we addressed the question of whether impairments producing the content and structure of spontaneously generated speech are related to different patterns of focal brain damage. Whether unique brain regions are required to produce content and structure during connected speech speaks to a central debate concerning the specificity of brain regions required for syntactic processing (Wilson and Saygın, 2004; Thothathiri *et al.*, 2012; Magnusdottir *et al.*, 2013; Blank *et al.*, 2016; Wilson *et al.*, 2016; Fedorenko *et al.*, 2018; Rogalsky *et al.*, 2018; *cf.* Dapretto and Bookheimer, 1999; Hagoort and Indefrey, 2014; Friederici *et al.*, 2017). To our knowledge, this is the first study of individual differences during the acute stage of stroke linking brain damage to deficits in the content and structure of spontaneous connected speech.

Our approach has significant methodological strengths. First, by examining lesion behaviour relationships in a large group of speakers identified with radiological signs of left hemisphere acute stroke, we increased variability in behavioural performance, lesion size, and lesion location and avoided the confound of brain-behaviour reorganization. Studies of speakers with chronic stroke often preselect subjects based on aphasia diagnosis (Schnur *et al.*, 2009;

Yourganov *et al.*, 2016; Halai *et al.*, 2017; Mirman *et al.*, 2019), which limits behavioural and lesion location variability. Speakers with chronic aphasia typically have large lesions spanning crucial language areas (e.g. inferior frontal gyrus, posterior superior temporal gyrus and angular gyrus) where adjacent cortical regions are often damaged and damage is correlated with overall lesion size (Ochfeld *et al.*, 2010; Yourganov *et al.*, 2016; Shahid *et al.*, 2017). As a result, it is difficult to disentangle how lesion size and damage to specific regions contribute to function. Further, function may have shifted to other regions after reorganization occurred (Weiller *et al.*, 1995; Marsh and Hillis, 2006; Saur *et al.*, 2006; Thompson and den Ouden, 2008; Turkeltaub *et al.*, 2012; Nardo *et al.*, 2017; Hartwigsen and Saur, 2019). Here, we increased power to assess whether damage to a region significantly affected behaviour in comparison to no damage to that region, by providing sufficient behavioural variability and heterogeneity for both between- and within- region lesion distributions (*cf.* Kimberg *et al.*, 2007; Sperber and Karnath, 2017; Lorca-Puls *et al.*, 2018; Pustina *et al.*, 2018; Sperber *et al.*, 2019). Second, we elicited speech using a spontaneous story-telling generation task, which provides the advantages of eliciting grammatically diverse spontaneous speech (Thompson *et al.*, 2010; den Ouden *et al.*, 2019) while providing excellent ecological validity with abilities to produce everyday speech (Olness and Ulatowska, 2017) and good test-retest reliability (Brookshire and Nicholas, 1994*a*, *b*; Roberts and Post, 2018). Third, we conducted detailed quantitative analyses to independently assess the content and structure of spontaneous speech using the rigorous quantitative production analysis approach (Saffran *et al.*, 1989; Rochon *et al.*, 2000; Gordon, 2006; Wilson *et al.*, 2010). Fourth, because single behavioural measures may consist of multiple cognitive components, we applied PCA to extract the underlying cognitive components across the quantitative production analysis measures (Rochon *et al.*, 2000; Butler *et al.*, 2014; Mirman *et al.*, 2015*a*; Lacey *et al.*, 2017). Lastly, we applied support vector regression based multivariate lesion behaviour mapping, which considers the pattern of all voxels as a single model to predict a behavioural outcome. In comparison to univariate approaches, multivariate lesion behaviour mapping ameliorates limitations from differential lesion distribution across voxels, type II error from applying statistical corrections across voxels and allows for interactions between different damaged areas to account for behaviour (Mah *et al.*, 2014; Zhang *et al.*, 2014; Yourganov *et al.*, 2016; Pustina *et al.*, 2018; Sperber *et al.*, 2019).

Materials and methods

Participants

Ninety acute stroke patients were consecutively recruited from the comprehensive stroke centres at the Memorial Hermann, Houston Methodist, and St. Luke hospitals in Houston, Texas as part of an ongoing project. Following previous studies (Karnath *et al.*, 2010; Corbetta *et al.*, 2015; Kristinsson *et al.*, 2020) we included participants diagnosed with an acute ischaemic or parenchymal haemorrhagic left hemisphere stroke. Further, we included patients if they were native English speakers and had no history of other significant neurological diseases (e.g. dementia, schizophrenia) as assessed by the clinical neurological care team and subsequently documented in electronic medical records. Informed consent was approved by the Baylor College of Medicine Institutional Review Board.

For the current study, 25 patients were excluded because they were not able to complete the spontaneous connected speech task ($n = 20$) or they produced unintelligible speech ($n = 5$). Sixty-five of the remaining patients (35 males; 54 right-handed; four haemorrhagic) had sufficient language production to complete the spontaneous connected speech task with no severe apraxia of speech to preclude accurate connected speech scoring. Mean apraxia of speech score (subtest 5 of the Second Edition of the Apraxia Battery for Adults; (Dabul, 2000) was 1.2 [standard deviation (SD) = 0.6; range = 1–3] where seven patients who did not complete this task were assessed using a picture description (picnic) and the story-telling task. Mean age and education were 61 (SD = 14; range = 20–85) and 14 (SD = 4; range = 6–33) years, respectively. To minimize effects of brain-behaviour reorganization, we completed behavioural testing within an average of 4 days after stroke onset (SD = 3; range = 1–13 days).

The control group consisted of 13 non-brain damaged participants (three male, 11 right-handed) matched in age and education with the patient group ($|t|$'s < 1.44, P-values > 0.16). Mean age and education were 55 (SD = 14, range = 37–78) and 16 (SD = 3; range = 12–22) years, respectively.

Story-telling assessment

Patients viewed a picture book of *Cinderella* with printed text occluded at bedside for as long as they wished (Saffran *et al.*, 1989; Zingeser and Berndt, 1990; Rochon *et al.*, 2000; MacWhinney *et al.*, 2011; Rogalski *et al.*, 2011; Thompson *et al.*, 2012; Martin and Schnur, 2019). Patients then closed the book and told the story in their own words. The experimenter encouraged participants to speak more if output was limited. Responses were recorded by a nearby digital device.

Quantitative production analysis

Spontaneous connected speech narratives were transcribed and scored by two research assistants according to the procedures in the quantitative production analysis (QPA) training manual (Berdnt *et al.*, 2000). We chose the QPA because it uses an objective rating approach to comprehensively quantify production deficits by analysing multiple measures at both structural and morphological levels of connected speech to identify differences across patients (Saffran *et al.*, 1989). As shown in a recent study, raters achieved high reliability for transcription and scoring (Martin and Schnur, 2019). We calculated 13 QPA measures (Table 1) and converted patients' QPA scores to z-scores to reflect the degree of deficit using the distribution of control subject QPA performance.

Principal component analysis

To identify independent coherent subsets of the 13 quantitative production analysis measures of spontaneous connected speech, we used PCA. Components with an eigenvalue exceeding one

Table 1 Summarized definitions of the quantitative production analysis measures following Saffran *et al.* (1989)

Dependent measures	Formula	Annotation
Speech measures		
Words per minute	Number of words produced / length of speech sample (min)	All intelligible words produced
Narrative measures		
Narrative words	Number of words directly contributing to narrative	Excluding direct responses to or repetition of examiner's speech, uninterpretable neologisms, consistently used stereotypes, utterances repaired, markers of direct discourse and conjunctions joining complete sentences
Morphological measures		
% Closed-class words produced	Number of closed-class words / number of narrative words	Closed-class words: all words excluding nouns, verbs, adjectives and -ly adverbs
% Pronouns produced	Number of pronouns / (number of nouns + pronouns)	Pronouns: personal, reflexive, and indefinite pronouns
% Verbs produced	Number of verbs / (number of nouns + verbs)	Verbs: all verb forms (e.g. infinitive, gerundive and copula)
Determiner index	Number of nouns requiring determiners, with determiners / number of nouns requiring determiners	Nouns requiring determiners (excluding proper nouns and plurals when in contexts not requiring determiners)
Auxiliary index	(Auxiliary score / number of matrix verbs) – 1	Matrix verb: the main verb in a sentence auxiliary score: each matrix verb itself and auxiliary element of the matrix verb was assigned 1 point Auxiliaries include modal, tense markers and inflections on main verb
Structural measures		
Mean utterance length	Number of words in utterances / number of utterances	Utterances defined by syntactic and prosodic boundaries
% Words in sentences	Number of words in sentences / number of narrative words	Sentence defined as an utterance including a subject/predicate structure: noun/pronoun + main verb or noun/pronoun + copula + adjective/prepositional phrase
Mean sentence length	Number of words in sentences / number of sentences	
% Well-formed sentences	Number of syntactically well-formed sentences / Number of sentences	Sentential ill-formedness could include omission of obligatory arguments, deleted elements, agreement errors or other syntactic anomalies
Embedding index	Number of embeddings / number of sentences	Embeddings: number of embedded clauses associated with a main clause
Elaboration index	Subject noun phrase elaboration + verb phrase elaboration	Phrase elaboration = (number of open class words and pronouns in a phrase / number of phrases) – 1

were extracted and rotated with the varimax method. To maximize power to detect significant effects for the PCA, we analysed language performance for all subjects who fit selection criteria and were able to complete the connected speech task ($n = 65$) independent of lesion demarcation constraints. We used varimax rotation because it only allows a small number of high variable loadings on each factor, yielding clearer interpretations (Kaiser, 1958). We calculated patient component scores for the lesion behaviour mapping analysis.

Image acquisition

We acquired diffusion weighted, apparent diffusion coefficient, and high-resolution structural scans (T_1 and T_2 FLAIR) along the axial direction with 4–5 mm slice thickness, and in cases where MRI was contraindicated, CT scans ($n = 5$) as part of the clinical protocols for admitted stroke cases. The voxel sizes of diffusion-weighted and structural images were $1 \times 1 \times 4.5$ mm, and $0.5 \times 0.5 \times 4.5$ mm, respectively.

Lesion tracing

To demarcate lesions, we first registered the diffusion weighted images with the high-resolution structural images (T_1 or T_2) using AFNI (https://afni.nimh.nih.gov/). Lesions were demarcated directly on the diffusion weighted images (by C.H.), using ITK-snap (http://www.itksnap.org/pmwiki/pmwiki.php) with reference to apparent diffusion coefficient and T_2 FLAIR images. Next, we normalized the individual structural images to the Colin-27 template in Montreal Neurological Institute (MNI) space using ANTS registration (http://stnava.github.io/ANTs/; Avants *et al.*, 2008). We used the corresponding affine parameter and diffeomorphic maps to warp individual masks to the MNI space (Holmes *et al.*, 1998; Avants *et al.*, 2006). For patients with CT images but no MRI, lesions which were clearly visible were directly demarcated onto the Colin 27 template based on the CT image.

Multivariate lesion behaviour mapping analysis

To investigate lesion behaviour relationships, we applied support vector regression (libsvm 3; https://www.csie.ntu.edu.tw/~cjlin/libsvm/) implemented with MATLAB 2018b (https://www.mathworks.com/products/matlab.html). For lesion behaviour mapping, 11 patients for whom lesions were difficult to identify were excluded [no clearly identifiable lesion ($n = 7$); subarachnoid haemorrhage ($n = 3$); and missing diffusion-weighted imaging

sequence ($n = 1$)]. Two patients were excluded because one or more component scores were identified as extreme outliers (beyond 3 SD from the patient average). Therefore, 52 patients were included in the lesion behaviour mapping analysis (three haemorrhagic; two with CT scans; eight left handers). Only voxels with a lesion ratio exceeding 5% (at least three subjects) were included (Sperber and Karnath, 2017). Regarding a potential influence of lesion size, we residualized component scores by considering the effect of lesion size (*cf.* LESYMAP; https://github.com/dorianps/LESYMAP). Then, we normalized the residual scores into an interval of [0, 1] to keep the same scale with the binary lesion pattern [normalized score = (component score − minimal value)/(maximal value − minimal value)]. We selected a non-linear radial basis function kernel to build the model (Zhang *et al.*, 2014; Mirman *et al.*, 2015*b*) because it provided a better data fit in comparison to linear-models. Because gamma and c parameters could affect the model, to determine the optimal parameter pairs, we carried out a grid search on c (10^{-2}–10^{9}) and gamma (10^{-9}–10^{3}) for model selection (the same range as scikit-learn: https://scikit-learn.org/stable/index.html). Specifically, for each parameter pair, 5-fold cross-validation was used to examine the model's prediction accuracy, i.e. data were split into five folds, and each time, we used four folds to train the model and predicted the remaining one. We averaged the mean squared errors between testing and predicted scores of the five folds to reflect the prediction accuracy. To assess model statistical inference, we further generated 1000 random models by permuting behavioural scores and compared the prediction accuracy from the original data to those from the random data. The *P*-value was the probability that the random models had a lower mean squared error than the original model. By comparing *P*-values for all the parameter pairs, we determined the optimal pair of c and gamma and their significance ($P < 0.05$). We then explored the behaviour-related lesion location. Following the approach of Zhang *et al.* (2014), the parametric values of the non-linear model were projected back to the original brain space, reflecting the statistical importance of the voxels to the model (i.e. the beta map). A similar permutation test for each voxel was further derived by shuffling behavioural scores 1000 times with the optimal model parameters to calculate random beta values associated with each voxel. As a result, the *P*-value of each voxel was generated by comparing random beta values with the original and thresholded at $P < 0.05$. Note that we did not apply multiple comparison correction because this is still a controversial issue (*cf.* Zhang *et al.*, 2014; Mirman *et al.*, 2015*b*). We reported significant voxel locations based on the Human Brain Connectome Atlas using a cluster threshold of >100 voxels (Fan *et al.*, 2016).

Data availability

Anonymized data that support the study findings will be available from the corresponding author upon reasonable request.

Results

Impairments of spontaneous connected speech

We calculated 13 aspects of spontaneous connected speech using quantitative production analysis (Saffran *et al.*, 1989; Rochon *et al.*, 2000; Gordon, 2006). Patients demonstrated wide variability and decreased performance in comparison to controls across most measures (Fig. 1). Because years age and education between patient and control groups were not significantly different (Welch $|t|$'s < 1.24, P's > 0.16), did not have unequal variance (Levine's test F's < 1.43, P's > 0.24) and did not correlate with patient performance on any QPA variable ($|r|$'s < 0.19; P's > 0.14; age was marginally correlated with words produced per minute, $r = -0.24$, $P = 0.053$) we did not control for these variables when comparing patient versus control connected speech performance. Patient performance was 1.5 SD from controls on an average of 3.9 of the 13 measures (30%), with wide variability across individuals (SD: 20%; range 0–85%). The degree of speech deficit as measured by the quantitative production analysis variables was similar for the left- and right-handed patients, i.e. left-handed patients ($n = 11$) demonstrated connected speech deficits (at least 1.5 SD below controls) on an average of four quantitative production analysis variables (range 1–11) which was similar to right-handed patients' performance (average = four variables, range = 0–10). As a group in comparison to controls, patients produced significantly fewer narrative words, required determiners, embeddings, well-formed sentences, proportionally more pronouns than nouns and an overall slower speech rate ($|t|$ values $>$ 3.20, P-values < 0.002, Bonferroni corrected). At an individual level, 27 of 65 patients (42%) scored more than 1.5 SD from control performance on these measures (SD: 18%; range 14–66%). Performance on other measures in comparison to controls was significantly impaired without correction for multiple comparisons. Specifically, patients produced proportionally more verbs compared to nouns, fewer elaborations, and shorter sentence and utterance lengths than controls ($|t|$ values: 2.61–2.64, P-values: 0.01–0.03). At an individual level, on average 19 of 65 patients (29%) scored more than 1.5 SD from control performance on these measures (SD: 10%; range 19–42%). There were no significant group differences in degree of auxiliary use, the proportion of closed-class words produced, and the number of words produced in sentences ($|t|$ values < 0.60, P-values > 0.55). However, at an individual level, on average 9 of 65 patients (14%) scored more than 1.5 SD from control performance on these measures (SD: 6%; range 11–22%). Therefore, as a whole, the patients were significantly impaired in comparison to controls on multiple aspects of spontaneous connected speech, with the sparing of few abilities.

Principal components of connected speech production

The quantitative production analysis measures of patients' impairments during spontaneous connected speech production were appropriate for PCA, as indicated by Kaiser-Meyer-Olkin (0.63) and Bartlett's tests ($P < 0.001$; Pechenizkiy *et al.*, 2004). Using PCA, we extracted four connected speech components that accounted for 71% of the

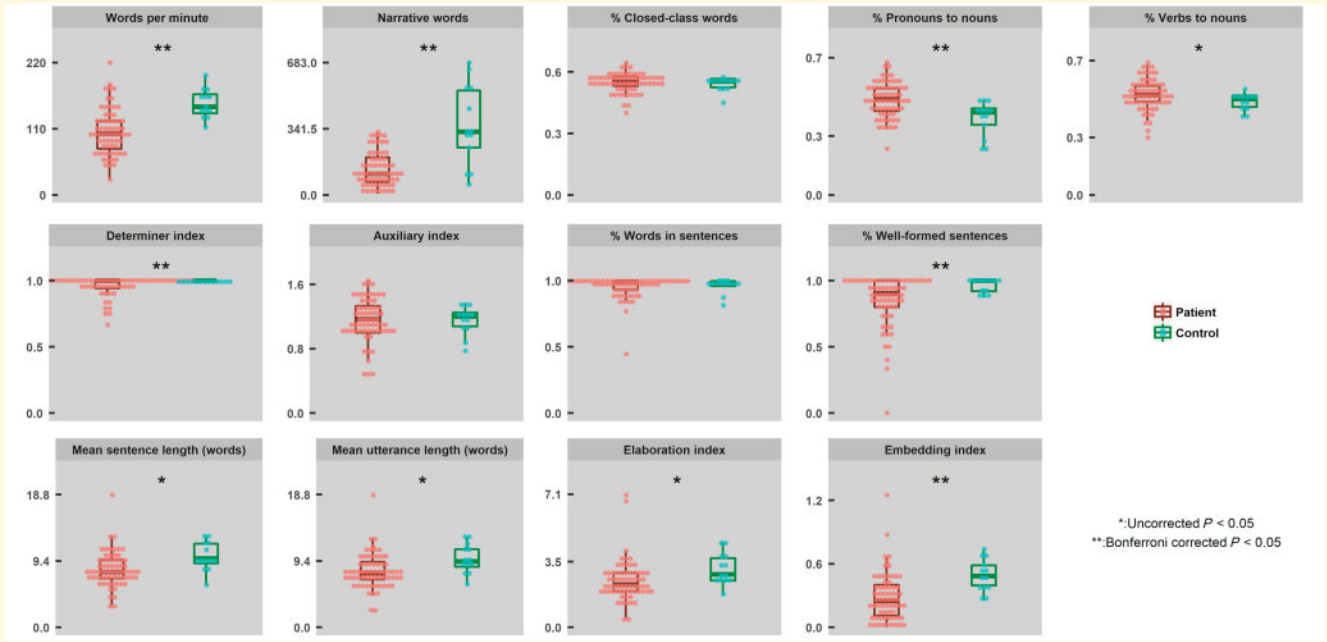

Figure 1 Individual performance and group distribution statistics for the quantitative production analysis measures of spontaneous connected speech. Individual performance is depicted by circles: red = patients; and green = controls. Group distribution statistics are provided as box and whisker plots depicting median, interquartile intervals, minimums, and maximums. */**Significant group differences as determined by Welch's unequal variances *t*-tests (Welch, 1947).

total variance, each component accounting for 30%, 16%, 15% and 10% of variance, respectively (Fig. 2). The first component had high loadings on the mean sentence length, mean utterance length, sentence elaboration index, embedding index and the number of narrative words produced (loadings > 0.6). We interpreted the first component as capturing the ability to produce words in increasingly complex combinations, which requires individuals to arrange words based on appropriate thematic and/or syntactic roles among other constraints. Hence, we call this component 'structural complexity'. The second component had high loadings on the proportion of pronouns, proportion of verbs and proportion of closed-class words produced (loadings > 0.7). To note, this component is bidirectional. Negative values indicate difficulties in producing words with grammatical function, i.e. closed-class words, pronouns and verbs in comparison to primarily nouns. Positive values indicate the reverse. Thus, we interpret the second component as capturing 'lexical selection' abilities for different kinds of words (e.g. from nouns to closed-class words and verbs), depending on the individual subject's component valence. The third component had high loadings on the proportion of well-formed sentences produced, proportion of words produced in sentences and the production of required determiners (loadings > 0.7). We refer to the third component as 'syntactic accuracy' as it captured the ability to generate accurate syntactic structure. The fourth component had high loading on a single variable, the number of words produced per minute (loading = 0.9). This component captured the fluency of overall spontaneous connected speech production. Hence, we referred to this component as 'production fluency'. In summary, we characterized via PCA the patients' multiple impairments of spontaneous connected speech production across four dimensions: structural complexity, lexical selection, syntactic accuracy and production fluency. The PCA results differ in some respects in comparison to the PCA Rochon *et al.* (2000) conducted likely because the current PCA involved different variables and more than twice as many subjects who were unselected for aphasia diagnosis during the acute (versus chronic) stage of stroke.

Lesion distribution

We examined the distribution of voxels damaged in at least 5% (n = 3) of the patient cohort. Figure 3 displays the lesion distribution across subjects. The highest degree of lesion overlap was in the left basal ganglia (n = 11; 21% of the patient group; peak coordinates: –20, –14, 22). Other sufficiently lesioned voxels for lesion behaviour mapping were distributed across the left middle and inferior frontal gyri, pre- and postcentral gyri, inferior and superior parietal lobes, posterior middle temporal gyrus, superior temporal sulcus, lateral occipital lobe, insula and thalamus.

Acute lesions were on average smaller (our data: 12 051 mm^3) than typically seen in chronic stroke (reported average chronic lesion sizes vary between 21 280 and 34 176 mm^3; *cf.* Corbetta *et al.*, 2015; Liew *et al.*, 2018), providing three distinct advantages for lesion behaviour mapping. First,

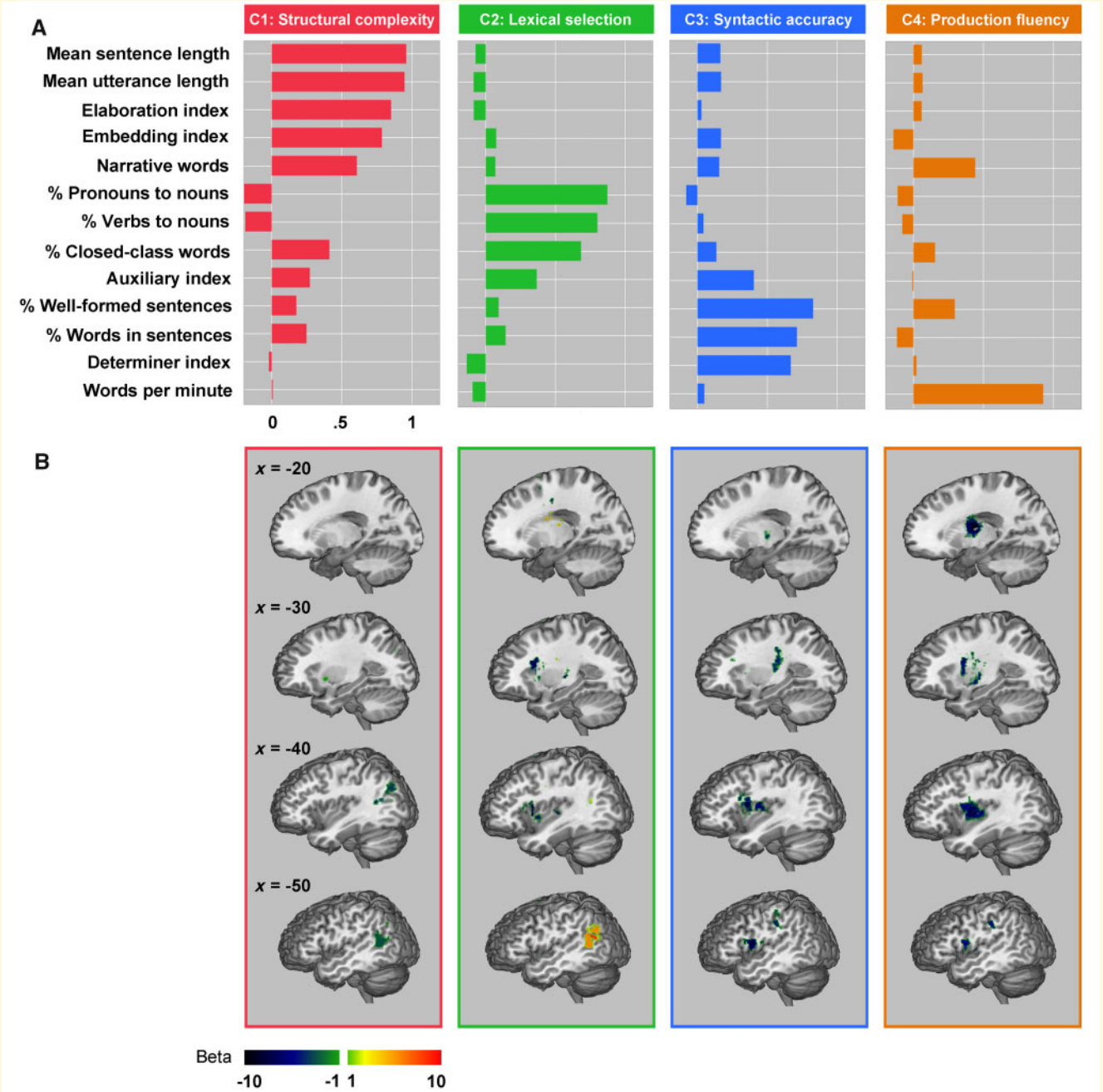

Figure 2 Principal components and multivariate lesion behaviour mapping results. (**A**) Quantitative production analysis measure component loadings C1–C4. (**B**) Significantly associated multivariate lesion behaviour mapping beta maps.

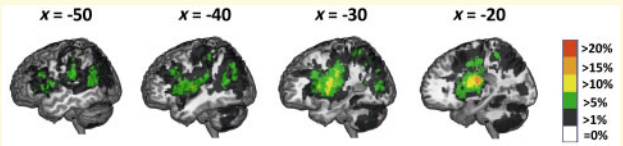

Figure 3 Lesion overlap across 52 patients. We conducted the lesion behaviour mapping on voxels with damage in >5% of the subject sample.

lesion volume was not significantly correlated with proportion damage to any brain connectome region (Bonferroni correction; average $r = 0.23$, range = –0.07 to 0.40; note that we controlled for lesion volume in the lesion behaviour mapping). This allowed us to examine the contribution of damage to function independent of overall stroke severity as measured by lesion volume. Second, regions farther apart from each other (Fig. 4A, lighter green to white) were less likely to be damaged in the same individuals (e.g. inferior

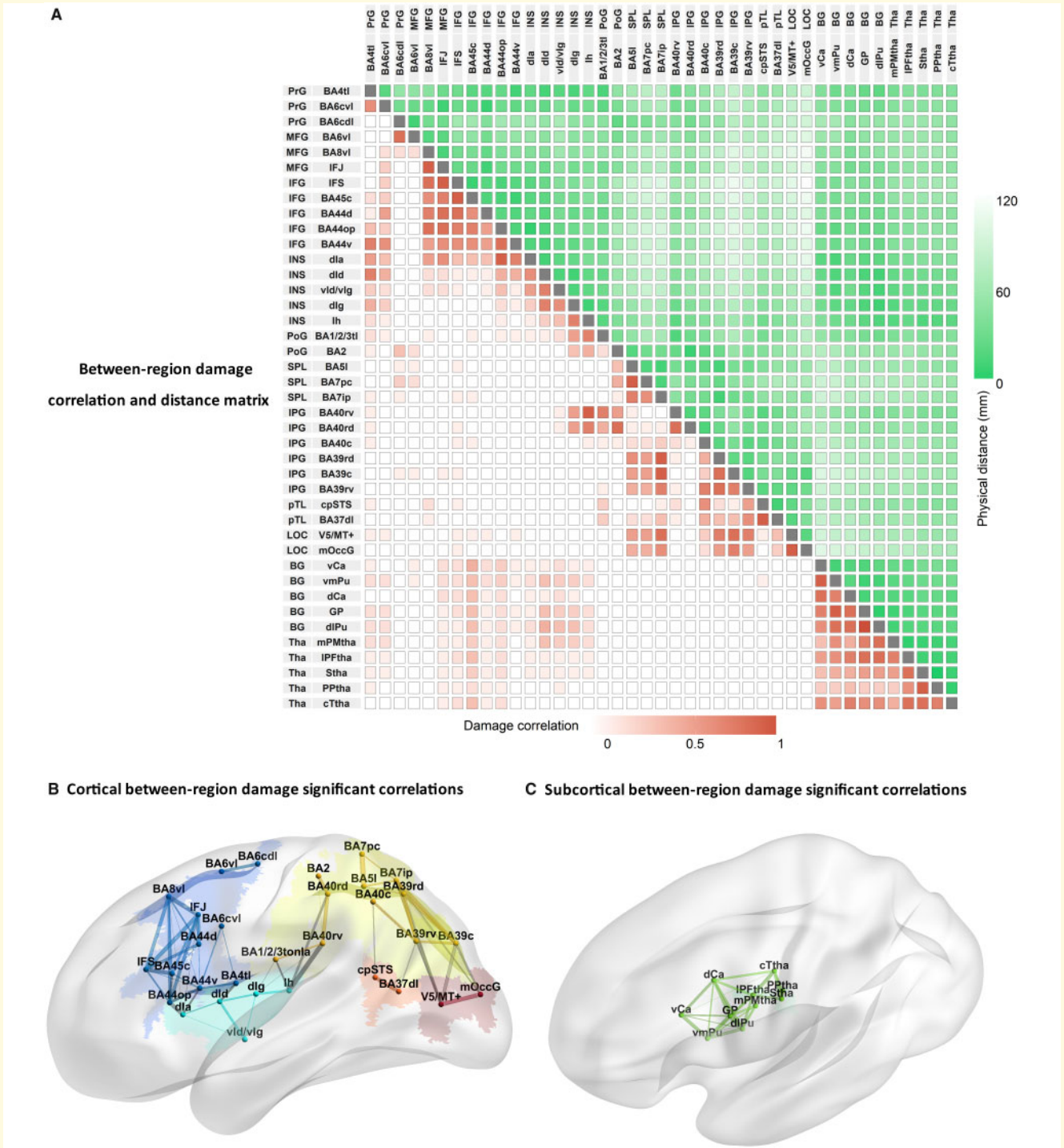

Figure 4 Proportion damage correlations between 41 brainconnectome regions (Fan *et al.*, 2016). (**A**) Matrix of between region damage correlations and distances. Label colours depict different lobules: blue = frontal lobe; cyan = insula; yellow: = parietal lobe; orange = temporal lobe; red = occipital lobe; green = thalamus. Increasing colour intensity reflects either increasing correlations of proportion damage between regions (in red) or decreasing distance between regions (in green). Depictions of correlations between proportion damage across cortical (**B**) and subcortical regions (**C**). Line thickness reflects the magnitude of correlation coefficients (where *r*'s > 0.53; $P < 0.05$ Bonferroni correction). Correlations between lobes are shown in grey and within lobule as the same colour as the lobule itself. BA44op = opercular BA44; BA4tl = BA4 (tongue and larynx region); BA7ip = intraparietal BA7; BA7pc = postcentral BA7; BG = basal ganglia; c = caudal; Ca = caudate; d = dorsal; GP = globus pallidus; Ia = dorsal agranular insula; Id = dysgranular insula; IFG = inferior frontal gyrus; IFS = inferior frontal sulcus; Ig = granular insula; Ih = hypergranular insula; INS = insula; IPL = inferior parietal lobe; l = lateral; LOG = lateral occipital lobe; m = medial; MFG = middle frontal gyrus; OccG = middle occipital gyrus; p = posterior; PFtha = pre-frontal thalamus; PoG = postcentral gyrus; PrG = precentral gyrus; Ptha = parietal thalamus; pTL = posterior temporal lobe; Pu = putamen; PMtha = pre-motor thalamus; r = rostral; s = superior; SPL = superior parietal lobe; STS = superior temporal sulcus; Tha = thalamus; IFJ = inferior frontal junction; Ttha = temporal thalamus; v = ventral.

frontal gyrus and inferior parietal lobe), as illustrated by lower correlations in lighter red to white (Fig. 4A) and thinner to no lines between regions (Fig. 4B and C). The reduced cross-regional correlation in proportion damage within individuals in this cohort allowed us to tease apart the independent contribution of damage to behaviour in frontal versus temporal-parietal regions. Third, although most regions are relatively near each other (Fig. 4A, darker green), proportion damage was correlated only between a few adjacent areas, and mostly within a brain connectome region (Fig. 4A, darker red; Fig. 4B and C, thicker lines between and within regions). As a result, we could discriminate between the contribution of damage to behaviour for some adjacent areas (e.g. inferior frontal gyrus and insula).

Neural substrates associated with impairments of spontaneous connected speech production

To explore lesion behaviour relationships across different spontaneous connected speech production impairments, we conducted multivariate lesion behaviour mapping using the support vector regression algorithm. For the lesion behaviour mapping, we used the patients' component scores and lesion masks. All the models were significant via cross-validation (P-values < 0.03). Figure 2 and Table 2 illustrate the clusters with significant β-values in each associated component model (voxel $P < 0.05$, cluster size > 100 voxels).

The production of reduced structural complexity was associated with damage to areas within the left middle temporal gyrus, posterior superior temporal sulcus, and inferior parietal lobe. Impaired lexical selection (reduced production of nouns) was positively associated with damage to areas overlapping with those associated with reduced structural complexity as well as the basal ganglia. In contrast, impaired lexical selection specific to words with grammatical function (closed-class words, pronouns, and verbs) was negatively associated with damage to areas within the left inferior frontal gyrus and insula. Reduced syntactic accuracy was associated with damage to areas within the left inferior frontal gyrus, insula, precentral gyrus, postcentral gyrus and inferior parietal lobe. Impaired production fluency was associated with damage to multiple regions including the left precentral gyrus, postcentral gyrus, posterior superior temporal sulcus, inferior parietal lobe, insula, thalamus, and basal ganglia.

Discussion

Here we asked whether producing words and organizing them into structure during spontaneous connected speech requires neuroanatomically and behaviourally distinct abilities. By quantifying brain damage and deficits in an unselected group of speakers during the acute phase of stroke we provide a detailed specification of the relationship between damage and deficits before brain-behaviour reorganization occurs. Using PCA and multivariate lesion behaviour mapping, our findings indicate that noun retrieval and the production of increasingly complex word combinations require exclusively left temporal-parietal regions while abilities to produce accurate syntactic structure and grammatical words (including verbs) are served primarily by left frontal regions. These results provide evidence from a rare and important patient population to clarify the debate as to whether morpho-syntactic processes are neuranatomically distinct from lexical processes required to produce spontaneous connected speech.

Lexically-driven versus syntactically-driven stages of production

A striking finding with important theoretical implications is the dissociation between damage to brain regions associated with the production of increasingly complex word combinations and damage associated with the production of syntactically accurate connected speech. Producing less complex structure (e.g. fewer words in utterances and phrases, fewer embedded clauses) was associated with damage to posterior temporal cortex [middle temporal gyrus (Brodmann area, BA 37) and superior temporal sulcus] and the angular gyrus of the inferior parietal lobe (rostroventral BA 39)]. In contrast, syntactic accuracy deficits (e.g. omitting required noun determiners, verb arguments, or morphemes) were associated with damage to the left inferior frontal gyrus (dorsal and ventral BA 44; pars opercularis), the insula, precentral gyri (BAs 6 and 4), postcentral gyri (BAs 1–3), and a single locus in the supramarginal gyrus of the inferior parietal lobe (rostrodorsal BA 40).

We observed a similar rostral-caudal dissociation associated with deficits producing different types of words. Producing fewer nouns was associated with damage to overlapping regions associated with generating complex structure including the middle temporal gyrus (BA 37), posterior superior temporal sulcus, and the angular gyrus (BA 39). In contrast, producing fewer words with grammatical functions (e.g. determiners, prepositions, pronouns, and verbs) was associated with damage to regions near but not overlapping with those associated with syntactic accuracy deficits including BA 44 (pars opercularis) within the left inferior frontal gyrus and the insula.

We hypothesize that the dissociation between brain regions necessary for noun retrieval and their combination into increasingly complex structure versus regions necessary for syntactic accuracy and grammatical word production reflect a neuroanatomical dissociation between different stages required to combine words into structure during connected speech. During initial stages of connected speech, a speaker generates a message she wishes to convey and activates the concepts associated with that message (Levelt, 1989; Levelt *et al.*, 1999). As Fromkin (1971) and Garrett (1975, 1980)

Table 2 Regions significantly associated with the component scores of speech production

Coordinates	Beta values	Cluster size	Lobe	Gyrus	Anatomical and cyto-architectonic description
C1 Structural complexity					
–43 –58 10	–4.53	1742	Parietal lobe	IPL	Rostroventral BA 39
–38 –75 15	–3.30	241		IPL	Caudal BA 39
–44 –54 13	–4.53	880	Temporal lobe	pSTS	Caudoposterior STS
–57 –58 8	–3.12	254		MTG	Dorsolateral BA 37
C2 Lexical selection (negative result – words with grammatical functions)					
–57 –58 8	–10	355	Frontal lobe	IFG	Opercular BA 44
–43 –58 10	–9.68	303	Insular lobe	INS	Dorsal dysgranular insula
–38 –75 15	–10	177		INS	Dorsal agranular insula
–44 –54 13	–8.84	120		INS	Hypergranular insula
C2 Lexical selection (positive result – nouns)					
–47 –65 12	8.71	1960	Parietal lobe	IPL	Rostroventral BA 39
–50 –52 29	5.38	127		IPL	Caudal BA 40
–51 –56 6	5.58	723	Temporal lobe	pSTS	Caudoposterior STS
–48 –57 6	5.58	201		MTG	Dorsolateral BA 37
–14 0 13	8.69	338	Subcortical nuclei	BG	Dorsal caudate
C3 Syntactic accuracy					
–46 3 17	–10	495	Frontal lobe	PrG	Caudal ventrolateral BA 6
–51 2 6	–7.85	474		PrG	BA 4 (tongue and larynx region)
–46 7 16	–10	260		IFG	Dorsal BA 44
–49 6 5	–7.85	232		IFG	Ventral BA 44
–41 –5 6	–7.85	622	Insular lobe	INS	Dorsal granular insula
–36 –4 9	–7.85	605		INS	Dorsal dysgranular insula
–56 –32 30	–4.30	471	Parietal lobe	IPL	Rostrodorsal BA 40
–46 –24 35	–4.30	324		PoG	BA 2
–43 –8 13	–5.70	133		PoG	BA 1/2/3 (tongue and larynx region)
C4 Production fluency					
–53 5 5	4.95	402	Frontal lobe	PrG	Caudal ventrolateral BA 6
–40 3 12	7.43	161		PrG	BA 4 (tongue and larynx region)
–36 –8 4	–9.91	1294	Insular lobe	INS	Dorsal granular insula
–35 –4 9	–9.91	891		INS	Dorsal dysgranular insula
–32 –18 9	–7.43	522		INS	Hypergranular insula
–54 –25 28	–4.95	500	Parietal lobe	IPL	Rostrodorsal BA 40
–46 –24 35	–4.95	163		PoG	BA 2
–56 –56 7	–4.95	143	Temporal lobe	STS	Caudoposterior superior STS
–33 –10 1	–7.52	532	Subcortical nuclei	BG	Dorsolateral putamen
–24 –8 7	–7.52	387		BG	Globus pallidus
–19 –11 10	–10	158		Tha	Caudal temporal thalamus
–19 –10 9	–10	106		Tha	Pre-motor thalamus
–18 –5 18	–9.34	106		BG	Dorsal caudate

All regions > 100 voxels.
BA = Brodmann area; BG = basal ganglia; IFG = inferior frontal gyrus; INS = insula; IPL = inferior parietal lobe; MTG = middle temporal gyrus; PoG = postcentral gyrus; PrG = precentral gyrus; pSTS = posterior superior temporal sulcus; Tha = thalamus.

originally formulated based on speech errors and others further developed (Dell, 1986; Levelt, 1989; Bock and Loebell, 1990; Bock and Levelt, 1994; Bock, 1995; Vigliocco and Hartsuiker, 2002; Matchin and Hickok, 2019), subsequent stages orchestrate word selection and the organization of words into larger multiword utterances. During functional encoding, a speaker must select the words associated with the concepts that form part of the intended message, potentially along with associated information like the word's grammatical class (Levelt *et al.*, 1999; *cf.* Caramazza, 1997; Rapp and Goldrick, 2000) and for heads of phrases (nouns and verbs) the number of arguments (Vigliocco and Hartsuiker, 2002; Thompson *et al.*, 2010; Kemmerer, 2015). To convey the event structure of the message, words are associated with each other via assignment of thematic roles (e.g. doer or receiver of an action) specified by the message and/or the lexical heads of phrase (*cf.* Vigliocco and Hartsuiker, 2002; Matchin *et al.*, 2019). Because thematic roles are not synonymous with syntactic functions (e.g. the agent of a message does not necessarily need to be the subject of the sentence, 'The glass slipper was found by the prince'), elements are also assigned syntactic roles (e.g. subject, direct object; Thompson *et al.*, 2007, 2010). We hypothesize that the middle temporal gyrus and posterior

superior temporal sulcus are regions required for successful lexical selection and generation of lexically-driven argument structure (Saffran *et al.*, 1989; Bock and Levelt, 1994; Pallier *et al.*, 2011; Kemmerer, 2015). Interestingly, lexical selection and structural complexity associated regions were largely non-overlapping in the angular gyrus. An intriguing possibility based on comprehension (Wu *et al.*, 2007; Thothathiri *et al.*, 2012; *cf.* Thompson *et al.*, 2010; Mack *et al.*, 2019) and production (den Ouden *et al.*, 2019) in stroke patients and functional MRI evidence in unimpaired speakers (Pallier *et al.*, 2011; Matchin *et al.*, 2019; *cf.* Thompson *et al.*, 2007) is that the angular gyrus uniquely associated with structural complexity may be necessary for a conceptually (message driven), not lexically driven assignment of thematic roles (also see Binder and Desai, 2011). We speculate that these posterior regions are not involved in the assignment of syntactic roles because deficits in verb production (and syntactic accuracy) were not associated with their damage, consistent with proposals by Thompson *et al.* (2007, 2010; *cf.* Matchin *et al.*, 2019). In sum, the overlap between temporal-parietal regions necessary for lexical selection and producing structural complexity suggests that these regions integrate content and structure corresponding to the functional encoding stage of connected speech production.

In contrast, we hypothesize that the regions associated with the production of accurate syntactic structure and grammatical words reflect a subsequent stage of grammatical encoding which positions utterance elements in order (Garrett, 1976; Bock and Levelt, 1994). This positional encoding stage involves the syntactic realization and linear ordering of lexical-semantic structures marked with bound- (e.g. inflectional markers indicating number, tense, and aspect, e.g. -ed in English) and free-standing closed-class morphemes (e.g. determiners, prepositions). Positional encoding may further involve the phonological specification of this linear order via the ordering of phonological segments (Garrett, 1975; Bock, 1987; Berndt *et al.*, 2002; Boeckx *et al.*, 2014) and the generation of prosodic structure (Vigliocco and Hartsuiker, 2002; *cf.* Ferreira, 1993). Although our analysis of connected speech did not quantify deficits at a phonological level of specificity, the postcentral and supramarginal gyri we identified are similar to damaged regions associated with phonological and phonetic errors during picture naming by speakers with chronic stroke (Mirman *et al.*, 2019). Damage to the left inferior frontal gyrus and insula was associated with deficits producing words with grammatical functions and nearby regions were also associated with reduced syntactic accuracy. We speculate that this association may reflect a verb-directed component of the positional assembly of lexical-semantic structures (*cf.* Hagoort, 2013; Takashima *et al.*, 2019). Together these patterns suggest that the regions associated with morpho-syntactic and grammatical word deficits are required for the syntactic realization of linearized lexical-semantic structure and potentially associated phonological processing during the positional stage of grammatical encoding.

Fluency

The last significant finding concerns fluency, the ability to produce connected speech without effort. We identified an extensive frontal-temporal-parietal and subcortical network of brain regions associated with fluency. This is potentially unsurprising as fluency as reflected here by the number of words produced per minute likely depends on multiple processes during language production including message generation, conceptual activation, grammatical encoding, phonological, articulatory and motor processes (*cf.* Nozari and Faroqi-Shah, 2017). These results overlap but are not synonymous with regions associated with fluency via lesion behaviour mapping in chronic stroke (Yourganov *et al.*, 2016; Halai *et al.*, 2017, 2018; Lacey *et al.*, 2017; *cf.* Borovsky *et al.*, 2007) and brain atrophy in patients with primary progressive aphasia (Wilson *et al.*, 2010; Rogalski *et al.*, 2011). Differences in our approach to measuring fluency and the results extend previous work in several ways. First, we examined fluency quantitatively (*cf.* Yourganov *et al.*, 2016) and without contribution of other cognitive measures not reflective of connected speech (Lacey *et al.*, 2017; Halai *et al.*, 2018). Second, the relationship between fluency and associated brain regions was not due to overall lesion volume (Borovsky *et al.*, 2007; Yourganov *et al.*, 2016; Halai *et al.*, 2017, 2018) or diffuse damage (Wilson *et al.*, 2010; Rogalski *et al.*, 2011). Lastly, we examined spontaneous connected speech during story-telling as opposed to picture description (Wilson *et al.*, 2010; Yourganov *et al.*, 2016; Halai *et al.*, 2017, 2018; Lacey *et al.*, 2017; *cf.* Borovsky *et al.*, 2007; Rogalski *et al.*, 2011), where speakers could not rely on visual input to guide message generation and concept identification. As a result, the fluency related brain network we identified is likely a more precise and ecologically relevant reflection of the multiple brain regions necessary for producing fluent connected speech.

Implications

This study demonstrated neuroanatomical and functional divisions within grammatical encoding during connected speech that dovetail with existing neurobiological sentence production evidence. In chronic stroke speakers, lesion behaviour mapping demonstrated that reduced numbers of words produced in sentences during story-telling was associated largely with damage to the left inferior frontal gyrus with smaller foci in the postcentral gyrus and inferior parietal lobe (Mirman *et al.*, 2019). When speakers undergoing awake craniotomies described pictures depicting actions, direct cortical stimulation of the pars opercularis and triangularis of the left inferior frontal gyrus resulted in 50% of patients producing grammatical errors during sentence production (Chang *et al.*, 2018). In speakers diagnosed with primary progressive aphasia, atrophy in primarily frontal regions was related with grammatical deficits during picture description (Wilson *et al.*, 2010). The role of these left

inferior frontal regions is also consistent with regions implicated for patients clinically assessed with classic agrammatic aphasia profiles using standard aphasia batteries (Saffran *et al.*, 1989; den Ouden *et al.*, 2019; *cf.* Rochon *et al.*, 2000).

However, neural localization of syntactic processing during sentence comprehension is more equivocal. On the one hand, that the posterior portion of the left inferior frontal gyrus (BA 44) is associated with deficits to produce appropriate syntactic markers and closed-class elements is consistent with proposed divisions of labour for the left inferior frontal gyrus in sentence comprehension (for reviews of evidence see Hagoort and Indefrey, 2014; Friederici *et al.*, 2017; *cf.* Hickok and Poeppel, 2007). However, deficits related to syntactic processing in sentence comprehension are also associated with left temporal-parietal regions (Rogalski *et al.*, 2011; Thothathiri *et al.*, 2012; Magnusdottir *et al.*, 2013; Caplan *et al.*, 2016; Fridriksson *et al.*, 2018; *cf.* Tyler *et al.*, 2010; Rogalsky *et al.*, 2018) raising the question of the role of anterior regions in syntactic processing (Thothathiri *et al.*, 2012; Magnusdottir *et al.*, 2013; Blank *et al.*, 2016; Wilson *et al.*, 2016; Fedorenko *et al.*, 2018; Rogalsky *et al.*, 2018).

Matchin and Hickok (2019) propose that the inherently different processing requirements for language production and comprehension account for the differences in syntactic processing related neural substrates. In production as well as comprehension, the generation of lexical-semantic structure is required both to understand the relationship between words during comprehension, as well as specify relationships between words to accurately reflect the message to be produced. However, only in production is morpho-syntactic (and phonological) processing necessary to order elements to accurately reflect the words and associated message to be conveyed. Matchin and Hickok advocate Ferreira and colleagues' proposal (Christianson *et al.*, 2001; Ferreira *et al.*, 2002) that understanding who did what to whom during sentence comprehension can often proceed based on discourse-level and semantic information without decoding of morphosyntactic elements, which may explain why anterior regions are not consistently engaged for syntactic processing during sentence comprehension. Our results significantly contribute to this theoretical debate by demonstrating that for spontaneous connected speech production, frontal and temporal-parietal regions are differentially necessary for lexically-driven (functional) versus morpho-syntactic (positional) encoding.

Limitations

Our study had limitations. First, because we examined connected speech deficits during the acute phase of stroke, individuals with minimal language output could not be included. This resulted in an under-representation of individuals with the most severe production impairments and potentially damage to additional regions within the left hemisphere language network (e.g. the anterior temporal lobe). This represents a trade-off between examining behaviour before brain-behaviour reorganization and avoiding lesion volume confounds versus assessing more severe deficits. Second, compared with univariate approaches, multivariate lesion behaviour mapping identifies multiple brain regions associated with function (Zhang *et al.*, 2014; Yourganov *et al.*, 2016). Within these brain region networks, not only the brain regions themselves but also their connections contribute to function. For example, impaired structural and functional connectivity of the arcuate fasciculus, which connects the inferior frontal gyrus and temporal parietal junction relates to multiple aspects of spontaneous speech production deficits in speakers with chronic stroke and primary progressive aphasia, such as informativeness, syntax or fluency (Marchina *et al.*, 2011; Wilson *et al.*, 2011; Yourganov *et al.*, 2018). However, the neuroimaging data available when we assessed behaviour during the acute phase of stroke were not sufficient to perform structural or functional connectivity analyses. Relatedly, we likely underestimated the behavioural contribution of some regions because focal lesions induce decreases in neuronal activation in distant regions (i.e. diaschisis; Carrera and Tononi, 2014). To minimize these limitations in future work, we anticipate investigating the structural connections (using diffusion tensor imaging) and functional connections (using resting state functional MRI) related with spontaneous speech deficits in individuals during the subacute phase of stroke, independent of their ability to produce speech acutely.

Conclusion

This work provides the first evidence in acute stroke of a functional and anatomical dissociation between temporal-parietal and frontal focal brain regions for lexically and syntactically driven processes required for spontaneous connected speech. These results are consistent with predictions from models of syntactic processing based primarily on evidence from language comprehension (Friederici *et al.*, 2017; Hagoort, 2019; Matchin and Hickok, 2019; *cf.* Wilson *et al.*, 2016; Fedorenko *et al.*, 2018; Rogalsky *et al.*, 2018). By including subjects unbiased as to clinical diagnosis of aphasia and lesion location in the acute stage of left hemisphere stroke, lesion sizes were generally smaller and more varied than in chronic stroke studies. As a result, we were able to dissociate lesion behaviour relationships between brain regions while minimizing the confound of overall lesion size and before brain behaviour reorganization occurred. Future directions should explore the mechanisms by which the critical components of connected speech identified here recover after acute stroke and whether other cognitive capacities, such as working memory (Martin and Schnur, 2019), play a role in recovery. The answers to these questions will be important to help with therapeutic interventions whereby improving other cognitive capacities may incidentally, and critically, improve language function thus contributing to improved daily quality of life after stroke.

Acknowledgements

The authors wish to thank Chia-Ming Lei, Danielle Rossi, and Miranda Brenneman for data collection. We thank Eric Johns, Bowie Lin, Hao Yan, and Rachel Zahn for transcription and analysis of narrative speech samples. We thank Lynn Maher for her help in diagnosing apraxia of speech in our subject population. We thank the clinical neurological intensive care unit teams at the University of Texas Health Sciences Center and Memorial Hermann Hospital, The Houston Methodist Hospital, and the Baylor St. Luke's Hospital for their assistance in patient recruitment and neurological assessment. We gratefully acknowledge and thank our research subjects and their caregivers for their willingness to participate in this research. This work was presented at the Academy of Aphasia in Macau, China (2019).

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award number R01DC014976 to the Baylor College of Medicine and an award from the Moody Endowment to Rice University.

Competing interests

The authors report no competing interests.

References

Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008; 12: 26–41.

Avants BB, Schoenemann PT, Gee JC. Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. Med Image Anal 2006; 10: 397–412.

Berndt RS, Haendiges AN, Burton MW, Mitchum CC. Grammatical class and imageability in aphasic word production: their effects are independent. J Neurolinguistics 2002; 15: 353–71.

Berdnt RS, Wayland S, Rochon E, Safffran E, Schwartz M. Quantitative production analysis: a training manual for the analysis of aphasic sentence production. Hove, East Sussex, UK: Psychology Press; 2000.

Binder JR, Desai RH. The neurobiology of semantic memory. Trends Cogn Sci 2011; 15: 527–36.

Blank I, Balewski Z, Mahowald K, Fedorenko E. Syntactic processing is distributed across the language system. Neuroimage 2016; 127: 307–23.

Bock K. An effect of the accessibility of word forms on sentence structures. J Mem Lang 1987; 26: 119–37.

Bock K. Producing agreement. Curr Dir Psychol Sci 1995; 4: 56–61.

Bock K, Levelt W. Language production: grammatical encoding. Handbook of psycholinguistics. San Diego, CA, USA: Academic Press; 1994. p. 945–84.

Bock K, Loebell H. Framing sentences. Cognition 1990; 35: 1–39.

Boeckx C, Martinez-Alvarez A, Leivada E. The functional neuroanatomy of serial order in language. J Neurolinguistics 2014; 32: 1–15.

Borovsky A, Saygin AP, Bates E, Dronkers N. Lesion correlates of conversational speech production deficits. Neuropsychologia 2007; 45: 2525–33.

Brookshire RH, Nicholas LE. Speech sample size and test-retest stability of connected speech measures for adults with aphasia. J Speech Lang Hear Res 1994a; 37: 399–407.

Brookshire RH, Nicholas LE. Test-retest stability of measures of connected speech in aphasia. Clin Aphasiol 1994b; 22: 119–33.

Butler RA, Lambon Ralph MA, Woollams AM. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain 2014; 137: 3248–66.

Caplan D, Michaud J, Hufford R, Makris N. Deficit-lesion correlations in syntactic comprehension in aphasia. Brain Lang 2016; 152: 14–27.

Caramazza A. How many levels of processing are there in lexical access? Cogn Neuropsychol 1997; 14: 177–208.

Carrera E, Tononi G. Diaschisis: past, present, future. Brain 2014; 137: 2408–22.

Chang EF, Kurteff G, Wilson SM. Selective interference with syntactic encoding during sentence production by direct electrocortical stimulation of the inferior frontal gyrus. J Cogn Neurosci 2018; 30: 411–20.

Christianson K, Hollingworth A, Halliwell JF, Ferreira F. Thematic roles assigned along the garden path linger. Cogn Psychol 2001; 42: 368–407.

Corbetta M, Ramsey L, Callejas A, Baldassarre A, Hacker Carl D, Siegel Joshua S, et al. Common behavioral clusters and subcortical anatomy in stroke. Neuron 2015; 85: 927–41.

Dabul B. Apraxia battery for adults, 2nd edn. Austin, TX: PRO-ED Inc; 2000.

Dapretto M, Bookheimer SY. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 1999; 24: 427–32.

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychol Rev 1986; 93: 283–321.

den Ouden D-B, Malyutina S, Basilakos A, Bonilha L, Gleichgerrcht E, Yourganov G, et al. Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. Hum Brain Mapp 2019; 40: 2153–73.

Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 2016; 26: 3508–26.

Fedorenko E, Mineroff Z, Siegelman M, Blank I. Word meanings and sentence structure recruit the same set of fronto-temporal regions during comprehension. bioRxiv 2018 (preprint): 477851. doi: 10.1101/477851.

Fedorenko E, Thompson-Schill SL. Reworking the language network. Trends Cogn Sci 2014; 18: 120–6.

Ferreira F. Creation of prosody during sentence production. Psychol Rev 1993; 100: 233–53.

Ferreira F, Bailey KGD, Ferraro V. Good-enough representations in language comprehension. Curr Dir Psychol Sci 2002; 11: 11–5.

Fridriksson J, den Ouden D-B, Hillis AE, Hickok G, Rorden C, Basilakos A, et al. Anatomy of aphasia revisited. Brain 2018; 141: 848–62.

Friederici AD, Chomsky N, Berwick RC, Moro A, Bolhuis JJ. Language, mind and brain. Nat Hum Behav 2017; 1: 713–22.

Fromkin VA. The non-anomalous nature of anomalous utterances. Language 1971; 47: 27–52.

Garrett M. Levels of processing in sentence production. In: Butterworth B, editor. Language production Vol. 1: Speech and Talk. London: Academic Press; 1980. p. 177–220.

Garrett MF. The analysis of sentence production. In: GH Bower, editor. Psychology of learning and motivation. New York: Academic Press; 1975. p. 133–77.

Garrett MF. Syntactic processes in sentence production. New approaches to language mechanisms 1976; 30: 231–256.

Gordon JK. A quantitative production analysis of picture description. Aphasiology 2006; 20: 188–204.
Hagoort P. MUC (memory, unification, control) and beyond. Front Psychol 2013; 4: 416.
Hagoort P. The neurobiology of language beyond single-word processing. Science 2019; 366: 55–8.
Hagoort P, Indefrey P. The neurobiology of language beyond single words. Annu Rev Neurosci 2014; 37: 347–62.
Halai AD, Woollams AM, Lambon Ralph MA. Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: revealing the unique neural correlates of speech fluency, phonology and semantics. Cortex 2017; 86: 275–89.
Halai AD, Woollams AM, Lambon Ralph MA. Triangulation of language-cognitive impairments, naming errors and their neural bases post-stroke. NeuroImage: Clinical 2018; 17: 465–73.
Hartwigsen G, Saur D. Neuroimaging of stroke recovery from aphasia—insights into plasticity of the human language network. Neuroimage 2019; 190: 14–31.
Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci 2007; 8: 393–402.
Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. Enhancement of MR images using registration for signal averaging. J Comput Assist Tomogr 1998; 22: 324–33.
Kaiser HF. The varimax criterion for analytic rotation in factor analysis. Psychometrika 1958; 23: 187–200.
Karnath H-O, Rennig J, Johannsen L, Rorden C. The anatomy underlying acute versus chronic spatial neglect: a longitudinal study. Brain 2010; 134: 903–12.
Kemmerer D. Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. Psychon Bull Rev 2015; 22: 1068–75.
Kertesz A. Western aphasia battery test manual. San Antonio, TX: Psychological Corp; 1982.
Kimberg DY, Coslett HB, Schwartz MF. Power in voxel-based lesion-symptom mapping. J Cogn Neurosci 2007; 19: 1067–80.
Kristinsson S, Thors H, Yourganov G, Magnusdottir S, Hjaltason H, Stark BC, et al. Brain damage associated with impaired sentence processing in acute aphasia. J Cogn Neurosci 2020; 32: 256–71.
Lacey EH, Skipper-Kallal LM, Xing S, Fama ME, Turkeltaub PE. Mapping common aphasia assessments to underlying cognitive processes and their neural substrates. Neurorehabil Neural Repair 2017; 31: 442–50.
Levelt WJ, Speaking: from intention to articulation. Cambridge, MA: The MIT Press; 1989.
Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behav Brain Sci 1999; 22: 1–38.
Liew S-L, Anglin JM, Banks NW, Sondag M, Ito KL, Kim H, et al. A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. Sci Data 2018; 5: 180011.
Lorca-Puls DL, Gajardo-Vidal A, White J, Seghier ML, Leff AP, Green DW, et al. The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. Neuropsychologia 2018; 115: 101–11.
Mack JE, Mesulam MM, Rogalski EJ, Thompson CK. Verb-argument integration in primary progressive aphasia: real-time argument access and selection. Neuropsychologia 2019; 134: 107192.
MacWhinney B, Fromm D, Forbes M, Holland A. AphasiaBank: methods for studying discourse. Aphasiology 2011; 25: 1286–307.
Magnusdottir S, Fillmore P, den Ouden DB, Hjaltason H, Rorden C, Kjartansson O, et al. Damage to left anterior temporal cortex predicts impairment of complex syntactic processing: a lesion-symptom mapping study. Hum Brain Mapp 2013; 34: 2715–23.
Mah Y-H, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. Brain 2014; 137: 2522–31.
Marchina S, Zhu LL, Norton A, Zipse L, Wan Catherine Y, Schlaug G. Impairment of speech production predicted by lesion load of the left arcuate fasciculus. Stroke 2011; 42: 2251–6.
Marsh EB, Hillis AE. Recovery from aphasia following brain injury: the role of reorganization. Progress in brain research 2006; 157: 143–156.
Martin RC, Schnur TT. Independent contributions of semantic and phonological working memory to spontaneous speech in acute stroke. Cortex 2019; 112: 58–68.
Matchin W, Hickok G. The cortical organization of syntax. Cerebral Cortex 2019; 00: 1–18.
Matchin W, Liao C-H, Gaston P, Lau E. Same words, different structures: an fMRI investigation of argument relations and the angular gyrus. Neuropsychologia 2019; 125: 116–28.
Mirman D, Chen Q, Zhang Y, Wang Z, Faseyitan OK, Coslett HB, et al. Neural organization of spoken language revealed by lesion–symptom mapping. Nat Commun 2015a; 6: 6762.
Mirman D, Kraft AE, Harvey DY, Brecher AR, Schwartz MF. Mapping articulatory and grammatical subcomponents of fluency deficits in post-stroke aphasia. Cogn Affect Behav Neurosci 2019;
Mirman D, Zhang Y, Wang Z, Coslett HB, Schwartz MF. The ins and outs of meaning: behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. Neuropsychologia 2015b; 76: 208–19.
Nardo D, Holland R, Leff AP, Price CJ, Crinion JT. Less is more: neural mechanisms underlying anomia treatment in chronic aphasic patients. Brain 2017; 140: 3039–54.
Nozari N, Faroqi-Shah Y. Investigating the origin of nonfluency in aphasia: a path modeling approach to neuropsychology. Cortex 2017; 95: 119–35.
Ochfeld E, Newhart M, Molitoris J, Leigh R, Cloutman L, Davis C, et al. Ischemia in broca area is associated with broca aphasia more reliably in acute than in chronic stroke. Stroke 2010; 41: 325–30.
Olness GS, Ulatowska HK. Aphasias. In: L Cummings, editor. Research in clinical pragmatics. Cham: Springer International Publishing; 2017. p. 211–42.
Pallier C, Devauchelle A-D, Dehaene S. Cortical representation of the constituent structure of sentences. Proc Natl Acad Sci USA 2011; 108: 2522–7.
Pechenizkiy M, Tsymbal A, Puuronen S. PCA-based feature transformation for classification: issues in medical diagnostics. Proceedings 17th IEEE Symposium on Computer-Based Medical Systems, 25 June 2004; 535–540.
Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann N Y Acad Sci 2010; 1191: 62–88.
Pustina D, Avants B, Faseyitan OK, Medaglia JD, Coslett HB. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. Neuropsychologia 2018; 115: 154–66.
Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. Psychol Rev 2000; 107: 460–99.
Roberts A, Post D. Information content and efficiency in the spoken discourse of individuals with Parkinson's disease. J Speech Lang Hear Res 2018; 61: 2259–74.
Rochon E, Saffran EM, Berndt RS, Schwartz MF. Quantitative analysis of aphasic sentence production: further development and new data. Brain Lang 2000; 72: 193–218.
Rogalski E, Cobia D, Harrison TM, Wieneke C, Thompson CK, Weintraub S, et al. Anatomy of language impairments in primary progressive aphasia. J Neurosci 2011; 31: 3344–50.
Rogalsky C, LaCroix AN, Chen K-H, Anderson SW, Damasio H, Love T, et al. The neurobiology of agrammatic sentence comprehension: a lesion study. J Cogn Neurosci 2018; 30: 234–55.
Saffran EM, Berndt RS, Schwartz MF. The quantitative analysis of agrammatic production: procedure and data. Brain Lang 1989; 37: 440–79.

Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, et al. Dynamics of language reorganization after stroke. Brain 2006; 129: 1371–84.

Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. Proc Natl Acad Sci USA 2009; 106: 322–7.

Shahid H, Sebastian R, Schnur TT, Hanayik T, Wright A, Tippett DC, et al. Important considerations in lesion-symptom mapping: illustrations from studies of word comprehension. Hum Brain Mapp 2017; 38: 2990–3000.

Sperber C, Karnath H-O. Impact of correction factors in human brain lesion-behavior inference. Hum Brain Mapp 2017; 38: 1692–701.

Sperber C, Wiesen D, Karnath H-O. An empirical evaluation of multivariate lesion behaviour mapping using support vector regression. Hum Brain Mapp 2019; 40: 1381–90.

Takashima A, Konopka A, Meyer A, Hagoort P, Weber K. Speaking in the brain: the interaction between words and syntax in producing sentences. bioRxiv 2019 (preprint): 696310. doi: 10.1101/696310.

Thompson CK, Bonakdarpour B, Fix SC, Blumenfeld HK, Parrish TB, Gitelman DR, et al. Neural correlates of verb argument structure processing. J Cogn Neurosci 2007; 19: 1753–67.

Thompson CK, Bonakdarpour B, Fix SF. Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. J Cogn Neurosci 2010; 22: 1993–2011.

Thompson CK, Cho S, Hsu C-J, Wieneke C, Rademaker A, Weitner BB, et al. Dissociations between fluency and agrammatism in primary progressive aphasia. Aphasiology 2012; 26: 20–43.

Thompson CK, den Ouden D-B. Neuroimaging and recovery of language in aphasia. Curr Neurol Neurosci Rep 2008; 8: 475.

Thompson CK, Meltzer-Asscher A, Cho S, Lee J, Wieneke C, Weintraub S, et al. Syntactic and morphosyntactic processing in stroke-induced and primary progressive aphasia. Behav Neurol 2013; 26: 35–54.

Thothathiri M, Kimberg DY, Schwartz MF. The neural basis of reversible sentence comprehension: evidence from voxel-based lesion symptom mapping in aphasia. J Cogn Neurosci 2012; 24: 212–22.

Turkeltaub PE, Coslett HB, Thomas AL, Faseyitan O, Benson J, Norise C, et al. The right hemisphere is not unitary in its role in aphasia recovery. Cortex 2012; 48: 1179–86.

Tyler LK, Wright P, Randall B, Marslen-Wilson WD, Stamatakis EA. Reorganization of syntactic processing following left-hemisphere brain damage: does right-hemisphere activity preserve function? Brain 2010; 133: 3396–408.

Vigliocco G, Hartsuiker RJ. The interplay of meaning, sound, and syntax in sentence production. Psychol Bull 2002; 128: 442–72.

Weiller C, Isensee C, Rijntjes M, Huber W, Müller S, Bier D, et al. Recovery from wernicke's aphasia: a positron emission tomographic study. Ann Neurol 1995; 37: 723–32.

Welch BL. The generalization of 'student's' problem when several different population variances are involved. Biometrika 1947; 34: 28–35.

Wilson SM, DeMarco AT, Henry ML, Gesierich B, Babiak M, Miller BL, et al. Variable disruption of a syntactic processing network in primary progressive aphasia. Brain 2016; 139: 2994–3006.

Wilson SM, Galantucci S, Tartaglia Maria C, Rising K, Patterson Dianne K, Henry ML, et al. Syntactic processing depends on dorsal language tracts. Neuron 2011; 72: 397–403.

Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, et al. Connected speech production in three variants of primary progressive aphasia. Brain 2010; 133: 2069–88.

Wilson SM, Saygın AP. Grammaticality judgment in aphasia: deficits are not specific to syntactic structures, aphasic syndromes, or lesion sites. J Cogn Neurosci 2004; 16: 238–52.

Wu DH, Waller S, Chatterjee A. The functional neuroanatomy of thematic role and locative relational knowledge. J Cogn Neurosci 2007; 19: 1542–55.

Yourganov G, Fridriksson J, Rorden C, Gleichgerrcht E, Bonilha L. Multivariate connectome-based symptom mapping in post-stroke patients: networks supporting language and speech. J Neurosci 2016; 36: 6668–79.

Yourganov G, Fridriksson J, Stark B, Rorden C. Removal of artifacts from resting-state fMRI data in stroke. Neuroimage Clin 2018; 17: 297–305.

Zhang Y, Kimberg DY, Coslett HB, Schwartz MF, Wang Z. Multivariate lesion-symptom mapping using support vector regression. Hum Brain Mapp 2014; 35: 5861–76.

Zingeser LB, Berndt RS. Retrieval of nouns and verbs in agrammatism and anomia. Brain and Language 1990; 39: 14–32.