Distinct Patterns of Syntactic Errors Doubly Dissociate in Chronic Poststroke Aphasia

Jeremy Yeaton[1,2*], Danielle Fahey[3*], Michelle Gomez[4,5], Brielle C. Stark[6], Julius Fridriksson[7], Dirk B. den Ouden[8], Gregory Hickok[1], and William Matchin[7]

Abstract

■ The lesion correlates of syntactic deficits in aphasia remain poorly understood. Previous studies have suggested that distinct error types in expressive syntax may be associated with damage to different brain regions, but the specific lesion correlates of these errors have not been fully delineated. To address this gap, we analyzed spontaneous speech samples from individuals with chronic aphasia, categorizing errors into hierarchical (paragrammatic-like) and linearization (agrammatic-like) errors. Lesion-symptom mapping was conducted to identify brain regions associated with these error types. Lesion clusters in the medial superior temporal sulcus and inferior parietal lobe were associated with hierarchical errors, whereas frontal lesions, including those in the inferior and middle frontal lobe, were associated with linearization errors. These results provide support for a two-stage model of syntactic encoding, with distinct neural correlates for hierarchical and linearization processes. Our findings suggest that distinct brain regions contribute to different stages of syntactic encoding. Hierarchical processing seems to be supported by posterior temporal–parietal regions, whereas linearization seems to be supported by frontal regions. These results suggest that the diagnosis and treatment of aphasia could take into account the existence of distinct syntactic production deficits resulting from different patterns of brain damage. ■

INTRODUCTION

Historically, the study of grammatical deficits in aphasia has focused overwhelmingly on the syndrome of *agrammatism*. Expressive agrammatism is characterized by syntactic reduction and simplification in speech output, typically with omitted verbs and functional elements, and difficulty producing noncanonical sentence structures, resulting in so-called telegraphic speech (e.g., *and he meet the prince at party* [participant M2006][1]; Kean, 1977; Goodglass, 1968; Jakobson, 1956). Lesion-deficit correlation studies in both stroke-based and primary progressive aphasia have found associations between expressive agrammatism and damage or degeneration to frontal cortex, including Broca's area (Matchin et al., 2020; Den Ouden et al., 2019; Whitwell et al., 2017; Thompson, Den Ouden, Bonakdarpour, Garibaldi, & Parrish, 2010; Wilson et al., 2010). By contrast, little is known about the lesion basis of paragrammatism—another syndrome of grammatical deficits in aphasia. Paragrammatism is characterized by transpositions, substitutions, and insertions often resulting in ungrammatical sentences, sometimes called "sentence monsters" (e.g., *oh actually before that they her Cindreller he had some nice and she saw that and she threw out take the necklace and all this stuff to not good* [participant M2226]; Matchin et al., 2020; Kleist, 1914). Although agrammatism has traditionally been associated with nonfluent aphasic syndromes (e.g., Broca's) and paragrammatism with fluent aphasic syndromes (e.g., Wernicke's, conduction), our focus here is the syntactic-level symptoms (error patterns), independent of clinical aphasia categories.

Although few in-depth studies of paragrammatic speech errors have been performed (Butterworth and Howard [1987] and Goodglass, Christiansen, and Gallagher [1993] being notable exceptions), some authors have argued that agrammatism and paragrammatism result from the same underlying syntactic deficit (Kolk & Heeschen, 1992; Heeschen, 1985). However, recently it has been proposed that these two syndromes result from damage to distinct brain systems (Matchin & Hickok, 2020), mapping onto distinct deficits in broad stages of syntactic encoding in prevailing psycholinguistic theories of sentence production (Yeaton, Fahey, Khosshal Mollasaraei, Krauska, & Matchin, 2025). Under these theories, the abstract syntactic structure of a sentence is generated in two phases: an initial abstract hierarchical phase, which constructs a relational hierarchy that encodes the meaning of an utterance (sometimes termed "functional" processing), and a subsequent linearization

[1]University of California, Irvine, [2]University of California, Davis, [3]University of Alabama, [4]University of California, San Diego, [5]Veterans Affairs San Diego Healthcare System, La Jolla, [6]Indiana University Bloomington, [7]University of South Carolina, [8]Chapman University
*Co-first authors.

https://doi.org/10.1162/JOCN.a.2684

and phonological encoding phase (sometimes termed "positional" processing; Krauska & Lau, 2023; Matchin & Hickok, 2020; Ferreira & Slevc, 2007; Bock & Levelt, 1994). During the first phase, the brain retrieves lexical and structural information that encodes the individual entities, semantic roles, and semantic relations of the message and joins these items into a hierarchically organized structure to encode the relationships between words and phrases. During the second phase, the elements of the tree are then linearized at the morphosyntactic and phonological levels and transformed into a linear sequence suitable for production by the speech motor system. These two phases have different names under different theoretical accounts, but we will refer to them as the *hierarchical* and *linearization* phases, respectively (Matchin & Hickok, 2020). Roughly speaking, we hypothesize that paragrammatic deficits are typically caused by breakdowns in the initial hierarchical stage, and agrammatic deficits are typically caused by breakdowns in the second, linearization stage.

While the two-stage hierarchical/linearization framework guides our predictions and error categorization, we note that other theoretical perspectives are not necessarily incompatible with the present approach. Lexicalist accounts, for instance, treat syntactic and lexical operations as deeply intertwined (see Krauska & Lau, 2023, for a review); under such accounts, what we term "hierarchical errors" may partly reflect failures at the lexical–syntactic interface rather than purely structural computation. Predictive coding accounts, on the other hand, treat syntactic structure as emerging from hierarchical predictions at multiple levels of abstraction (Schrimpf et al., 2021), and these would likely map onto our hierarchical stage. Our error coding scheme is deliberately agnostic to the specific computational mechanism underlying each stage, focusing instead on observable surface consequences.

Consistent with these hypotheses, Matchin and colleagues (2020) used a perceptual approach to identify aphasic patients as predominantly agrammatic or paragrammatic in character and found that the lesion correlates of agrammatism and paragrammatism doubly dissociate: Large frontal lesions were associated with agrammatism, and posterior temporal–parietal lesions were associated with paragrammatism (see also Casilio et al. [2025] for converging evidence). However, this holistic, perceptual approach could have conflated nonsyntactic deficits (such as prosodic, lexical, and phonological issues) with syntactic ones. In addition, this approach focused on the traditional clinical impressions of deficits rather than more clearly categorizing patients or utterances according to underlying functional mechanisms encoded by the brain, which likely do not map cleanly onto those clinical syndromes.

The present study improves upon previous approaches by presenting lesion-mapping analyses of grammatical deficits in aphasia based on an utterance-level coding scheme motivated by a psycholinguistic model of sentence production. This analysis was performed on an utterance-by-utterance basis—akin to what is common practice in scoring picture naming errors—rather than the global impression ratings used previously (Fahey et al., 2025). We predicted, consistent with previous results, a double dissociation in the lesion profiles of grammatical production deficits in people with aphasia: Linearization-level errors (those which generally typify agrammatism) should be associated with frontal lobe lesions, whereas hierarchical-level errors (those which generally typify paragrammatism) should be associated with lesions in the posterior temporal lobe.

METHODS

Participants

Participants were drawn from a database of individuals with chronic, poststroke aphasia who have completed testing for various studies conducted at the University of South Carolina and the Medical University of South Carolina over the last 15 years. All were native speakers of English, had suffered an ischemic stroke to the left hemisphere at least 6 months before the study, and presented with language difficulties (all participants had presented with aphasia in the acute phase following their stroke, which formed the basis of their enrollment in the respective studies, and most participants were classified as aphasic according to the WAB-R [Kertesz, 2007] during tests in the chronic phase, although some scored outside of the aphasic range by the time of examination). From this database, 79 participants were selected (mean age = 59.7 years old [range: 29–80 years old], 31 women). Participants were diagnosed with the following aphasia types: 21 with anomia, 25 with Broca's aphasia, 18 with conduction aphasia, 2 with global aphasia, 1 with transcortical motor aphasia, 1 with transcortical sensory aphasia, and 5 with Wernicke's aphasia. Six participants were classified as not aphasic according to the WAB-R. Fifty-one of the included participants were selected because of their inclusion in Matchin and colleagues (2020). Selection for that study was based on the amount of intelligible speech output recorded during their retelling of the story of Cinderella; see below). The remaining 28 participants were selected from the database based on the availability of imaging data and a transcribed Cinderella discourse sample with at least 10^2 intelligible, nonfiller utterances (Fahey et al., 2025). All procedures were approved by the internal review board at the University of South Carolina (Pro00053559), informed consent was obtained, and data were de-identified.

Elicitation Protocol

To elicit speech, we followed the story-retelling protocol from AphasiaBank (MacWhinney, Fromm, Forbes, & Holland, 2011). Participants first reviewed a picture book of the Cinderella story (text omitted). The book

was then removed, and they were asked to retell the Cinderella story in their own words. Video recordings were made, lasting from a few seconds to more than 5 min. Participants were instructed to use content from the book and their own recollection. Videos were transcribed by speech-language pathology master's students following CHAT-CLAN conventions (MacWhinney et al., 2011).

Coding Scheme

Utterances that contained linguistic content (i.e., not filled pauses) were coded according to the type of errors they contained (Fahey et al., 2025). If an utterance was not a well-formed sentence or phrase, the syntactic error was determined by estimating the least changes that would have rectified the objective ungrammaticality. Errors in hierarchical processing ("Hier") were typically identifiable by the overuse or misuse over hierarchical structuring elements and included subcategorization violations, fusion errors, exchange errors, and unfilled nodes or branches, whereas errors in linear processing ("Lin") were usually identifiable by reduced structural elements and included word order violations, functional morpheme errors (e.g., inflection errors), or telegraphic speech (evidence of defective frame building). Single omission errors ("O"), utterances in which a single functional morpheme omission caused ungrammaticality, were included as a separate category because of low confidence in their categorization as either hierarchical or linear errors. For example, in the utterance "*the boy walk*," the third-person present tense *–s* is absent from the verb. This could be a linearization error whereby the morpheme *–s* was omitted at the linearization stage, with the hierarchical relation required for establishing agreement between the subject and verb phrase intact. On the other hand, the omission may have arisen due to correct linearization of degraded structure at the hierarchical level (e.g., the verb agreement relations were not well formed). This general uncertainty in many cases motivated the omission "O" category to limit noise in our categorization. Finally, utterances could be identified as having both hierarchical and linearization errors (H + L), exhibiting apparent deficits in both levels. For the list of syntactic utterance codes as well as examples, see Table 1 (see Supplement 1 for a more extensive list of annotated examples). All utterances received two additional codes: presence or absence of each phonological and semantic errors/paraphasias (regardless of the presence of syntactic errors).

Three expert raters (D. F., J. Y., and M. G.)—each blind to participants' aphasia type, psychological assessment scores, prior ratings from Matchin and colleagues (2020), and lesions—viewed the transcripts and coded each utterance where the target could reasonably be inferred according to the schema outlined above. Consensus was reached between the initial raters for each utterance. Two additional expert raters (W. M. and B. S.) rated a subset of samples. The subset of samples was the same for each additional rater and was selected to represent a variety of aphasia types and evaluations. These expert raters were also blinded to participants' aphasia types, psychological assessment scores, lesions, and ratings in Matchin and colleagues. All raters were experienced language scientists with a mix of backgrounds in aphasia research, psycholinguistics, and theoretical linguistics.

To determine rater agreement between the original consensus codes and the codes provided by the additional raters, Cohen's kappa was used. Overall, secondary raters agreed with the consensus codes on grammaticality in

Table 1. System of Coding Syntactic Errors and Example Utterances

Utterance Codes	*Example*
Hier: deficits in hierarchical and grammatical function processing; subcategorization violation, fusion errors, exchange errors, unfilled nodes or branches (e.g., lexical items or phrases required by syntactic relations), ungrammatically inserted branches	*"She was thinking about that things that she was for everything she was goings something"*
Lin: deficits in linearization/positional processing; word order violations, functional morpheme errors (e.g., inflection errors), telegraphic speech (evidence of defective frame building)	*"And he meet the prince at party"*
O: ungrammatical utterance (a/p indeterminate due to single **omission** of a functional morpheme)	*"And the first guy wasn't going them"*
H + L: combination of hierarchical and linearization errors	*"Anyway it's right at midnight she realize she's supposed to go back at midnight back"*
N: phonological errors **paraphasia**/neologism only	*"Well they ripped her outfit apout"*
S: semantic errors; grammatical but **semantically** incoherent	*"The slippers went off"*
FP: Filled Pauses/ hesitation	*"Okay, umm"*
G: **grammatical** utterance	*"And they were all trying on the shoes to see which of them could get into the shoes"*

79.6% of utterances and exact code in 69.8% of utterances. Raters agreed on the presence/absence of a hierarchical error on 80.1% of utterances and the presence/absence of a linearization error on 91.7% of utterances.

There was "moderate" agreement between the secondary raters and the consensus codes regarding grammaticality (k = 0.557, 95% CI [0.494, 0.620]) and "moderate" agreement between raters with regard to specific codes (unweighted k = 0.426, 95% CI [0.372, 0.481]). Consensus was most difficult differentiating between grammatical sentences from those containing hierarchical errors. The lack of agreement around what constitutes a hierarchical error is consistent with the interrater reliability results on paragrammatism from Matchin and colleagues (2020) as well as the lower interrater reliability reported by Casilio, Rising, Beeson, Bunton, and Wilson (2019) for paragrammatism relative to stereotypical agrammatic features.

Scanner Protocol and Lesion Mapping

High-resolution MRI data (T1- and T2-weighted images) were collected at the University of South Carolina and the Medical University of South Carolina.[3] Lesions were demarcated onto each participant's T2 image by an expert neurologist or an expert cognitive scientist, each blind to the behavioral data. Lesion maps were then aligned to the high resolution T1 image. Lesions were replaced with the corresponding brain structure from the intact hemisphere, and this image as well as the lesion map in participant space were subsequently warped to the Montreal Neurological Institute space (Nachev, Coulthard, Jäger, Kennard, & Husain, 2008) using SPM12 (Ashburner & Friston, 2005). The warped lesion map was then binarized with a 50% probability threshold, which was used to perform voxel-wise analyses.

Analysis

For each participant, the total number of utterances of each error type were summed and divided by the participant's total number of nonhesitation utterances. This resulted in a proportion of each error type for each participant. Because the number of intelligible, nonhesitation utterances varied drastically by participant (range = 10–148, mean = 39.0), a simple count of each error type would not be appropriate. We also calculated the proportion of utterances containing phonological (Phon) and semantic (Sem) paraphasias, respectively. Note that we only coded for the presence or absence of errors in each utterance, rather than counting errors of each kind in each utterance. Given that mixed errors contain both hierarchical and linearization errors, we calculated two different proportions for hierarchical and linearization errors each: one that contains mixed errors (all) and one that excludes mixed errors ("pure" errors). As such, mixed errors are excluded from the "pure" error analyses. The (all) proportions allow the broadest possible analyses of each error type, whereas the ("pure") proportions allow the analyses that identify the most separable lesion correlates of each error type.

Lesion-behavior Mapping

For each error type, we ran a whole-brain voxel-based lesion-symptom mapping analysis using SVRLSM (DeMarco & Turkeltaub, 2018). Images were resampled to 2-mm^3 isotropic voxels. We corrected for lesion volume by regressing it out of both behavior and lesion data and included only voxels with damage in at least eight participants (10.1%; coverage shown in Figure 2A). Resulting voxel-wise SVR-β values were thresholded at $p < .005$ based on 10,000 permutations. For each syntactic error type, we used phonological and semantic error rates, number of utterances produced, and WAB-AQ Fluency and object naming scores as nuisance covariates. In the resulting thresholded statistical maps, we used the connected_components method in nilearn (Abraham et al., 2014) to extract coordinates and volume for groups of suprathreshold voxels (Table 3). Only clusters with a volume of 800 mm^3 or greater are shown in the table.

To assess the robustness of the SVR-LSM results to individual participant influence, we conducted a subsampling analysis in which, for each behavioral variable, four participants were randomly removed from the full sample and the SVR-LSM analysis was rerun on the remaining participants. This procedure was repeated 50 times per variable, each time drawing a new random subsample without replacement. For each iteration, permutation-based thresholding was applied using 1000 permutations at a voxelwise threshold of $p < .005$ and a clusterwise threshold of $p < .05$. To quantify the spatial consistency of results across iterations, we computed pairwise Pearson correlations between the unthresholded beta maps from all 50 runs. In addition, each subsampled beta map was correlated with the beta map from the full-sample analysis to assess how closely the reduced-sample solutions approximated the complete data set solution. All robustness analyses were implemented in MATLAB using a custom scripted pipeline that interfaced with the SVRLSM toolbox.

Predictive Validity Comparison

To determine whether the different observed behavioral patterns in our data are the result of the same underlying lesion damage, we conducted supplemental predictive validity comparison (PVC) analyses (Magnotti, Patterson, & Schnur, 2023) based on sparse canonical correlation analysis for neuroimaging (Pustina, Avants, Faseyitan, Medaglia, & Coslett, 2018), using the default settings in the PVC software from Magnotti and colleagues (2023). PVC establishes that two lesion-behavior maps are distinct if and only if they provide unique predictive power for the behaviors being assessed: PVC tests whether participants' behavioral scores can best be predicted by a single (null)

Table 2. Distribution of Error Types across Participants

Score	*M*	*SD*	*Max*	*Min*
# Utterances	39.0	30.9	148	10
% Grammatical	0.566	0.230	1.0	0
% Mixed	0.045	0.080	0.538	0
% All hierarchical	0.265	0.156	0.654	0
% Pure hierarchical	0.220	0.134	0.60	0
% All linearization	0.148	0.229	0.944	0
% Pure linearization	0.103	0.217	0.944	0
% Omissions	0.066	0.073	0.286	0
% Phonological errs	0.281	0.228	0.852	0
% Semantic errs	0.177	0.148	0.731	0

lesion model or distinct (alternative) lesion models for each behavior. To investigate whether our different syntactic error types, as well as phonological and semantic deficits, might be due to the same underlying neural tissue, we performed these PVCs pairwise between the syntactic error type rates, as well as between each of the syntactic error types and proportions of phonological and semantic errors.

Atlas-based Fiber Tracking

To investigate the extent to which the significant lesion clusters were associated with disruption of white matter tracts, we extracted tractograms associated with significant clusters derived from our voxel-based lesion-symptom mapping analyses (Table 3) using DSI Studio (https://dsi-studio.labsolver.org). We ran deterministic fiber tracking[4] with each cluster as a ROI to filter out fibers that do not pass through that region. Identified fibers were automatically clustered into their corresponding tracts. Based on the tractography associated with each ROI, we next

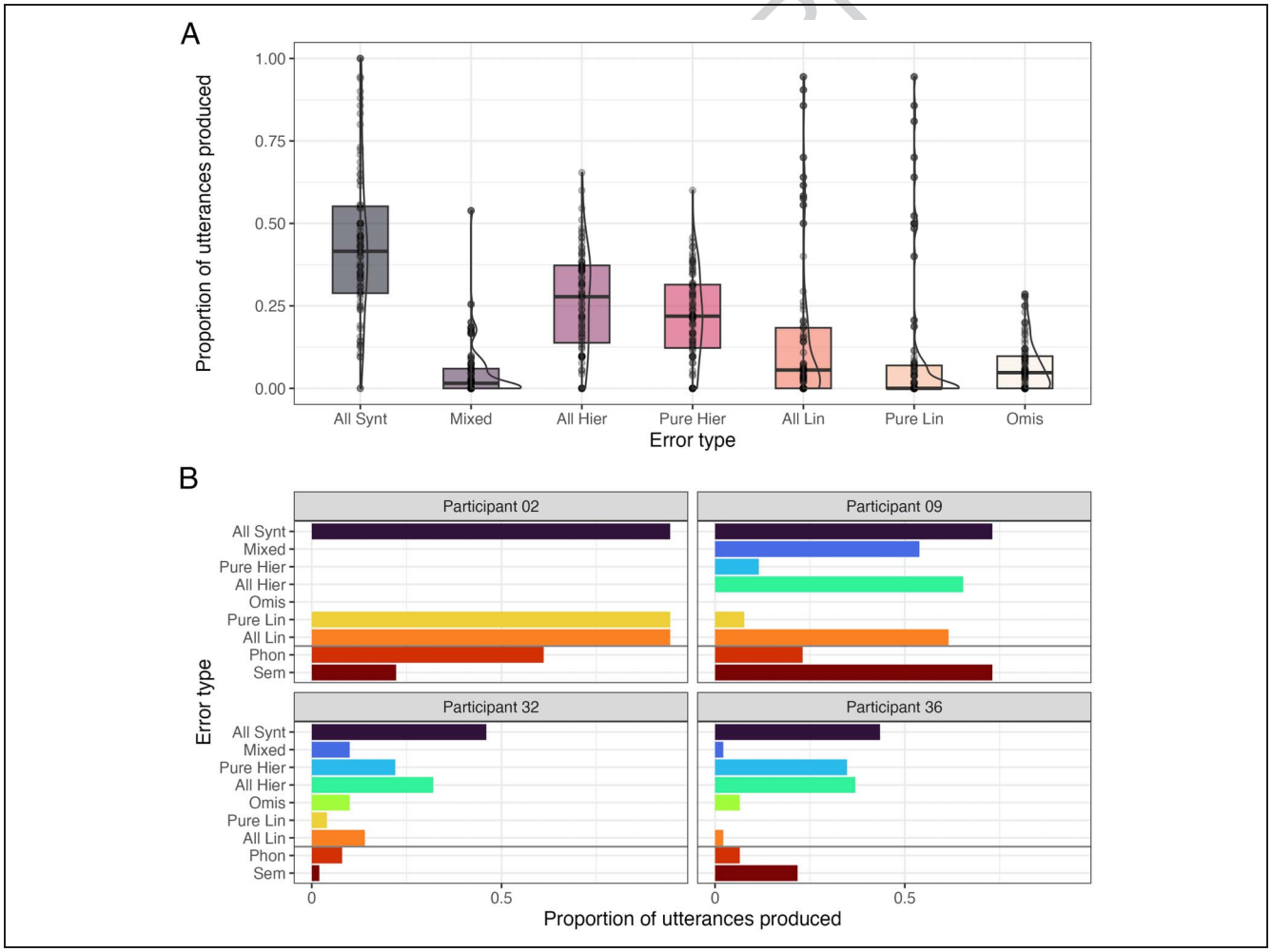

Figure 1. Behavioral results. (A) Overall distribution of error types across participants. (B) Error distributions for select participants illustrating divergent patterns of errors. Scores above the horizontal gray line are syntactic (all syntactic errors, mixed errors, pure hierarchical errors, all hierarchical errors, omission errors, pure linearization errors, all linearization errors), whereas scores below the line are nonsyntactic (phonological and semantic errors).

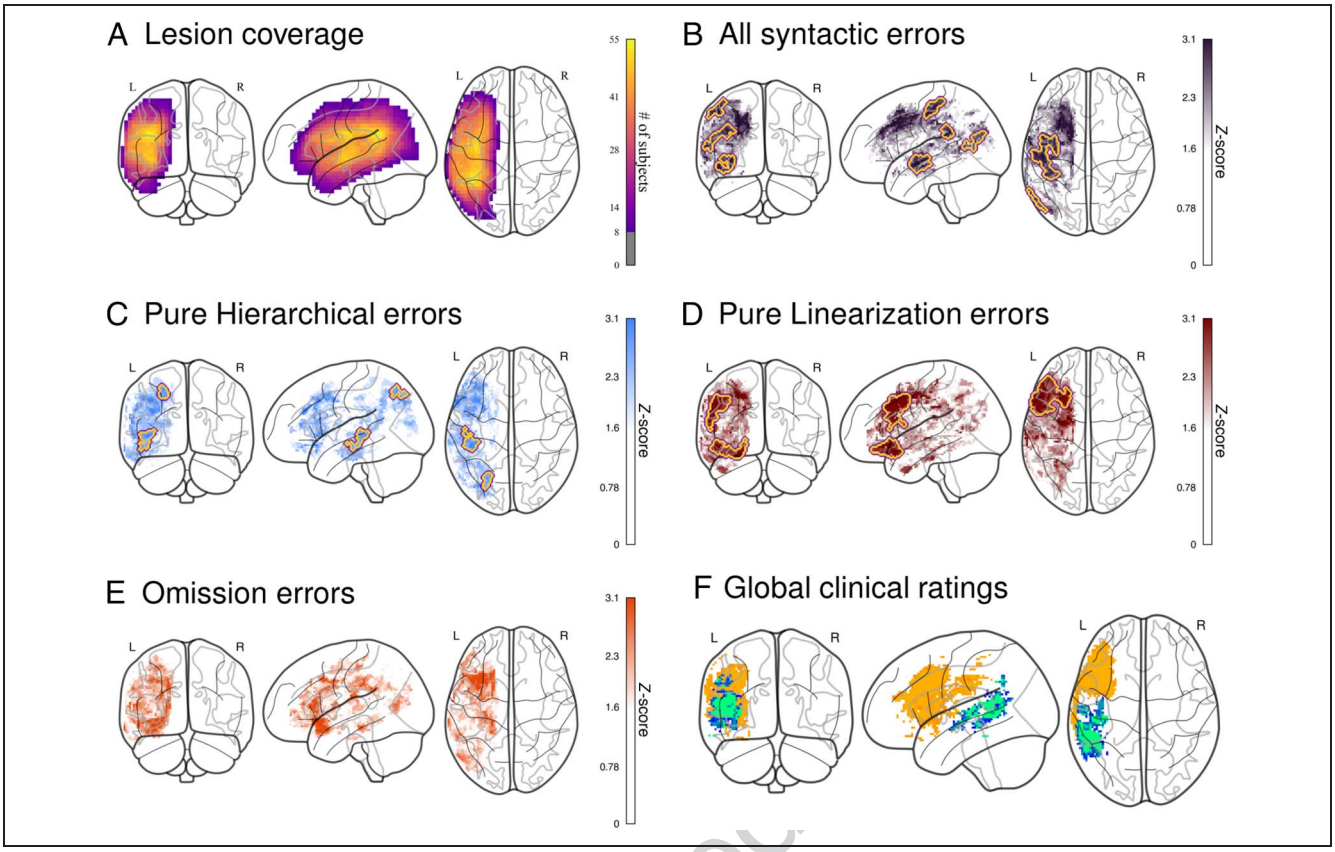

Figure 2. Lesion-behavior mapping results. (A) Lesion overlap showing voxels included in the analysis with damage in at least five participants. (B–E) Unthresholded lesion maps corresponding to (B) overall syntactic error rate by collapsing across all syntactic error types (i.e., hierarchical, linear, mixed, or omissions), (C) pure hierarchical error rate (i.e., excluding mixed hierarchical + linearization errors), (D) pure linearization error rate (i.e., excluding mixed hierarchical + linearization errors), and (E) omission error rate. Color scale denotes *z* value. Extracted clusters are outlined in yellow (NB: no clusters of sufficient size were found for omission errors). (F) Significant voxels associated with impressionistic ratings of agrammatic (warm colors) or paragrammatic (cool colors) from Matchin and colleagues (2020).

computed a disconnection index, defined as the number of fibers intersecting each ROI divided by its volume. This allows us to quantify the extent to which an ROI was interconnected to other parts of the brain and/or the extent to which an ROI contributes to disconnections between other regions, controlling for size of the ROI.

RESULTS

Behavior

Group average proportions of errors in each category can be found in Table 2 and are visualized in Figure 1A. Participants demonstrated a wide range of behavioral profiles, with some presenting with a complete or near-complete inability to produce a syntactically well-formed utterance and others who committed few-to-no syntactic errors in their elicitations (Figure 1B). Overall, participants produced an average of 39.0 utterances (SD = 30.9). Within those utterances, participants ranged widely over the proportion of error-containing utterances they produced as well as how their errors distributed over our categories. Figure 1B shows the error distributions for a few exemplar participants who showed clear patterns of linearization (Participant 02), mixed (Participant 09), or hierarchical (Participants 32 and 36) errors. More detailed analyses and examples of error types are available in Supplement 1.

Lesions Associated with Overall Syntactic Error Rate

We first conducted a lesion-symptom mapping analysis looking at overall syntactic error rate by aggregating over the different syntactic error types (hierarchical, linear, mixed, or omissions). The lesion map corresponding to overall syntactic error rate identifies scattered voxels across much of the canonical left-hemisphere language network, extending through parts of the frontal lobe, superior temporal lobe, and parts of the superior and inferior parietal lobes; however, only clusters in the temporal and parietal lobes met our cluster size cutoff (Figure 2B).

Lesions Associated with Overall Hierarchical and Linearization Error Rates

Next, we asked whether the broad network implicated in syntactic ability could be partitioned into computational

subcomponents. To do so, we examined the lesions corresponding to the overall rate of hierarchical and linearization errors. In each of these, we used the proportion of utterances presenting with hierarchical errors (i.e., utterances coded as hierarchical or mixed hierarchical + linearization errors) or proportion of utterances with linearization errors (i.e., utterances coded as containing linearization errors or mixed hierarchical + linearization errors). As such, utterances coded as mixed hierarchical + linearization were included in both scores. Under this approach, a dissociation between hierarchical and linearization errors emerged: Hierarchical errors were associated with damage to the middle temporal gyrus (MTG), as well as parts of the posterior and superior parietal lobe. Linearization errors, on the other hand, were associated with damage to two disjoint clusters: in the middle frontal gyrus (MFG) and the pars orbitalis of the inferior frontal gyrus (IFG; Table 3).

Lesions Associated with "Pure" Hierarchical and Linearization Errors

Because the above analyses included mixed errors in both rates—which added power but also added noise—we conducted further analyses for the rates of hierarchical and linearization errors, using only the rate of "pure" errors of each kind, that is, excluding mixed errors. The results of this analysis largely agree with the results observed above, which include mixed errors. Crucially, we do not observe any meaningful overlap between the voxels associated with hierarchical and linearization errors. The lesion map associated with hierarchical errors included a cluster in the MTG and mid-to-posterior superior temporal sulcus (STS) and in the medial parietal lobe (Figure 2C). Similar to above, "pure" linearization errors were associated with two frontal clusters in the MFG and IFG orbitalis (Figure 2D).

Lesions Associated with Errors of Omission

Although we found a few scattered suprathreshold voxels associated with errors of omission of functional syntactic elements (e.g., determiners, tense marking), none comprised a large enough cluster to meet our threshold of 800 mm^3 (Figure 2E).

Lesions Associated with Rates of Nonsyntactic Errors

Finally, we examined the lesions associated with nonsyntactic (i.e., phonological and semantic) error rates. Phonological errors (e.g., phonological paraphasias, neologisms) were associated with a posterior–inferior frontal and anterior inferior parietal lesion, extending into the superior insula (Supplemental Figure S1A). Semantic errors, on the other hand, were associated with two clusters encompassing most of the STS and MTG (Supplemental Figure S1B).

Table 3. Coordinates in Montreal Neurological Institute Space and Volume for Connected Components of Lesions

Error Type	*x*	*y*	*z*	*Size (mm^3)*	*Region*
All syntactic errors	−55.83	−69.46	17.91	1128	Inf Par
All syntactic errors	−50.13	−23.6	53.21	1968	poC
All syntactic errors	−41.48	−7.79	−7.19	2320	STS/MTG
All syntactic errors	−39.79	−39.89	27.93	1192	SMG
Pure hierarchical errors	−44.22	−27.47	−2.32	1544	STS/MTG
Pure hierarchical errors	−25.56	−68.99	48.24	952	Med Par
All hierarchical errors	−55.61	−13.9	−16.86	4488	MTG
All hierarchical errors	−39.61	−74.22	36.65	2680	Inf Par
All hierarchical errors	−29.81	−10.12	48.18	2056	poC
All hierarchical errors	−27.48	−61.11	51.68	1632	Med Par
All hierarchical errors	−31.09	−42.09	61.29	880	Sup Par
Pure linearization errors	−51.08	16.39	32.69	5616	MFG
Pure linearization errors	−38.18	25.87	−14.41	5144	IFG Orb
All linearization errors	−52.69	16.09	33.95	2816	MFG
All linearization errors	−52	21.73	−11.64	1456	IFG Orb

The lesions corresponding to the error types in italics are visualized in Figure 2. Inf Par = inferior parietal; poC = postcentral; SMG = supramarginal gyrus; Med Par = medial parietal; Sup Par = superior parietal.

Table 4. Results of Robustness Analyses

	Beta Map Cross-correlations				*Cross-correlation with Full Sample*	
Variable	*Mean r*	*SD*	*Max*	*Min*	*Mean r*	*SD*
All syntactic	.83	0.11	0.993	0.516	.888	0.105
Pure hierarchical	.891	0.041	0.995	0.773	.929	0.035
All hierarchical	.871	0.072	0.997	0.616	.901	0.075
Pure linearization	.826	0.107	0.999	0.505	.828	0.1
All linearization	.75	0.137	0.999	0.445	.757	0.15
Omissions	.916	0.044	1	0.654	.953	0.029

Robustness of Lesion-behavior Mapping Results

We find moderate-to-strong consistency under repeated subsampling of four participants across 50 iterations (Table 4), with mean pairwise correlations ranging from $r = .750$ (all linearization) to $r = .916$ (omissions). Correlations against the full-sample beta maps were uniformly higher than pairwise interrun correlations, ranging from $r = .757$ (all linearization) to $r = .953$ (omissions), suggesting that although individual subsampled runs vary somewhat from one another, they collectively remain close to the full-sample solution. Variability was greatest for linearization errors, which showed the widest ranges in pairwise r (.445–.999 and .505–.999, for all and pure, respectively) and the largest standard deviations, indicating occasional runs that diverged notably from the group. Overall, the results support the robustness of the SVR-LSM findings.

Comparison of Our Results to Findings from Matchin and Colleagues (2020)

To assess our results against prior work, we compared the SVRLSM results for our proportions of syntactic errors against the clinical consensus classifications from Matchin and colleagues (2020). We found that our lesion associated with an increased proportion of linearization errors was very similar to the lesion associated with impressionistic agrammatic ratings (compare Figure 2D with warm colors in Figure 2E). On the other hand, although one of the regions associated with the proportion of hierarchical error utterances overlapped with the lesion associated with a clinical classification of paragrammatism, the error-based lesion identified here was divided into two clusters: one in the mid- to posterior STS (overlapping with paragrammatism) and another in the medial parietal lobe.

Dissociation between Syntactic and Nonsyntactic Errors

We found that the model that assumes that hierarchical and linearization error rates arise from distinct cortical foci significantly outperformed the null model, which assumes they arise from the same underlying lesion (AIC of 1703.86 where values greater than 10 are considered decisive evidence in favor). The same dissociation was found in models comparing the lesions underlying syntactic error rates (pure hierarchical and linearization) to phonological (AIC = 816.63 and 766.60, respectively) and semantic (AIC = 1030.01 and 1224.60, respectively) error rates. In fact, we found that all pairwise comparisons of our error rates produced reliable dissociations, except (unsurprisingly, and reassuringly) for comparing all hierarchical with

Table 5. Pairwise PVC Results (AIC Values)

	All Hierarchical	*Pure Hierarchical*	*All Linearization*	*Pure Linearization*	*Omission*	*Semantic*
Pure hierarchical	−137.48					
All linearization	1062.89	1593.66				
Pure linearization	1517.79	1703.86	−250.69			
Omission	602.95	674.84	1730.75	1757.29		
Semantic	649.2	1030.01	765.15	1224.6	834.89	
Phonological	555.7	816.63	701.26	766.6	1036.44	906.8

All pairwise comparisons resulted in dissociations between the lesion correlates for the two tested behaviors (i.e., AIC > 10, indicating that a model that assumes different neural correlates for the two behaviors provides better prediction accuracy than a model assuming a single, shared neural correlate) except for all hierarchical versus pure hierarchical and all linearization versus pure linearization.

Table 6. Clusters Used for Fiber-tracking Analyses and Corresponding Tracts

Error Type	*Region*	*Size (mm^3)*	*Cluster Fibers*	*DI*	*Tract*	*Percent*	*Tract Fibers*	
Pure hierarchical	pSTS	1544	14,370	9.31	Inferior longitudinal fasciculus	45.78%	6579	
					Inferior fronto-occipital fasciculus	26.77%	3847	
					Arcuate fasciculus	11.21%	1611	
					Middle longitudinal fasciculus	10.70%	1538	
					Superior Longitudinal Fasciculus 2	0.02%	2	
	Med Par	952	655	0.69	Middle longitudinal fasciculus	40.15%	263	
						Superior Longitudinal Fasciculus 2	0.76%	5
						Superior Longitudinal Fasciculus 1	0.15%	1
						Inferior fronto-occipital fasciculus	0.15%	1
All hierarchical	MTG	4488	14,323	3.19	Inferior longitudinal fasciculus	67.63%	9686	
					Arcuate fasciculus	18.10%	2592	
					Inferior fronto-occipital fasciculus	8.09%	1158	
					Middle longitudinal fasciculus	3.23%	461	
	Inf Par	2680	3567	1.33	Superior Longitudinal Fasciculus 2	11.07%	394	
						Middle longitudinal fasciculus	8.80%	314
						Inferior fronto-occipital fasciculus	7.96%	283
						Inferior longitudinal fasciculus	2.27%	81
						Superior Longitudinal Fasciculus 3	0.56%	19
						Arcuate fasciculus	0.31%	10
	poC	2056	15,120	7.35	Superior Longitudinal Fasciculus 2	23.32%	3525	
					Frontal aslant tract	10.46%	1582	
					Arcuate Fasciculus	0.11%	17	
					Superior Longitudinal Fasciculus 1	0.05%	8	
	Med Par	1632	3520	2.16	Middle longitudinal fasciculus	27.19%	956	
						Superior Longitudinal Fasciculus 2	10.91%	384
						Inferior fronto-occipital fasciculus	5.60%	196
						Arcuate fasciculus	0.43%	14
						Superior Longitudinal Fasciculus 1	0.03%	1
						Inferior longitudinal fasciculus	0.03%	1

Table 6. (*continued*)

Error Type	*Region*	*Size (mm^3)*	*Cluster Fibers*	*DI*	*Tract*	*Percent*	*Tract Fibers*
	Sup Par	880	498	0.57	Superior Longitudinal Fasciculus 2	20.68%	102
					Arcuate fasciculus	2.01%	9
					Middle longitudinal fasciculus	1.81%	9
					Frontal aslant tract	0.60%	3
Pure linearization	MFG	5616	6550	1.17	Arcuate fasciculus	35.22%	2307
					Frontal aslant tract	33.54%	2197
					Superior Longitudinal Fasciculus 2	22.43%	1469
					Inferior fronto-occipital fasciculus	0.27%	17
					Superior Longitudinal Fasciculus 3	0.05%	2
	IFG Orb	5144	9713	1.89	Inferior fronto-occipital fasciculus	45.85%	4453
					Uncinate fasciculus	32.33%	3139
					Middle longitudinal fasciculus	0.08%	7
					Inferior longitudinal fasciculus	0.05%	4
					Arcuate fasciculus	0.02%	2
All Linearization	MFG	2816	2385	0.85	Arcuate fasciculus	32.96%	786
					Frontal aslant tract	30.90%	737
					Superior Longitudinal Fasciculus 2	27.97%	667
					Inferior fronto-occipital fasciculus	0.08%	1
	IFG Orb	1456	313	0.21	Inferior fronto-occipital fasciculus	69.01%	216
					Middle longitudinal fasciculus	17.57%	55
					Uncinate fasciculus	4.79%	14
					Inferior longitudinal fasciculus	3.51%	11
					Arcuate fasciculus	2.56%	7

Inf Par: inferior parietal; poC = postcentral; SMG = supramarginal gyrus; Med Par = medial parietal; Sup Par = superior parietal.

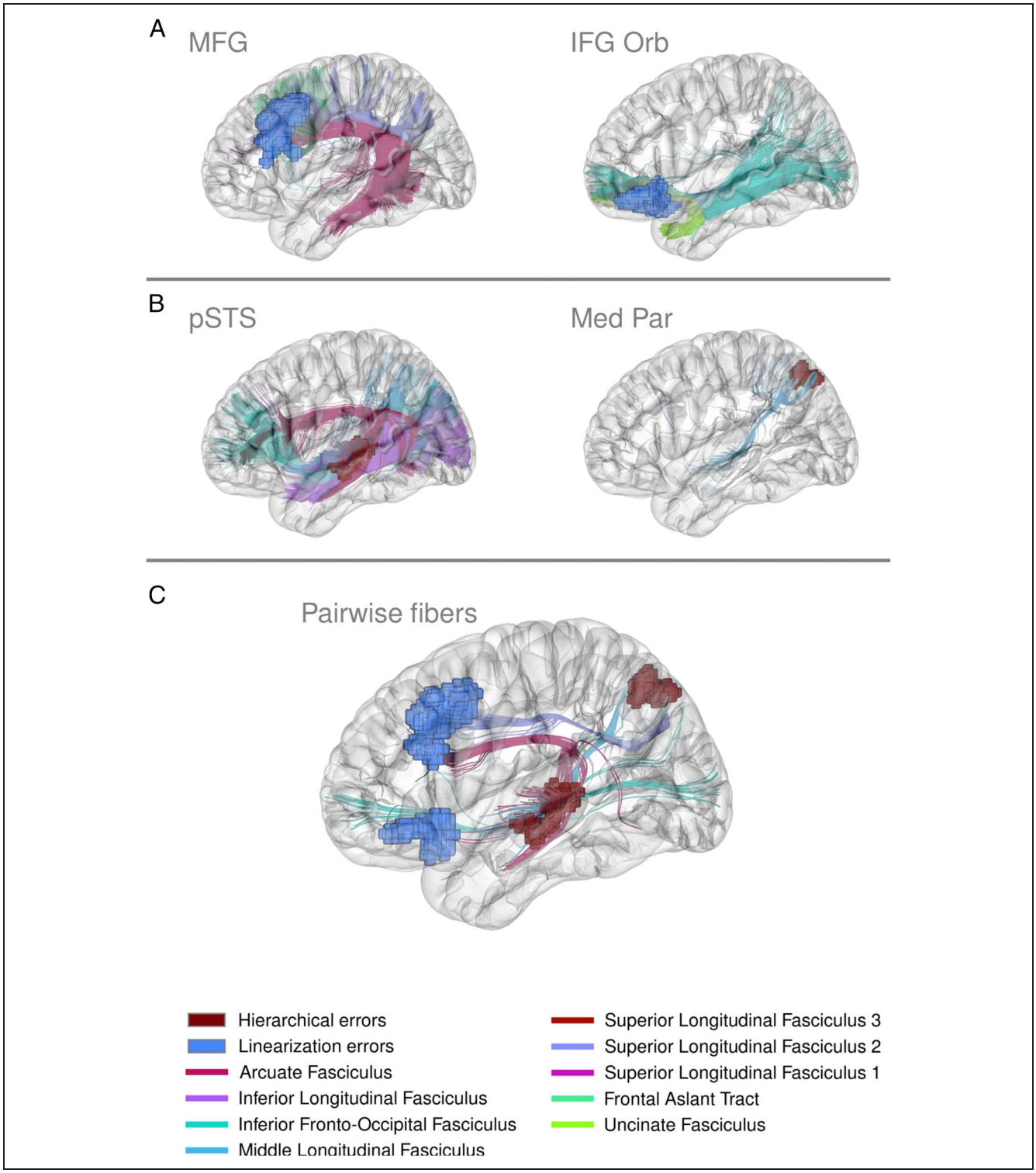

Figure 3. Atlas-based fiber-tracking results. (A) Columns IFGorb (IFG – pars orbitalis), pFC, and pMFG (posterior MFG)/prC/poC (precentral gyrus and postcentral gyrus) are from the clusters derived from the lesion maps for linearization errors. (B) Columns pSTS (posterior superior temporal sulcus) and Inf Par (inferior parietal) are from the clusters derived from the lesion maps from hierarchical errors. (C) Tracts connecting the pSTS cluster to the MFG cluster.

pure hierarchical and all linearization with pure linearization (Table 5).

Atlas-based Fiber Tracking

Cortical association tracts identified by our analyses, as well as the percentage and count of fibers interrupted by a given cluster, are given in Table 6. The highest disconnection index was associated with the STS/MTG cluster for pure hierarchical errors (9.31), followed by the post central cluster for all hierarchical errors (7.35). The tracts with the highest number of fibers interrupted by the hierarchical error ROIs were the inferior longitudinal, inferior fronto-occipital, middle longitudinal, and arcuate fasciculi. Tracts most interrupted by the linearization ROIs were the frontal aslant tract, arcuate, superior longitudinal, inferior fronto-occipital, and uncinate fasciculi. Finally, we identified tracts that passed through two or more of the identified ROIs (Figure 3C). We find that the middle longitudinal fasciculus has fiber that pass through both the hierarchical error posterior temporal ROI and the medial parietal hierarchical ROI. The arcuate fasciculus has fiber that passes through both the posterior temporal and MFG (linearization) ROIs. The inferior fronto-occipital fasiculus and inferior longitudinal fasciculus have fibers that pass through both the posterior temporal ROI and the IFG Pars Orbitalis (Orb) ROI. The superior longitudinal fasciculus has fibers that pass through both the medial parietal (hierarchical) and MFG ROIs.

DISCUSSION

This work aimed to identify the lesion correlates associated with different expressive syntactic error types in a cohort of individuals with chronic poststroke aphasia. We built on prior work by introducing a novel utterance-level analysis of grammatical errors in spontaneous speech by people with aphasia to characterize and quantify syntactic output rather than relying on holistic clinical impressions resulting in binary classifications (Casilio et al., 2019, 2025; Matchin et al., 2020). Our results support the double dissociation identified by Matchin and colleagues (2020) using a more granular and higher-dimensional behavioral measure in an expanded sample of participants. We found that hierarchical and linearization deficits in syntactic production characterized at the utterance level doubly dissociate. More precisely, hierarchical errors, theorized to be typical of paragrammatism, were associated with posterior temporal–parietal lesions, and linearization errors, theorized to be typical of agrammatism, were associated with middle and inferior frontal lesions. This neural dissociation of grammatical errors converges with the cognitive-behavioral dissociation encoded in popular two-stage models of grammatical encoding in sentence production (Yeaton et al., 2025), which posit distinct processing stages associated with hierarchical and linearization mechanisms.

Neural Basis of Sentence Planning

Previous lesion-symptom mapping work using clinical classifications of agrammatism and paragrammatism found that individuals classified as agrammatic had lesions predominantly in the inferior and middle frontal lobe, whereas participants classified as paragrammatic had lesions in the posterior temporal lobe (Casilio et al., 2025; Matchin et al., 2020; Figure 2F). Our lesion-symptom mapping results from rates of linearization errors corroborate prior findings associating expressive agrammatism with lesions to the frontal lobe (Figure 2D). Consistent with Matchin and colleagues' (2020) findings, we found that lesion correlates associated with the proportion of hierarchical errors were spatially disjoint from the lesion correlates of linearization errors and localized to posterior temporoparietal cortex (Figure 2C). Although Matchin and colleagues' (2020) finding for paragrammatic classification covers significant portions of the posterior temporal lobe, our result is more circumscribed and we see two disjoint clusters emerge: one in the MTG, STS, and underlying white matter and the other in the medial parietal lobe.

Corroborating our findings, atrophy to both white matter and cortex in both of the regions associated with pure hierarchical errors has recently been shown to associate with complex syntactic output in discourse in primary progressive aphasia (Matias-Guiu et al., 2022). In addition, we note that the extensive number of white matter tracts that pass through the posterior temporal cluster (Figure 3B) suggests that this may be the most consistent lesion site that produces behavioral deficits due to the functional disruption caused by disconnecting core parts of the language processing network. It is possible that the two lesion sites in MTG/STS and parietal lobe (and the corresponding white matter disconnections) both reflect disruption to the same underlying hierarchical processing system; however, it could be the case that each lesion site reflects disruption to distinct underlying subsystems, with breakdowns in either subsystem resulting in similar-appearing hierarchical deficits in output. One of these areas might underlie the actual hierarchical computations needed to construct a well-formed syntactic tree (Matchin & Hickok, 2020). The other might support sensorimotor transformations between the abstract hierarchical trees generated by the first zone and the linearization system in the frontal lobe—where damage leads to a sort of syntactic conduction aphasia (Hickok, 2025). Although we do not have comprehensive sentence comprehension assessments from the participants we report here, data from recent work in sentence comprehension in aphasia points to lesion correlates nearby our STS site (Biondo, Ivanova, Pracar, Baldo, & Dronkers, 2024). We would therefore predict that individuals who present with high numbers of hierarchical errors and comprehension deficits should

show lesion correlates in the posterior temporal lobe. By contrast, we would predict that individuals who present with high numbers of hierarchical errors but not comprehension deficits should show lesion correlates in the inferior parietal lobe.

In a similar way, it is possible that agrammatic output might arise from different underlying causes. Recent meta-analytic (Yeaton, 2025) and intracranial (Morgan et al., 2025a, 2025b) evidence reveals distinct areas in the frontal lobe associated with speech planning: one involving the posterior MFG, including parts of Area 55b, and another in the inferior frontal lobe, in the territory of the canonical Broca's area (Hickok, Venezia, & Teghipco, 2023). These regions are posited to have distinct roles in the sentence production pipeline: The more dorsal area supports phrasal prosody, which likely serves as a planning frame for morphosyntactic sequencing, whereas the ventral area supports speech sequencing (Hickok et al., 2023). Whereas it seems that damage to the more dorsal area may cause agrammatic output due to a breakdown in the computations supporting linearization, damage to the more ventral area seems to be more implicated in phonological errors, fluency disruptions (Supplemental Figure S1C), and apraxia of speech (Kurteff, Walker, Fridriksson, & Hickok, 2025). In addition, each of the frontal lesion sites implicated in linearization errors were associated with distinct tractographies, suggesting distinct functional roles and contributions to the overall syndrome of expressive agrammatism.

Somewhat surprisingly, we find that the clusters based on pure linearization errors are larger than those found from linearization + mixed errors (10,760 mm^3 vs. 4,272 mm^3, respectively); however, the inverse is true for hierarchical errors: Clusters for hierarchical + mixed errors are larger than for pure hierarchical errors (11,736 mm^3 vs. 2,496 mm^3, respectively). Intuitively, this suggests that mixed errors might be better characterized as a type of hierarchical error, rather than both a hierarchical and a linearization error occurring in the same sentence; however, more data are needed to investigate this. We further note that the clusters associated with all hierarchical errors are somewhat distributed across the posterior language network. This may be attributable to hierarchical processing being a more complex computational process that requires multiple subcomputations to be executed effectively. Given that "hierarchical processing" encompasses multiple computational stages of the sentence production pipeline (e.g., lexical selection, role assignment, phrase construction), we speculate that a finer-grained version of this system might capture errors at more precise levels of processing that could then be more precisely localized to one of the clusters we find here.

A Localized but Distributed Syntactic Network

Our results further provide some explanation for findings of a distributed syntactic system (Hu et al., 2023; Fedorenko, Blank, Siegelman, & Mineroff, 2020; Siegelman, Blank, Mineroff, & Fedorenko, 2019) or disparate patterns of brain damage underlying syntactic deficits (e.g., Caplan et al., 2007; Dick et al., 2001; Caplan, Hildebrandt, & Makris, 1996). Despite the relatively small size of the cluster related to pure hierarchical processing in the MTG/STS, our tractography results show that this region is widely connected across the language network, primarily within temporal–parietal regions, but also including fibers connecting to the MFG region associated with linearization error rates. This connection lends credence to the two-stage model of grammatical encoding: Hierarchical trees are generated first in the posterior temporal lobe and then converted to linear strings in the posterior inferior-middle frontal lobe.

Interestingly, the IFG Orb region associated with pure linearization errors is more connected with semantic processing regions in the anterior temporal lobe and inferior angular gyrus/TPJ (via the uncinate and inferior frontal-occipital fasciculi). This organization makes sense if this anterior–inferior frontal region supports motor coordination based on semantic information rather than syntax, consistent with previous suggestions about this region's involvement in controlled semantic retrieval (Binder & Desai, 2011; Lau, Phillips, & Poeppel, 2008). Proposals of a distributed syntactic system, however, tend to approach syntactic processing as a monolithic phenomenon. Although indeed our results point to syntactic processing being distributed in different parts of the language network, they also strongly suggest spatially differentiated functions of some kind. This is consistent with recent theoretical models regarding the cortical localization of syntactic processing, which posit a distribution of labor between posterior temporal sites that support hierarchical computation and frontal sites that support linearization and other aspects of lexical processing and cognitive control (Murphy, 2024; Matchin & Hickok, 2020; see Yeaton, 2025, for a review). These results are inconsistent with the predictions of an undifferentiated network-based approach (Fedorenko et al., 2020). Under such an account, syntax is epiphenomenal to semantic and lexical computations that are localized but distributed across the language network. This would predict that damage to any of the various semantic or lexical nodes that comprise the distributed network supporting syntactic computations should produce a single pattern of syntactic errors (as opposed to errors due to other processing deficits). This is not borne out in our data: Not only do we observe distinct patterns of errors, but we also observe distinct neural correlates. To account for the findings we present here, a network-based approach to syntax would need to integrate subnetworks or network states that support different aspects of (morpho-)syntactic computation.

On the other hand, it remains possible that agrammatic output in some talkers might be a resource-rational adaptation to high costs of production: An individual for whom speaking is very costly (e.g., due to severe anomia or other

cognitive or memory/planning limitations) might strategically elect to produce utterances that resemble linearization errors due to strategic omission of function words with limited conceptual content, thus allocating more resources to semantically rich content words, while still maintaining canonical word order (Fedorenko, Ryskin, & Gibson, 2023; Rezaii, Mahowald, Ryskin, Dickerson, & Gibson, 2022; Kolk & Heeschen, 1992). That said, a detailed account of how impairments in nonlinguistic processing mechanisms produce the specific error types identified here is currently lacking. Moreover, it is hard to capture the error types we observe from paragrammatic patients under a resource rational approach, as they are not shorter, more efficient, or semantically richer.

Nonimplication of Broca's Area in Syntactic Deficits

Conspicuously absent from our suite of our analyses are findings pointing to canonical Broca's area (i.e., IFG pars opercularis and pars triangularis) playing a key role in hierarchical or linear syntactic processing in production. This converges with previous evidence that damage restricted to this region alone does not cause Broca's aphasia (Pracar, Biondo, Dronkers, & Ivanova, 2025; Andrews et al., 2023; Fridriksson, Fillmore, Guo, & Rorden, 2015; Mohr et al., 1978) and recent lesion-symptom mapping analyses that failed to identify any significant lesion correlates for a large suite of aphasia assessment measures for this region (Herron et al., 2024). This suggests that, contrary to historical proposals, this region may not be implicated in chronic agrammatic speech deficits (Hickok et al., 2023). On the other hand, some possible explanations for the historical association of agrammatic deficits with Broca's area arise from our findings. The most simplistic of these is that Broca's area is sandwiched between the MFG and IFG Orb regions we found to be associated with linearization deficits and therefore lumped in with previous findings in this vicinity. Another possibility is that this Broca's area supports fluency that is often confounded with agrammatic output. Indeed, we find that the WAB-R Fluency scores from our sample are associated with damage to the posterior portion of the IFG and inferior premotor area (Supplemental Figure 1C). Our fluency lesion finding also overlaps with regions shown to cause apraxia of speech (Kurteff et al., 2025; Lorca-Puls et al., 2024), which may contribute to the nonfluency that is typically comorbid with agrammatic deficits.

Notably, Broca's area is not a cytoarchitectonically unified region but instead consists of multiple side-by-side subregions that support different domain-general and linguistic functions (Fedorenko & Blank, 2020; Fedorenko, Duncan, & Kanwisher, 2012) including syntactic (Amici et al., 2007; Fiebach, Schlesewsky, Lohmann, Von Cramon, & Friederici, 2005) and nonsyntactic working memory (Rogalsky & Hickok, 2011; Rogalsky, 2008); however, others have proposed that subregions support different syntactic functions (Zaccarella, Papitto, & Friederici, 2021). These proposals—especially those positing a role for Broca's area in (syntactic) working memory—are parsimonious with our findings. Given the context of the elicitation method used in our study, participants might adapt a metacognitive strategy not to attempt highly complex embedded structures, which would unnecessarily tax working memory and cognitive control resources and increase the risk of error commission. We acknowledge that the framework we use here does not take syntactic complexity into account (Peristeri, Nerantzini, Drakoulaki, Boznou, & Varlokosta, 2024; Rezaii et al., 2022; Gleichgerrcht et al., 2021) and might therefore be blind to Broca's area's role in supporting the production of more syntactically complex sentences. Assuming participants have adopted this resource-rational approach could then explain the lack of Broca's area findings (i.e., syntactic working memory is not implicated), while not ruling out that linearization deficits remain (Fedorenko et al., 2023).

Limitations and Open Questions

A significant limitation of this approach is that error types must be inferred from the annotator's best guess at what the intended target sentence was, as the ground truth is not available. As with Matchin and colleagues (2020), interrater reliability was lower on samples associated with paragrammatism/hierarchical processing deficits. Theoretically, by nature of occurring earlier in the production process, such deficits would have more chances to variably affect the surface form of the utterance, potentially resulting in a wider variety of surface-level errors. We acknowledge a potential circularity in using lesion correlates to partially validate our coding methodology: The presence of meaningful lesion correlates does not per se guarantee that the codes capture the intended theoretical constructs. Importantly, however, the lesion findings converge with prior results obtained using an entirely independent methodology (impressionistic clinical ratings; Matchin et al., 2020) and are supported by recent computational modeling and lesion-parameter mapping evidence (Yeaton et al., 2026), lending some weight to the interpretation that the codes capture a genuine distinction. We view the moderate interrater agreement—particularly for hierarchical errors—as both a limitation and an indication that greater theoretical consensus around the definition of paragrammatism is needed before the coding scheme can be considered fully objective. Future work should prioritize the development of a formalized, operationally specified coding protocol with documented decision rules and evaluate whether such a protocol yields higher reliability.

The categories we present here offer a practical, if coarse-grained, operationalization of psycholinguistic theory that advances the field by promoting utterance-level, error-based characterization of syntactic deficits over broader syndrome-based classification. Extending this

framework cross-linguistically will require integration of typologically diverse language data to develop a unified coding approach applicable across languages. Future improvements and refinements may allow for greater ease in identifying deficit types from short discourse samples; however, the existing debate regarding the distinction between agrammatic and paragrammatic errors will need further resolution before wider translational applications can be implemented.

Beyond limitations of the coding scheme, the theoretical framework itself carries important caveats. There are many computations and cognitive operations that go into generating a well-formed sentence besides simply generating a syntactic tree and subsequently linearizing it. It is possible that other cognitive or linguistic constraints (e.g., error monitoring, cognitive control) might in some cases give rise to ungrammatical output despite intact syntactic processing, as is observed in sentence production errors among neurologically healthy adults (Butterworth & Howard, 1987). To the extent that executive and monitoring processes interact with the syntactic system, their disruption might be expected to produce error patterns that partly mimic either hierarchical or linearization deficits depending on which stage of the production pipeline they affect. For instance, failures in error monitoring might allow malformed hierarchical structures to pass through to output without correction, superficially resembling paragrammatic errors; impaired cognitive control or working memory might result in abandoned or truncated syntactic frames, resembling agrammatic output. The extent to which error monitoring itself relies on linguistic mechanisms of production, comprehension, or other cognitive systems is also somewhat unclear (Nozari & Novick, 2017; Nozari, Dell, & Schwartz, 2011; Levelt, 1989). Disentangling these interactions will require paradigms that orthogonally manipulate syntactic and executive demands, and ideally incorporate measures of error self-correction and repair behavior. As a clearer theoretical and computational picture emerges of the relationship between executive, working memory, and error monitoring processes and the syntactic system (Hagoort, 2016; Peristeri et al., 2024), future investigations may consider integrating additional metrics to address these paralinguistic influences and their interactions with syntactic processing.

Conclusions

This study provides promising evidence for the dissociation of hierarchical and linearization errors in aphasia, linking distinct patterns of syntactic errors to specific brain regions. The neural correlates of these errors support the two-stage model of syntactic encoding, with posterior temporal–parietal lesions associated with hierarchical errors and frontal lesions with linearization deficits. These findings contribute to a more nuanced understanding of the neurobiological underpinnings of syntactic processing in aphasia, offering valuable insights for clinical assessment and intervention. By distinguishing between the neural bases of different error types, this work paves the way for more targeted treatments that address the specific cognitive and neural mechanisms involved in sentence production deficits. Future research should aim to further explore the relationship between lesion location, error type, and treatment outcomes, and expand these findings to other forms of aphasia and neurodegenerative conditions.

Corresponding author: Jeremy Yeaton, Department of Neurological Surgery, University of California, Davis, CA, e-mail: jyeaton@health.ucdavis.edu.

Data Availability Statement

Data can be shared contingent upon an institutional data-sharing agreement due to the sensitive nature of Clinical Neuroimaging data. Supplemental Material can be accessed on this article's homepage: https://doi.org/10.1162/JOCN.a.2684.

Author Contributions

Jeremy Yeaton: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Writing—Original draft; Writing—Review & editing. Danielle Fahey: Conceptualization; Data curation; Formal analysis; Investigation; Writing—Original draft; Writing—Review & editing. Michelle Gomez: Data curation; Formal analysis; Writing—Review & editing. Brielle C. Stark: Conceptualization; Validation; Writing—Review & editing. Julius Fridriksson: Funding acquisition; Supervision; Writing—Review & editing. Dirk B. den Ouden: Funding acquisition; Supervision; Writing—Review & editing. Gregory Hickok: Funding acquisition; Methodology; Supervision; Writing—Review & editing. William Matchin: Conceptualization; Formal analysis; Funding acquisition; Methodology; Supervision; Visualization; Writing—Original draft; Writing—Review & editing.

Declaration of AI and AI-assisted Technologies

The authors declare that the following tool(s) or service(s) were used: Claude. AI use for research tasks: Used to produce custom code for requested analysis; AI use for manuscript preparation Original.

Funding Information

National Institute on Deafness and Other Communication Disorders (https://dx.doi.org/10.13039/100000055), grant number: DC014664. Arnold School of Public Health, University of South Carolina (https://dx.doi.org/10.13039/100018427). William Orr Dingwall Foundation Dissertation Fellowship awarded to JY.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this paper report its proportions of citations by gender category to be: M/M = .593; W/M = .074; M/W = .148; W/W = .185.

Notes

1. Quotes are drawn from our own database. Transcription annotations and filled pauses have been removed for clarity.
2. Some participants—especially those who were severely nonfluent—did not produce many utterances; however, we found that the 10-utterance minimum permitted sufficient insight into the distribution of errors.
3. See "Scanner protocol" in supplemental materials for scan parameters.
4. See supplemental section "Tractography atlas construction" for additional information.

REFERENCES

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., et al. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014, PubMed: 24600388

Amici, S., Brambati, S. M., Wilkins, D. P., Ogar, J., Dronkers, N. L., Miller, B. L., et al. (2007). Anatomical correlates of sentence comprehension and verbal working memory in neurodegenerative disease. *Journal of Neuroscience*, *27*, 6282–6290. https://doi.org/10.1523/JNEUROSCI.1331-07.2007, PubMed: 17554002

Andrews, J. P., Cahn, N., Speidel, B. A., Chung, J. E., Levy, D. F., Wilson, S. M., et al. (2023). Dissociation of Broca's area from Broca's aphasia in patients undergoing neurosurgical resections. *Journal of Neurosurgery*, *138*, 847–857. https://doi.org/10.3171/2022.6.JNS2297, PubMed: 35932264

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*, 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018, PubMed: 15955494

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*, 527–536. https://doi.org/10.1016/j.tics.2011.10.001, PubMed: 22001867

Biondo, N., Ivanova, M. V., Pracar, A. L., Baldo, J., & Dronkers, N. F. (2024). Mapping sentence comprehension and syntactic complexity: Evidence from 131 stroke survivors. Brain. *Communications*, *6*. https://doi.org/10.1093/braincomms/fcae379, PubMed: 39554380

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. *Handbook of Psycholinguistics*, *5*, 405–452.

Butterworth, B., & Howard, D. (1987). *Paragrammatisms. Cognition*, *26*, 1–37. https://doi.org/10.1016/0010-0277(87)90012-6, PubMed: 3608394

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. https://doi.org/10.1093/brain/119.3.933, PubMed: 8673503

Caplan, D., Waters, G., Kennedy, D., Alpert, N., Makris, N., DeDe, G., et al. (2007). A study of syntactic processing in aphasia II: Neurological aspects. *Brain and Language*, *101*, 151–177. https://doi.org/10.1016/j.bandl.2006.06.226, PubMed: 16997366

Casilio, M., Kasdan, A. V., Bryan, K., Shibata, K., Schneck, S. M., Levy, D. F., et al. (2025). Four dimensions of naturalistic language production in aphasia after stroke. *Brain*, *148*, 291–312. https://doi.org/10.1093/brain/awae195, PubMed: 38889230

Casilio, M., Rising, K., Beeson, P. M., Bunton, K., & Wilson, S. M. (2019). Auditory-perceptual rating of connected speech in aphasia. *American Journal of Speech-Language Pathology*, *28*, 550–568. https://doi.org/10.1044/2018_AJSLP-18-0192, PubMed: 31136232

DeMarco, A. T., & Turkeltaub, P. E. (2018). A multivariate lesion symptom mapping toolbox and examination of lesion-volume biases and correction methods in lesion-symptom mapping. *Human Brain Mapping*, *39*, 4169–4182. https://doi.org/10.1002/hbm.24289, PubMed: 29972618

Den Ouden, D., 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*, 2153–2173. https://doi.org/10.1002/hbm.24514, PubMed: 30666767

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, *108*, 759–788. https://doi.org/10.1037/0033-295x.108.4.759, PubMed: 11699116

Fahey, D., Yeaton, J., Stark, B., Fridriksson, J., Den Ouden, D., Hickok, G., et al. (2025). An objective coding scheme for grammatical production deficits in aphasia reveals a categorical divide between agrammatism and paragrammatism. https://doi.org/10.31234/osf.io/7jw62

Fedorenko, E., & Blank, I. A. (2020). Broca's area is not a natural kind. *Trends in Cognitive Sciences*, *24*, 270–284. https://doi.org/10.1016/j.tics.2020.01.001, PubMed: 32160565

Fedorenko, E., Blank, I. A., Siegelman, M., & Mineroff, Z. (2020). Lack of selectivity for syntax relative to word meanings throughout the language network. *Cognition*, *203*, 104348. https://doi.org/10.1016/j.cognition.2020.104348, PubMed: 32569894

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*, 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011, PubMed: 23063434

Fedorenko, E., Ryskin, R., & Gibson, E. (2023). Agrammatic output in non-fluent, including Broca's, aphasia as a rational behavior. *Aphasiology*, *37*, 1981–2000. https://doi.org/10.1080/02687038.2022.2143233, PubMed: 38213953

Ferreira, V. S., & Slevc, L. R. (2007). Grammatical encoding. In *The Oxford handbook of psycholinguistics* (pp. 453–470).

Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198568971.013.0027, PubMed: 22832793
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. https://doi.org/10.1002/hbm.20070, PubMed: 15455462
Fridriksson, J., Fillmore, P., Guo, D., & Rorden, C. (2015). Chronic Broca's aphasia is caused by damage to Broca's and Wernicke's areas. *Cerebral Cortex*, *25*, 4689–4696. https://doi.org/10.1093/cercor/bhu152, PubMed: 25016386
Gleichgerrcht, E., Roth, R., Fridriksson, J., den Ouden, D., Delgaizo, J., Stark, B., et al. (2021). Neural bases of elements of syntax during speech production in patients with aphasia. *Brain and Language*, *222*, 105025. https://doi.org/10.1016/j.bandl.2021.105025, PubMed: 34555689
Goodglass, H. (1968). Studies on the grammar of aphasics. *Developments in Applied Psycholinguistics Research*, 177–208.
Goodglass, H., Christiansen, J. A., & Gallagher, R. (1993). Comparison of morphology and syntax in free narrative and structured tests: Fluent vs. nonfluent aphasics. *Cortex*, *29*, 377–407. https://doi.org/10.1016/S0010-9452(13)80250-X, PubMed: 8258282
Heeschen, C. (1985). Agrammatism versus paragrammatism: A fictitious opposition. In *Agrammatism* (pp. 207–248). Elsevier. https://doi.org/10.1016/B978-0-12-402830-2.50012-2
Herron, T. J., Schendel, K., Curran, B. C., Lwi, S. J., Spinelli, M. G., Ludy, C., et al. (2024). Is Broca's area critical for speech and language? Evidence from lesion-symptom mapping in chronic aphasia. *Frontiers in Language Sciences*, *3*. https://doi.org/10.3389/flang.2024.1398616
Hickok, G. (2025). *Wired for words: The neural architecture of language*. MIT Press. https://doi.org/10.7551/mitpress/9000.001.0001
Hickok, G., Venezia, J., & Teghipco, A. (2023). Beyond Broca: Neural architecture and evolution of a dual motor speech coordination system. *Brain*, *146*, 1775–1790. https://doi.org/10.1093/brain/awac454, PubMed: 36746488
Hu, J., Small, H., Kean, H., Takahashi, A., Zekelman, L., Kleinman, D., et al. (2023). Precision fMRI reveals that the language-selective network supports both phrase-structure building and lexical access during language production. *Cerebral Cortex*, *33*, 4384–4404. https://doi.org/10.1093/cercor/bhac350, PubMed: 36130104
Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbances. In *Studies on child language and aphasia* (pp. 49–74). De Gruyter Mouton.
Kean, M.-L. (1977). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition*, *5*, 9–46. https://doi.org/10.1016/0010-0277(77)90015-4
Kertesz, A. (2007). *Western Aphasia Battery–Revised*.
Kleist, K. (1914). Aphasie und geisteskrankheit. *Münchener Medizinische Wochenschrift*, *61*, 8–12.
Kolk, H., & Heeschen, C. (1992). Agrammatism, paragrammatism and the management of language. *Language and Cognitive Processes*, *7*, 89–129. https://doi.org/10.1080/01690969208409381
Krauska, A., & Lau, E. (2023). Moving away from lexicalism in psycho- and neuro-linguistics. *Frontiers in Language Sciences*, *2*. https://doi.org/10.3389/flang.2023.1125127
Kurteff, G. L., Walker, G. M., Fridriksson, J., & Hickok, G. (2025). *Mapping of critical prosodic and phonetic networks in post-stroke apraxia of speech*. https://doi.org/10.64898/2025.12.16.694760, PubMed: 41446164
Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*, 920–933. https://doi.org/10.1038/nrn2532, PubMed: 19020511
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001
Lorca-Puls, D. L., Gajardo-Vidal, A., Mandelli, M. L., Illán-Gala, I., Ezzes, Z., Wauters, L. D., et al. (2024). Neural basis of speech and grammar symptoms in non-fluent variant primary progressive aphasia spectrum. *Brain*, *147*, 607–626. https://doi.org/10.1093/brain/awad327, PubMed: 37769652
MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, *25*, 1286–1307. https://doi.org/10.1080/02687038.2011.589893, PubMed: 22923879
Magnotti, J. F., Patterson, J. S., & Schnur, T. T. (2023). Using predictive validity to compare associations between brain damage and behavior. *Human Brain Mapping*, *44*, 4738–4753. https://doi.org/10.1002/hbm.26413, PubMed: 37417774
Matchin, W., Basilakos, A., Stark, B. C., den Ouden, D.-B., Fridriksson, J., & Hickok, G. (2020). Agrammatism and paragrammatism: A cortical double dissociation revealed by lesion-symptom mapping. *Neurobiology of Language*, *1*, 208–225. https://doi.org/10.1162/nol_a_00010, PubMed: 34296193
Matchin, W., & Hickok, G. (2020). The cortical organization of syntax. *Cerebral Cortex*, *30*, 1481–1498. https://doi.org/10.1093/cercor/bhz180, PubMed: 31670779
Matias-Guiu, J. A., Suárez-Coalla, P., Yus, M., Pytel, V., Hernández-Lorenzo, L., Delgado-Alonso, C., et al. (2022). Identification of the main components of spontaneous speech in primary progressive aphasia and their neural underpinnings using multimodal MRI and FDG-PET imaging. *Cortex*, *146*, 141–160. https://doi.org/10.1016/j.cortex.2021.10.010, PubMed: 34864342
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*, 311–311. https://doi.org/10.1212/WNL.28.4.311, PubMed: 565019
Morgan, A. M., Devinsky, O., Doyle, W. K., Dugan, P., Friedman, D., & Flinker, A. (2025a). Decoding words during sentence production with ECoG reveals syntactic role encoding and structure-dependent temporal dynamics. Communications. *Psychology*, *3*, 87. https://doi.org/10.1038/s44271-025-00270-1, PubMed: 40461573
Morgan, A. M., Devinsky, O., Doyle, W. K., Dugan, P., Friedman, D., & Flinker, A. (2025b). *From single words to sentence production: Shared cortical representations but distinct temporal dynamics*. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2024.10.30.621177
Murphy, E. (2024). ROSE: A neurocomputational architecture for syntax. *Journal of Neurolinguistics*, *70*, 101180. https://doi.org/10.1016/j.jneuroling.2023.101180
Nachev, P., Coulthard, E., Jäger, H. R., Kennard, C., & Husain, M. (2008). Enantiomorphic normalization of focally lesioned brains. *Neuroimage*, *39*, 1215–1226. https://doi.org/10.1016/j.neuroimage.2007.10.002, PubMed: 18023365
Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, *63*, 1–33. https://doi.org/10.1016/j.cogpsych.2011.05.001, PubMed: 21652015
Nozari, N., & Novick, J. (2017). Monitoring and control in language production. *Current Directions in Psychological Science*, *26*, 403–410. https://doi.org/10.1177/0963721417702419

Peristeri, E., Nerantzini, M., Drakoulaki, K., Boznou, A., & Varlokosta, S. (2024). Narrative production and executive functions in post-stroke agrammatic aphasia. *Aphasiology*, *1–24*. https://doi.org/10.1080/02687038.2024.2393221

Pracar, A. L., Biondo, N., Dronkers, N. F., & Ivanova, M. V. (2025). The neuroanatomy of Broca's aphasia. *Frontiers in Language Sciences*, *4*, 1496209. https://doi.org/10.3389/flang.2025.1496209, PubMed: 41737654

Pustina, D., Avants, B., Faseyitan, O. K., Medaglia, J. D., & Coslett, H. B. (2018). Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. *Neuropsychologia*, *115*, 154–166. https://doi.org/10.1016/j.neuropsychologia.2017.08.027, PubMed: 28882479

Rezaii, N., Mahowald, K., Ryskin, R., Dickerson, B., & Gibson, E. (2022). A syntax–lexicon trade-off in language production. *Proceedings of the National Academy of Sciences*, *119*, e2120203119. https://doi.org/10.1073/pnas.2120203119, PubMed: 35709321

Rogalsky, C. (2008). Broca's area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*, 14. https://doi.org/10.3389/neuro.09.014.2008, PubMed: 18958214

Rogalsky, C., & Hickok, G. (2011). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*, *23*, 1664–1680. https://doi.org/10.1162/jocn.2010.21530, PubMed: 20617890

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., et al. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences, U.S.A.*, *118*, e2105646118. https://doi.org/10.1073/pnas.2105646118, PubMed: 34737231

Siegelman, M., Blank, I. A., Mineroff, Z., & Fedorenko, E. (2019). An attempt to conceptually replicate the dissociation between syntax and semantics during sentence comprehension. *Neuroscience*, *413*, 219–229. https://doi.org/10.1016/j.neuroscience.2019.06.003, PubMed: 31200104

Thompson, C. K., Den Ouden, D.-B., Bonakdarpour, B., Garibaldi, K., & Parrish, T. B. (2010). Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia*, *48*, 3211–3227. https://doi.org/10.1016/j.neuropsychologia.2010.06.036, PubMed: 20603138

Whitwell, J. L., Duffy, J. R., Machulda, M. M., Clark, H. M., Strand, E. A., Senjem, M. L., et al. (2017). Tracking the development of agrammatic aphasia: A tensor-based morphometry study. *Cortex*, *90*, 138–148. https://doi.org/10.1016/j.cortex.2016.09.017, PubMed: 27771043

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., et al. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*, 2069–2088. https://doi.org/10.1093/brain/awq129, PubMed: 20542982

Yeaton, J. (2025). The neurobiology of sentence production: A narrative review and meta-analysis. *Brain and Language*, *264*, 81–94. https://doi.org/10.1016/j.bandl.2025.105549, PubMed: 39983635

Yeaton, J., Fahey, D., Khosshal Mollasaraei, Z., Krauska, A., & Matchin, W. (2025). *Aligning psycholinguistics, neuroscience, and aphasiology with respect to grammatical encoding*. https://doi.org/10.31234/osf.io/ydj9z

Yeaton, J. D., Walker, G. M., Fahey, D., Matchin, W., Fridriksson, J., & Hickok, G. (2026). From structure to sequence: A multinomial processing tree model of syntactic encoding. *Psychonomic Bulletin & Review*. https://doi.org/10.31234/osf.io/w53f9_v2

Zaccarella, E., Papitto, G., & Friederici, A. D. (2021). Language and action in Broca's area: Computational differentiation and cortical segregation. *Brain and Cognition*, *147*, 105651. https://doi.org/10.1016/j.bandc.2020.105651, PubMed: 33254030