








RESEARCH ARTICLE

Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia

Dirk-Bart den Ouden¹  | Svetlana Maljutina¹  | Alexandra Basilakos¹  |
Leonardo Bonilha²  | Ezequiel Gleichgerrcht²  | Grigori Yourganov³ | Argye E. Hillis⁴  |
Gregory Hickok⁵ | Chris Rorden³  | Julius Fridriksson¹

¹Department of Communication Sciences and Disorders, University of South Carolina, Columbia, South Carolina

²Department of Neurology, Medical University of South Carolina, Charleston, South Carolina

³Department of Psychology, University of South Carolina, Columbia, South Carolina

⁴Department of Neurology, Johns Hopkins University School of Medicine and Department of Cognitive Science, Johns Hopkins University, Baltimore, Maryland

⁵School of Social Sciences, University of California, Irvine, California

Correspondence

Dirk-Bart den Ouden, Department of Communication Sciences and Disorders, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Columbia, SC 29208.

Email: denouden@sc.edu

Present address

Dr. Svetlana Maljutina is now at the National Research University Higher School of Economics.

Funding information

National Institute on Deafness and Other Communication Disorders, Grant/Award Numbers: P50 DC014664, T32 DC0014435

Abstract

Agrammatism in aphasia is not a homogeneous syndrome, but a characterization of a nonuniform set of language behaviors in which grammatical markers and complex syntactic structures are omitted, simplified, or misinterpreted. In a sample of 71 left-hemisphere stroke survivors, syntactic processing was quantified with the Northwestern Assessment of Verbs and Sentences (NAVS). Classification analyses were used to assess the relation between NAVS performance and morphosyntactically reduced speech in picture descriptions. Voxel-based and connectivity-based lesion-symptom mapping were applied to investigate neural correlates of impaired syntactic processing. Despite a nonrandom correspondence between NAVS performance and morphosyntactic production deficits, there was variation in individual patterns of syntactic processing. Morphosyntactically reduced production was predicted by lesions to left-hemisphere inferior frontal cortex. Impaired verb argument structure production was predicted by damage to left-hemisphere posterior superior temporal and angular gyrus, as well as to a ventral pathway between temporal and frontal cortex. Damage to this pathway was also predictive of impaired sentence comprehension and production, particularly of noncanonical sentences. Although agrammatic speech production is primarily predicted by lesions to inferior frontal cortex, other aspects of syntactic processing rely rather on regional integrity in temporoparietal cortex and the ventral stream.

KEYWORDS

agrammatism, aphasia, connectome, lesion-symptom mapping, magnetic resonance imaging, syntax

1 | INTRODUCTION

1.1 | Agrammatism and the neural correlates of syntactic processing

It has been recognized since the mid-nineteenth century that some persons with aphasia after brain damage (PWA) show language performance that can be characterized as “agrammatic” (Kleist, 1916; Kussmaul, 1877; Pick, 1913). Speech output lacks function words and bound morphemes (Goodglass & Berko, 1960) and there may be an overuse of nouns, coupled with specific difficulties with verbs and verb argument structure (Bastiaanse, Rispens, Ruigendijk, Rabadan, &

Thompson, 2002; Caplan, 1987; Kim & Thompson, 2000; Zingeser & Berndt, 1990). In general, sentence structure is simplified (Alajouanine, 1968; Goodglass, 1997), though paragrammatic performance has also been noted, characterized by (erroneous) substitutions rather than deletions of syntactic material (Butterworth & Howard, 1987; Heeschen, 1985; Kleist, 1916). Syntactic comprehension deficits (receptive agrammatism) may occur as well, but these are not necessarily limited to speakers with agrammatic output (Caramazza & Zurif, 1976).

Impaired production of sentences by PWA manifests itself more prominently in some sentence types than in others. Longer sentences, as counted by number of words or syntactic constituents, are more likely to be simplified or distorted, but there are also other factors that

play a role, pertaining to sentence structure. Typically, sentence types with noncanonical word orders are more prone to error than other sentence types. In English, these are sentences with word orders other than subject-verb-object. Such vulnerable sentences include object-extracted structures, such as object wh-questions, object clefts and object relatives, and passive sentences. Complexity in verb argument structure has also been shown to contribute to production and comprehension problems in PWA. This relates to the number of arguments a verb can take, the number of different thematic roles it can assign, and the number of subcategorization options it has, that is, the different syntactic frames with which the verb is compatible (for a discussion, see Malyutina, Richardson, & Den Ouden, 2016). Rehabilitation programs have therefore been developed to specifically target production and comprehension of these relatively “complex” sentence types and verb argument structures in speakers with aphasia (cf., Bastiaanse, Hurkmans, & Links, 2006; Bazzini et al., 2012; Schwartz, Saffran, Fink, Myers, & Martin, 1994; Thompson, 2001; Thompson, Shapiro, & Roberts, 1993).

Under the assumption that the difficulty with noncanonical sentence types reflects an impaired ability to perform specific linguistic construction operations, functional neuroimaging studies have targeted the neural substrates of this ability in unimpaired speakers. A few production studies exist (Den Ouden, Hoogduin, Stowe, & Bastiaanse, 2008; Grande et al., 2012; Indefrey et al., 2001; Indefrey, Hellwig, Herzog, Seitz, & Hagoort, 2004), but most neuroimaging studies in this area investigate sentence comprehension, rather than production. These studies tend to show involvement of left-hemisphere perisylvian regions in comprehension and production of structurally complex sentence types. The regions and the network(s) in which these regions participate show increased activation correlated with noncanonical sentence production or comprehension. Broca's area (Brodmann's Areas (BA) 44 and 45) in the inferior frontal gyrus (IFG) is generally held to play a prominent role in sentence-level production and comprehension. Its opercular part (BA44) has been argued to be central to hierarchical phrase structure building in comprehension (Friederici, 2017; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006), while resolution of syntactic ambiguity has been associated with a more anterior region (BA45/47, Tyler et al., 2011). By contrast, the association between Broca's area and sentence comprehension has also been argued to be driven by its function as a phonological short-term memory resource (Rogalsky & Hickok, 2011). With respect to verb argument structure, functional neuroimaging studies routinely implicate posterior superior temporal cortex and angular gyrus in the processing of more complex verbs, in both comprehension (Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Thompson, Bonakdarpour, & Fix, 2010) and verb retrieval (Den Ouden, Fix, Parrish, & Thompson, 2009).

Contemporary neurolinguistic models of language production are starting to converge on a dual-stream network architecture, in which syntactic structure-building and deconstruction is primarily supported by dorsal connections between inferior frontal cortex and posterior superior temporal cortex (Bornkessel-Schlesewsky & Schlesewsky, 2013; Den Ouden et al., 2012; Friederici, 2012), while the ventral route plays a more prominent role in lexical-level semantic contributions to sentence interpretation (e.g., Saur et al., 2008). This is

supported by studies of white matter atrophy in speakers with an agrammatic form of primary progressive aphasia (Wilson et al., 2011). Nevertheless, ventral pathways and anterior temporal cortex cannot be completely ruled out to play a role in syntactic processing either, based on studies of unimpaired syntactic processing (Den Ouden et al., 2012), as well as lesion-symptom mapping in aphasia (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Henseler, Regenbrecht, & Obrig, 2014; Ivanova et al., 2016; Magnusdottir et al., 2013) and primary progressive aphasia (Grossman et al., 2013).

Rather than comprising a homogeneous module, syntactic processing is a complex function of different components: morphosyntactic processing, phrase-structure building/deconstruction, and thematic role assignment. It also interacts with lexical-semantic and pragmatic processing. To different extents, deficits restricted to each of these components might result in an “agrammatic” performance pattern. It is therefore equally unlikely that a single homogeneous syndrome of agrammatism exists, as it is unlikely that there is a single specifically localized neural correlate of “syntactic processing.” Subcomponents of syntactic processing and agrammatic language performance, however, may still be identified, and this is relevant both to our understanding of the neurobiology of language and to our ability to optimize diagnostics and treatment approaches for speakers with aphasia.

1.2 | Lesion-symptom mapping of syntactic processing

Functional neuroimaging studies typically allow for inference about the *involvement* of regions in task execution, not about whether such regions are *critical* to the task. That question can be addressed by lesion studies, which assess the relation between structural brain damage and behavior. Additionally, such lesion-symptom-mapping studies allow for the investigation of how predictive certain patterns of brain injury are for diagnostics and prognosis of recovery.

Lesion-symptom-mapping studies have found damage to left-hemisphere anterior superior temporal gyrus, as well as superior temporal sulcus, middle temporal gyrus, angular gyrus, mid-frontal (BA 46), and IFG (BA 45; 47) to be associated with sentence comprehension difficulty in general, that is, irrespective of canonicity (Dronkers et al., 2004; Henseler et al., 2014; Pillay, Binder, Humphries, Gross, & Book, 2017; Tyler, Wright, Randall, Marslen-Wilson, & Stamatakis, 2010). In a previous study (Magnusdottir et al., 2013), we investigated the correlation of structural lesions with sentence comprehension deficits in a group of Icelandic stroke survivors ($n = 50$) and identified damage to left anterior superior and middle temporal gyrus to be predictive of patients having greater problems with processing noncanonical sentences than canonical sentences in a sentence-picture verification task. Specific problems with syntactically complex sentence processing have also been associated with damage to angular gyrus and temporoparietal cortex (Dickerson, Michaud, Hufford, & Caplan, 2013; Rogalsky et al., 2018; Thothathiri, Kimberg, & Schwartz, 2012), though observed individual differences as well as task-related effects suggest a level of fluidity in the neural substrates to syntactic processing (cf., Caplan, Michaud, Hufford, & Makris, 2016).

Fewer studies have investigated the neural correlates of sentence *production* deficits in aphasia. This may partly be due to the general

difficulty of assessing sentence production in a controlled manner, as that requires a procedure to elicit specific sentence types that do not naturally occur in aphasic speech. Lukic, Bonakdarpour, Den Ouden, Price, and Thompson (2013) used the *Northwestern Assessment of Verbs and Sentences* (NAVS; Cho-Reyes & Thompson, 2012) to investigate structural damage associated with sentence production problems. In a sample of 31 aphasic speakers, this study found damage to left-hemisphere IFG, superior and middle temporal gyrus, as well as the insula to be predictive of problems with primed sentence production. In a large study ($n = 102$) by Henseler et al. (2014), damage to middle and posterior superior temporal and inferior parietal cortex was also correlated with impaired “syntactic structure” in spontaneous speech, as scored gradiently according to criteria from the German *Aachener Aphasia Test*. In a recent contribution, Gleichgerrcht et al. (submitted) used natural language processing to quantify the grammatical quality of picture descriptions in a sample of 65 stroke survivors and 10 unimpaired control speakers. Measures of syntactic degradation were associated with structural lesions in posterior IFG and inferior parietal cortex, as well as structural-connectivity damage to dorsal pathways between these regions, again suggesting a primary role for the dorsal stream in syntactic production.

1.3 | Current study

The present study investigates structural gray matter as well as white matter predictors not only of general performance on tasks with a sentence-processing component, but also of specific problems with noncanonical sentence production and comprehension (relative to difficulty with canonical sentences), as well as complex verb production (relative to production of verbs with less complex argument structure), using partial correlations to covary out nuisance variables. We make use of the power offered by a large sample of stroke survivors to conduct voxel-based lesion-symptom mapping (VLSM; Bates et al., 2003) to assess the predictive value of local structural brain damage for performance on tasks that require different forms of syntactic computation. These tasks draw on construction (in production) and deconstruction (in comprehension) of sentences that may be complex in terms of their word order and/or the argument structure of their verbs.

In addition, we have included the more recent method of connectome-based lesion-symptom mapping (CLSM; see Gleichgerrcht, Fridriksson, Rorden, & Bonilha, 2017), using structural connectivity matrices generated from diffusion tensor imaging along prespecified tracts in a white matter atlas. It is increasingly acknowledged that damage to specific connections between cortical areas may be just as predictive of cognitive symptoms as damage to the cortical regions themselves. Impaired connectivity, however, may occur due to anatomically distinct lesions along a pathway in different individuals and is therefore typically underestimated in analyses that focus only on regional structural damage, whether to gray or white matter. We examined correlations between our behavioral measures and the structural connectivity in a matrix of perisylvian regions shown to be involved in language processing and part of a dual-stream architecture (Bornkessel-Schlesewsky & Schlesewsky, 2013; Fridriksson et al., 2016; Friederici, 2012; Hickok & Poeppel, 2007).

For these VLSM and CLSM analyses, we hypothesize that performance on different NAVS subtests, focusing on verb argument structure (*Argument Structure Production Test* [ASPT]), sentence construction (*Sentence Production Priming Test* [SPPT]), and sentence deconstruction (*Sentence Comprehension Test* [SCT]), will be associated with structural integrity of separate brain regions, as we do not presume the existence of a unique neural correlate to underlie such a vastly complex function as “syntactic processing.” We also expect regions associated with general performance on these tests to be distributed rather than focal, as low overall task performance may have many different origins, ranging from conceptual, lexical, and semantic issues to articulatory problems. Nevertheless, in line with current neurolinguistics models, we expect performance on the ASPT to load primarily onto posterior superior temporal and temporoparietal cortex, the SPPT to load primarily onto areas along the dorsal pathway between IFG and posterior superior temporal gyrus (pSTG), and the SCT to show both ventral and dorsal correlates, between these same regions. We hypothesize that overlap between these predictive lesion patterns is located along the dorsal route between IFG and pSTG. We will also directly contrast aspects of sentence and argument structure that are syntactically more versus less complex, and we expect lesion predictors of scores on these derived measures to be more focal. By their very nature, these measures should capture specific syntactic processing abilities more focally, canceling variance in the confounds that are not directly related to structural complexity in sentence production and processing (e.g., hearing ability, word retrieval/comprehension, attention span, etc.). Alternatively, an absence of neural correlates of our derived measures may reflect the absence of a sufficiently homogeneous functional process that underlies “syntactic” ability in aphasia (as argued in Caplan, Hildebrandt, and Makris (1996)).

Whereas our main interest in the current paper was in investigating structural and connectivity damage associated with specific problems in production and/or comprehension of noncanonical sentences and verb argument structure, a question that always looms over this field is whether the notion of “agrammatic aphasia” should be considered a homogeneous and specific syndrome (Caplan, 1986, 1991; Grodzinsky, 1984), particularly given the widely reported variability in patterns of syntactic impairment between and even within individual stroke survivors (Berndt, Mitchum, & Haendiges, 1996; Caplan, 1987; Goodglass, 1997; Kean, 1985; Kolk & Van Grunsven, 1985; Miceli, Mazzucchi, Menn, & Goodglass, 1983). Alternatively, agrammatism might be used merely as a descriptive term for certain language symptoms, without hinging on the assumption that there is a consistent “syndrome,” let alone a specific neural correlate to the basic processing ability that is impaired in such a syndrome (Badecker & Caramazza, 1985). The reason why this is of interest is because it could have implications for how agrammatic language behavior in aphasia is approached and ultimately treated, but also because it relates to a debate about the autonomy of syntax, which has been waged based on arguments and data from formal linguistics, functional neuroimaging, and lesion studies.

For the sake of argument, then, we included a categorization of patients who show an “agrammatic” output pattern, strictly defined as

reduced morphosyntactic production (Goodglass, 1973; Goodglass & Berko, 1960) and based on clinical judgment of descriptive-speech samples. This binary categorization was compared to a hierarchical clustering analysis of patients based on their NAVS subtest performance. As noted above, the classic definition of agrammatism referred not to complex sentence structures or verb argument structure, but to morphosyntactic production, that is, omission of inflections and function words (e.g., Goodglass & Berko, 1960). The criteria we used to classify patients as having a morphosyntactic production problem are therefore intentionally different from the behaviors evaluated by the NAVS, so that it is of interest to what extent these components to syntactic processing pattern together. The outputs of both analyses were also subject to VLSM and CLSM. The NAVS is intended as an evaluation tool complementary to other diagnostic measures, to identify speakers with deficits in syntactic production and/or comprehension. It is important that it provides meaningful qualitative information about patients' performance, above and beyond what can be concluded from standard diagnostic tests and the general impact of stroke. We therefore start our analysis with an investigation of correlations between NAVS subtests, and between these tests and general aphasia severity and lesion size.

2 | MATERIALS AND METHODS

2.1 | Participants

Participants were 71 right-handed native speakers of English (mean age 59.6 years, *SD* 10.1, range 37–80 years; 23 females) who had suffered a left-hemisphere ischemic stroke (63 single and 8 multiple strokes; mean number of months since first stroke 52.8, *SD* 52.5, range 8–290 months; mean lesion size 93.5 cc, *SD* 86.0 cc, range 0.16–365.4 cc). Forty-nine participants were diagnosed with aphasia, based on the *Western Aphasia Battery-Revised* (WAB-R; Kertesz, 2007: mean WAB-R aphasia quotient [AQ] 65.8, *SD* 21.8, range 20.1–93.8), and 22 stroke survivors were not aphasic according to the WAB-R (AQ > 93.8). Three of the participants were only included in the behavioral analyses, but not in the VLSM and CLSM analyses: one person had a lesion in the right hemisphere, one person had a lesion that was restricted to the cerebellum, and one person only had a CT scan taken, limiting the accuracy of the lesion map. Figure 1 provides a lesion overlay map of the 68 patients included in the LSM analyses. Distribution of aphasia classifications and further patient variables as recorded at the time of testing is provided in Table 1. Participants had no premorbid history of other neurological or psychiatric disorders, and all had normal or corrected to normal vision and hearing. The study was approved by the Institutional Review Boards at the University of South Carolina and at the Medical University of South Carolina and all participants signed informed consent.

2.2 | Behavioral data

2.2.1 | Northwestern Assessment of Verbs and Sentences

Participants were tested on three tasks from the NAVS (Cho-Reyes & Thompson, 2012): the ASPT, the SPPT and the SCT. All tests were

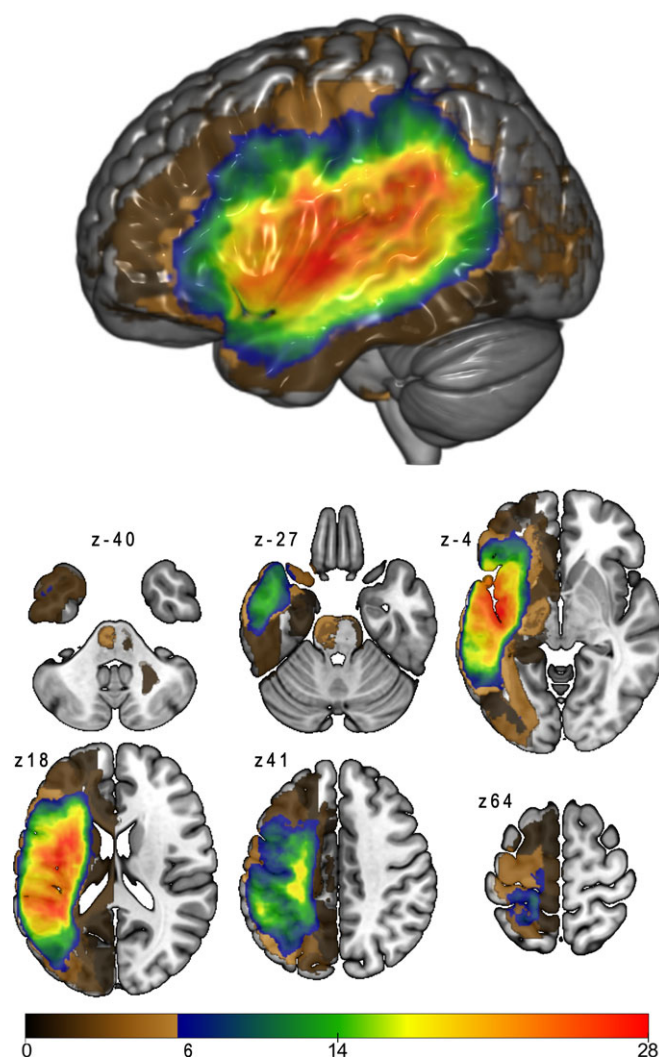


FIGURE 1 Lesion overlap map, for the participants included in lesion-symptom mapping analyses ($n = 68$). Brown-shaded areas show lesion overlap for fewer than six participants; these regions were not included in lesion-symptom mapping analyses [Color figure can be viewed at wileyonlinelibrary.com]

administered during the same session. All participants completed at least one subtest of the NAVS, but three participants were unable to complete the ASPT and six participants were unable to complete the SPPT. These participants' inability to complete the production tasks was not due to time constraints but due to the difficulty they experienced with these tasks, yielding no scorable responses. Their data were treated as "missing data" for these subtests. The NAVS also contains an action naming and a verb comprehension task, but our focus here was on sentence-level performance and most patients were not given these tasks, to reduce testing time.

For detailed procedure instructions and scoring criteria, we refer to the NAVS manual. Scoring was performed by two trained research assistants, with daily supervision by author S.M., and any uncertainties were discussed and resolved in weekly group meetings that included two authors of the paper, D.-B.d.O. and S.M. All resolutions were recorded in a logbook, so that new responses that were similar to previously scored responses could be evaluated in the same spirit. Finally, to ensure consistency, all NAVS scores were submitted to a check by author D.-B.d.O., who had 10 years of experience with NAVS testing

TABLE 1 Participant characteristics by WAB-R aphasia classification

	N	Gender (% female)	Age, years (SD)	Months P.O. (SD)	WAB-R AQ (SD)
Anomic	15	26.7%	59.3 (9.7)	43.5 (50.7)	89.1 (4.5)
Broca's	18	27.8%	57.7 (10.3)	68.6 (73.0)	54.5 (17.1)
Conduction	11	36.4%	64.9 (8.2)	44.2 (40.9)	65.3 (15.2)
Global	2	0%	52 (11.3)	57 (11.0)	24.5 (1.2)
Transcortical sensory	1	0%	60 (na)	22.1 (na)	57.8 (na)
Wernicke's	2	0%	63 (4.2)	61.2 (12.9)	42.0 (15.2)
None	22	45.5%	59 (11.1)	50.8 (44.3)	97.9 (1.6)
Total	71	32.4%	59.6 (10.1)	52.8 (52.5)	75.8 (23.5)

Months P.O. = months postonset; WAB-R AQ = Western Aphasia Battery-Revised, aphasia quotient.

and scoring at the time of the study. Cho-Reyes and Thompson (2012) report a point-to-point agreement between two independent raters' NAVS scores ranging from 97 to 100%, with overall agreement of 99.8%.

In the ASPT, subjects were presented with a line drawing of an intransitive or transitive action, including the written names of the participants in the action (the arguments) and the action word (the verb). Their task was to produce a full sentence describing the picture that included all the target words. For example, a picture might depict a boy tickling a girl, with the arguments identified overtly by the experimenter as *boy* and *girl*, and the action given as *tickle*. The participant would then have to respond: "*The boy is tickling the girl.*" Test items included one-place verbs (intransitive: *The man is laughing*), two-place verbs (transitive: *The boy is tickling the girl*), and three-place verbs (ditransitive: *The woman reads a book to the girl*). For the ASPT-Code A (ASPT-A) score, responses were correct if the verb as well as all required arguments were present, in the correct order. Phonemic paraphasias did not influence the score, as long as the correct target was identifiable. The ASPT is specifically intended to assess applied knowledge of verb argument structure (identifying the roles that the verb assigns to the various arguments), so that absence of required prepositions or articles, errors in verb inflection, and other morphological errors did not affect the score. For this reason, we added a custom category to the scoring, ASPT-Code G (ASPT-G), which was only given in case the response sentence was grammatically well-formed, including appropriate inflection and morphology, but still not penalizing phonemic paraphasias. In addition, the ASPT-Code W (ASPT-W) score represents the number of word-level errors (phonemic, verbal, and semantic paraphasias) produced, as a measure of lexical output problems outside of the syntax domain. The participants' ASPT-W scores were included in the hierarchical clustering analysis (see below).

In the SPPT, subjects were cued with an example sentence describing a picture and asked to produce a sentence of the same grammatical structure describing another picture with reversed roles. For example, participants had to produce *The boy is tickling the girl* when cued with *The girl is tickling the boy*. Sentence types included three canonical types (active sentences, *The boy is pulling the girl*; subject-extracted Wh-questions, *Who is kissing the man?*; and subject relatives, *Pete saw the man who is saving the woman*) and three noncanonical sentence types (passive sentences, *The cat is chased by the dog*; object-extracted Wh-questions, *Who is the woman kissing?*; and object relatives, *Pete saw the girl who the boy is pulling*). Responses were scored as correct

if they were (a) grammatically correct sentences of the target grammatical structure (so, including verb inflections and correctly placed prepositions, articles, and wh-words) and (b) described the picture appropriately (i.e., the verb arguments were assigned the correct roles).

In the SCT, participants had to select a picture, out of two presented options, that matched an auditorily presented sentence. The incorrect picture portrayed the same situation with reversed roles (e.g., the incorrect picture for *The boy is tickling the girl* portrayed a girl tickling the boy). The SCT used the same six sentence types as the SPPT. Responses were scored as correct if participants pointed to the correct picture. The values used for the ASPT, SPPT, and SCT scores in all analyses were the proportions of correct responses.

Total scores on any of these NAVS subtests can be affected by a range of functional deficits that are not necessarily syntactic in nature. Motor speech problems, impaired lexical processing/representations, problems with auditory or visual processing, reduced general working memory, as well as reduced resources for attentional control may all contribute to lower overall scores on individual subtests. Therefore, to define relative difficulties with complex syntactic structures, rather than absolute scores of task performance, we calculated differences in performance between more and less complex item types. For the ASPT, we calculated a difference score between proportion correct ASPT-A scores for one-place versus two-place verb stimuli (ASPT-2-1), as well as on one-place versus three-place verb stimuli (ASPT-3-1). For both the SPPT and the SCT, we calculated the scores (proportion correct) on noncanonical and canonical sentence types in addition to overall test scores. To characterize relative difficulty for a participant on noncanonical sentences, we subtracted canonical from noncanonical sentence scores, so that equal performance would yield a value close to zero, whereas a positive value would indicate relatively better performance on noncanonical than on canonical sentence types and a negative value would indicate the reverse pattern. This yielded the SPPT-noncanonical minus canonical (NC-C) score and the SCT-NC-C score, respectively, that we used for our behavioral classification analysis. For all these derived difference scores, lower negative scores indicated a greater vulnerability to error for more complex syntactic structures.

A downside of using difference scores is that they equalize speakers who have a deficit that is specific to one structure type (e.g., canonical error rates at 90% vs. noncanonical error rates at 50%) with speakers who have difficulty across the board, but simply

relatively more severe for one structure type (e.g., canonical at 55% vs. noncanonical at 15%; see Rogalsky et al., 2018). Our perspective, however, is that the general severity of aphasia may impact various components that affect sentence processing, leading to generally lower scores also on canonical structures, whereas a disproportionate problem with noncanonical structures is indicative of the syntactic processing difficulties we aim to target here. For inclusion of multiple behavioral variables in the same correlation and classification analyses, therefore, we used the single-value difference scores. These scores are interpretable and lend themselves well to the hierarchical clustering analysis we applied (see below). In our VLSM and CLSM analyses, we did not use the differences scores, but instead regressed out performance on canonical structures from scores on noncanonical structures to identify damage uniquely predictive of problems with noncanonical sentences (see Section 2.4.4). Table 2 provides a list of the variables obtained from the NAVS, with brief descriptions.

2.2.2 | Agrammatism

In addition to the objective continuous measures of production and comprehension of certain sentence types on the ASPT, SPPT, and SCT, we created a binary distinction between patients we labeled as showing “agrammatic” speech output and those that did not. This distinction was based on a clinical judgment, much like in informal reports in clinical practice. In our assessment, we focused on morphosyntactic errors, as these are the classic hallmark of agrammatic output, whereas word order and phrase structure problems have been shown to dissociate from the morphosyntactic error pattern, which is more consistent (Goodglass, 1973; Goodglass & Berko, 1960; Saffran, Berndt, & Schwartz, 1989). Whereas the NAVS might actually be used to aid in informal clinical diagnosis of “agrammatism,” our classification here was based solely on the speakers' description of the Cookie-Theft picture that is part of the WAB-R, in order to achieve a noncircular comparison between NAVS outcomes and speech output characteristics. Two authors (D.-B.d.O. and A.B.) independently categorized the patients' speech output according to the following criteria:

1. Agrammatic pattern of morphosyntactic reduction (PD-MR): Omission and/or substitution of grammatical morphemes (verb inflection [tense and agreement errors]; plural markers), articles, and prepositions.
2. No morphosyntactically reduced speech in picture description (PD-Other): These patients provide sufficient output to warrant classification of their morphosyntactic production. They may have other production problems, such as anomia, phonemic or semantic paraphasias, but their morphosyntactic production is not specifically affected.
3. Cannot be categorized (PD-NC): Too little output, for example, because of too severe apraxia of speech, but also in case severe anomia does not yield sufficient data to establish agrammatism.

This scoring method allowed us to exclude patients for which the presence or absence of agrammatism could not be reasonably established (PD-NC) from the classification and lesion-symptom mapping

TABLE 2 Variables obtained from the NAVS

Abbreviation	Test	Summary
ASPT-A	Argument Structure Production Test, Code A	Verb and its arguments produced in the correct order on a picture description task.
ASPT-G	Argument Structure Production Test, Code G	Verb and its arguments produced in the correct order on a picture description task, plus the full sentence is grammatically and morphosyntactically correct
ASPT-W	Argument Structure Production Test, Code W	Absence of word-level errors (phonemic, verbal, and semantic paraphasias) in picture description
ASPT-2-1	Argument Structure Production Test, 2-1	Proportion correct ASPT-A scores for two-place minus one-place verb stimuli. A negative number indicates relatively greater difficulty with two-place verbs than with one-place verbs.
ASPT-3-1	Argument Structure Production Test, 3-1	Proportion correct ASPT-A scores for three-place minus one-place verb stimuli. A negative number indicates relatively greater difficulty with three-place verbs than with one-place verbs.
SPPT	Sentence Production Priming Test	Correctly produced target structures on a picture description task in which reversible sentences are used to prime the elicitation of specific sentence types
SCT	Sentence Comprehension Test	Correct scores on an auditory-sentence-picture matching task in which reversible sentences are used to provide binary choices
NC-C	Noncanonical minus canonical	Proportion correct scores on noncanonical structures minus canonical structures, on the SPPT or SCT. A negative number indicates relatively greater difficulty with noncanonical than with canonical sentence structures.

NAVS = Northwestern Assessment of Verbs and Sentences.

analyses. In total, the two raters agreed on the ratings for 59/71 samples (83% agreement ratio), and the diverging classifications were resolved by mutual agreement. Ultimately, 12 speakers in the group were classified as showing a morphosyntactic deficit in production (PD-MR), 52 speakers in the group as not showing a morphosyntactic deficit in production (PD-Other), and seven were deemed to be not classifiable (PD-NC), based on descriptive speech output alone.

As an external validation of our subjective grouping based on an impressionistic analysis, Table S1 (Supporting Information) shows how the PD-MR and PD-Other groups turned out to differ quantitatively on diagnostic measures as well as linguistic content of their picture descriptions. The PD-MR group had more severe aphasia and apraxia of speech ($p < 0.05$), but was not different from the PD-Other group in severity of dysarthria ($p = 0.862$). The PD-MR group also scored lower

on measures of fluency (words-per-minute) and utterance length and on almost all measures of grammatical content, such as the produced number of auxiliaries, prepositions, and conjunctions. Notable exceptions were the type/token ratio ($p = 0.123$) and number of verbs per total words ($p = 0.208$). The noun/verb ratio was indeed lower in the PD-MR group ($p < 0.05$), but that was driven by a higher number of nouns per total words produced ($p < 0.05$), as opposed to a lower number of verbs. These more detailed analyses of aphasic discourse will be the subject matter of forthcoming work from our group, but are presented here primarily to support the classification we adopted for this study.

2.3 | Neuroimaging data

2.3.1 | Magnetic resonance imaging data acquisition and preprocessing

For each participant, several types of brain images were obtained with a 3.0T Siemens Tim Trio magnetic resonance imaging (MRI) scanner equipped with a 12-channel head coil: T1-weighted anatomical scan (TR = 2,250 ms, TE = 4.15 ms, matrix = 256×256 , field-of-view (FOV) = $256 \times 256 \text{ mm}^2$, slice thickness = 1 mm, 192 sagittal slices), T2-weighted anatomical scan (TR = 3,200 ms, TE = 212 ms, matrix = 256×256 , FOV = $256 \times 256 \text{ mm}^2$, slice thickness = 1 mm, 160 sagittal slices), diffusion tensor imaging scan (two scans with a 180° flip, TR = 4,987 ms, TE = 79.2 ms, matrix = 90×90 , FOV = $207 \times 207 \text{ mm}^2$, slice thickness = 2.3 mm, 50 transversal slices). All images were converted to NIFTI format using *dcm2nii* (Li, Morgan, Ashburner, Smith, & Rorden, 2016).

2.3.2 | Lesion maps

Lesions were manually delineated on the T2 images by one of the authors (L.B.), a neurologist blinded to the behavioral data. Spatial normalization of the lesions to standard space was performed using SPM12 and open source MATLAB scripts developed in-house (Rorden, Bonilha, Fridriksson, Bender, & Karnath, 2012): the patient's T2 image was linearly coregistered with their T1 image and the same transformation was applied to the lesion mask. After coregistration and reslicing, the lesion map was smoothed with a 3 mm full width at half maximum Gaussian kernel. Prior to normalization of the T1 image, a new image was created in which the lesioned area was replaced by its (mirrored) equivalent intact region taken from the contralateral right hemisphere. This new T1 image was then normalized through SPM12's unified segmentation-normalization (Ashburner & Friston, 2005) and the resulting nonlinear transformation parameters were also applied to place the lesion mask in the same standard space. Finally, we re-binarized the smoothed standardized lesion mask by thresholding at a signal intensity of 0.5, which provides a good balance between the preservation of lesion volume and shape, relative to the binary lesion mask prior to smoothing.

2.3.3 | Structural brain connectivity

Individual structural white matter whole-brain connectomes were constructed based on probabilistic diffusion tensor imaging. Probabilistic gray and white matter maps were obtained during the unified segmentation-normalization of the T1 image, described above. The

probabilistic gray matter map was segmented into regions of interest (ROIs) based on the John's Hopkins University (JHU) atlas (Faria et al., 2012; Mori et al., 2008), which parcellates gray matter into 189 regions. All connectome preprocessing steps were performed in diffusion space, so the standard-space connectome template was first warped to each individual's native intensity-normalized diffusion-based fractional anisotropy image, and the same transformation parameters were used to normalize the lesion mask and the probabilistic maps to the individual's diffusion space. In our analysis, we zoomed in on potentially dorsal versus ventral connections that might support syntactic processing by conducting a CLSM that only included connections between left-hemisphere perisylvian nodes that we deemed to be potentially involved in language processing, in a broad sense (see below). The 17 nodes we selected from the JHU atlas are listed in Table 3. As for the whole-brain connectome, we assumed potential full connectivity between these nodes, for a total of 136 unique connections.

Pair-wise connectivity between all 17 gray matter ROIs was obtained through probabilistic tractography, using FMRIB's Diffusion Toolbox (FDT) in FMRIB's Software Library (FSL) (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007), with 5,000 individual pathways drawn through the probability distributions on principle fiber direction, curvature threshold set at 0.2, 200 maximum steps, step length 0.5 mm and distance correction (see Yourganov, Fridriksson, Rorden, Gleichgerrcht, & Bonilha, 2016). Each individual's probabilistic white matter map excluding the stroke lesion was used as a waypoint mask for the estimation of their probabilistic tractography. For each pair of ROIs, the number of streamlines arriving in one ROI when another ROI was seeded was computed, and the connectivity was defined as the average between (a) the number of streamlines arriving in ROI_j when ROI_i was seeded, and (b) the number of streamlines arriving in ROI_i when ROI_j was seeded. The connectivity between two ROIs was corrected based on the distance traveled by the streamlines and further divided by the sum of the volumes of the two ROIs, resulting in a 17×17 adjacency matrix of weighted connections (with a total of 136 unique values of connectivity). Connections to ROIs within the lesion site were not tracked and given a weight of zero in this matrix.

Based on the dual-stream model for speech processing proposed by Hickok and Poeppel (2007) and the anatomical identification of these streams outlined in Fridriksson et al. (2016), we labeled these nodes a priori as being more likely to be part of the dorsal versus the ventral stream, or vice versa. Also, given their primary roles in syntactic processing and the models proposed by Friederici (2012) and Bornkessel-Schlesewsky and Schlesewsky (2013), we treat the triangular and opercular parts of the IFG on the one hand and the pSTG on the other hand as transitional between these streams. As noted by Fridriksson et al. (2016), there is no doubt that structural connections exist between regions in IFG and anterior temporal lobe, but the functional connectivity of both the triangular and the opercular parts of IFG is primarily dorsal, whereas the orbital part of IFG may be more involved with the ventral stream. In addition, because of its structural connectivity to anterior (dorsal) as well as ventral regions (Uddin et al., 2010), we also treat angular gyrus as transitional between the two streams. In Fridriksson et al. (2016), 55% of voxels in the angular gyrus patterned with other ventral-stream regions, but it must be noted that the principal-component analysis applied in this study was not

TABLE 3 Nodes along dorsal and ventral routes connecting IFG and posterior temporal/inferior parietal cortex

JHU atlas #	ROI/node	Abbreviation	Dorsal/ventral
29	Supramarginal gyrus left	SMG	Dorsal
71	Insular left	Ins	Dorsal
182	Posterior insula left	pIns	Dorsal
11	IFG pars opercularis left	IFG_op	Dorsal/ventral ^a
15	IFG pars triangularis left	IFG_tri	Dorsal/ventral ^a
184	pSTG left	pSTG	Dorsal/ventral ^a
31	Angular gyrus left	AG	Dorsal/ventral ^a
13	IFG pars orbitalis left	IFG_orb	Ventral
17	Lateral fronto-orbital gyrus left	LFOG	Ventral
19	Middle fronto-orbital gyrus left	MFOG	Ventral
35	Superior temporal gyrus left	STG	Ventral
37	Pole of superior temporal gyrus left	STG_pole	Ventral
39	Middle temporal gyrus left	MTG	Ventral
41	Pole of middle temporal gyrus left	MTG_pole	Ventral
43	Inferior temporal gyrus left	ITG	Ventral
186	Posterior middle temporal gyrus left	pMTG	Ventral
188	Posterior inferior temporal gyrus left	pITG	Ventral

IFG = inferior frontal gyrus; JHU = John's Hopkins University; ROI = region of interest.

^a Nodes that may have a transitional function between dorsal and ventral streams.

equipped to identify transitional voxels themselves (voxels that are involved with both streams), as the lesion data were necessarily split to load onto separate behavioral components.

2.4 | Statistical analyses

2.4.1 | Correlations

We calculated Pearson's multiple correlations between the NAVS subtests, including the derived complexity measures, WAB-R AQ (as a measure of overall aphasia severity), and lesion size. For a corrected α -level of 0.05, the Bonferroni correction for multiple comparisons set the acceptable probability value at 0.0009.

2.4.2 | Hierarchical clustering based on NAVS scores

To assess to what extent performance on the NAVS subtests, including on derived measures, might reveal patterns of homogeneous behavior in our participant group, we used hierarchical clustering (SPSS, version 24), based on the Squared Euclidian Distance interval, with all variables scaled to a 0–1 range. Hierarchical clustering proceeds through an algorithm that starts with each case in a separate cluster and merges clusters that are located near each other until only one is left.

2.4.3 | Linear discriminant analysis

In order to establish which, if any, of the NAVS subscores are predictive of group membership based on (a) our descriptive-speech categorization and (b) the hierarchical clustering, we conducted stepwise linear discriminant analyses (LDAs, SPSS v. 24), with leave-one-out cross-validation. Variables were entered when they minimized Wilk's lambda at an F value with $p < 0.05$, and removed at $p > 0.10$. Naturally, the second analysis has an element of circularity, but the goal here was to establish which variables contribute primarily and independently to the clustering of participants by NAVS scores. The variables entered in these analyses were the following scores: ASPT-A, ASPT-W, SPPT, SCT, ASPT-2-1, ASPT-3-1, SPPT-NC-C, and SCT-NC-C. We also included WAB-AQ, as it is important to establish whether the NAVS contributes to patient categorization beyond what might be derived from general aphasia severity. Our principal analyses did not include the ASPT-G measure, as this is not part of the original NAVS coding system, but we conducted secondary LDAs which did include ASPT-G, for comparison. The LDA's were supplemented with independent samples t tests, to assess direction and significance of differences in NAVS scores between the groups.

2.4.4 | Lesion-symptom mapping

Analyses of brain-behavior correlations were conducted with in-house *NiiStat* software (<https://github.com/neurolabusc/NiiStat>), running on a MATLAB platform. Voxel-wise lesion-symptom mapping (VLSM; Bates et al., 2003) analyses were run on (a) the two binary classifications (based on hierarchical clustering [NAVS] and our own categorization of morphosyntactic production deficits [PD-MR vs. PD-Other]), as well as on the following behavioral variables: (b) overall ASPT-A, ASPT-G, SPPT, and SCT scores; (3a) scores on three-place and two-place verb production on the ASPT, with scores on one-place verbs partialled out as a nuisance regressor, using the Freedman-Lane procedure (Winkler, Ridgway, Webster, Smith, & Nichols, 2014) incorporated in *NiiStat*; (3b) scores on noncanonical sentences, with scores on canonical scores partialled out as a nuisance regressor on the SPPT and the SCT. We first performed these analyses on the raw test scores, in order to get a rough idea of regional damage that is predictive of behavior on our measures of interest. However, because of expected high correlations between impairment severity and lesion size, for our primary analyses we regressed out overall lesion size from the binary classifications in Analysis 1 (using binary logistic regression), as well as from the overall NAVS scores used in Analysis 2 (using linear regression), before performing the VLSM analyses.

Only voxels in which six or more participants had lesions were included in the analyses (see Figure 1), which represented about 10% of the number of participants included in each analysis. The number of participants included in the VLSM analyses varied between 62 and 68, due to some participants' inability to complete certain NAVS subtests. For each relevant voxel in the template whole-brain map, a t test was performed to compare the average scores in subjects with a lesion involving that voxel versus those whose lesions did not involve that voxel. All analyses were corrected for multiple comparisons through permutation testing (5,000 permutations). Analyses

were one-tailed (since injured tissue is predicted to only cause poorer, not better, performance), with $\alpha = 0.05$.

2.4.5 | Connectivity analyses

The procedure and behavioral tests used for CLSM were very similar to those described above for VLSM, except that the relation between the integrity of connectivity and behavioral scores was not assessed with *t* tests, but rather with correlations between individuals' test scores and their pair-wise connectivity weights, for all connections in the adjacency matrix. As for VLSM, all analyses were univariate and corrected for multiple comparisons through permutation testing (5,000 permutations). Analyses were one-tailed (since injured tissue is predicted to only cause poorer, not better, performance), with $\alpha = 0.05$.

For the connectivity analyses, the regression of overall lesion size, as well as the regression of lesion size restricted to white matter from the behavioral data, did not leave results that survived the statistical threshold, so we only report the results based on the "raw" behavioral data. At this point, it is not yet clear whether the regression of the complete lesion size in CLSM analyses may be a form of overreach, and our group is in the process of establishing standards that allow an optimal balance between specificity and sensitivity. Our approach here does leave the possibility that results are confounded with effects of general aphasia severity, but since this is a relatively novel approach to these data, we considered it important to report the analyses and results, after all.

3 | RESULTS

3.1 | Multiple correlations

Group scores on NAVS subtests, including scores per sentence and verb type, are provided in Table 4. Paired *t* tests showed general effects of sentence canonicity for the SPPT task ($t(64) = 5.205, p < 0.001$) and the SCT task ($t(70) = 3.651, p < 0.001$), with significantly lower scores for production and comprehension of noncanonical structures. On the ASPT task, scores of "A" were significantly lower for three-place verbs than for two-place verbs ($t(67) = 5.662, p < 0.001$) and one-place verbs ($t(67) = 5.861, p < 0.001$), without a difference between one- and two-place verbs ($t(67) = 1.203, p = 0.233$).

Table 5 provides the Pearson's correlation coefficients for the multiple correlations between NAVS scores, WAB-R AQ, and lesion size. Absolute scores on the three NAVS subtests (1–5) are all highly correlated with one another, as well as with aphasia severity and lesion size (in number of 1 mm^3 voxels). Not surprisingly, aphasia severity and lesion size themselves are also highly correlated; the larger the lesion, the more severe is the language deficit. The derived measures of grammatical impairment (6–9) are not correlated with general aphasia severity, or with absolute lesion size. However, a more severe impairment on three-place verbs than on one-place verbs is correlated with the same pattern for two-place verbs, as well as with the ASPT-G measure (grammatically correct production on the ASPT) and with the total score on the SPPT.

TABLE 4 NAVS performance

	N	Range (%)	Mean (%)	SD (%)
ASPT-A	68	0–100	73.7	36.7
ASPT-G	68	0–100	63.6	41.0
ASPT-W	68	0–100	82.5	29.4
SPPT total	65	0–100	63.6	39.1
SCT total	71	13–100	82.7	18.2
ASPT-A 1 arg. Verbs	68	0–100	80.3	35.2
ASPT-A 2 arg. Verbs	68	0–100	78.6	36.1
ASPT-A 3 arg. Verbs	68	0–100	68.9	37.7
SPPT canonical	65	0–100	68.2	39.4
SPPT noncanonical	65	0–100	58.9	39.8
SCT canonical	71	40–100	86.8	17.2
SCT noncanonical	71	20–100	79.8	19.4

ASPT-A = Argument Structure Production Test, Code A; ASPT-G = Argument Structure Production Test; Code G; ASPT-W = Argument Structure Production Test, Code W; ASPT-2-1 = Argument Structure Production Test, 2-1; ASPT-3-1 = Argument Structure Production Test, 3-1; NAVS = Northwestern Assessment of Verbs and Sentences; SCT = Sentence Comprehension Test; SPPT = Sentence Production Priming Test.

3.2 | Hierarchical clustering and reduced morphosyntactic production

Based on the NAVS scores, participants were clustered into two groups, of 48 (Group A) versus 17 (Group B) members. Six participants were automatically excluded from this analysis, as they had too few NAVS subscores to aid in the classification. Group B contained 8 of the 12 participants categorized as showing morphosyntactically reduced output in descriptive speech, as well as five of the seven participants whose output could not be categorized and 4 of the 52 participants categorized as not having a morphosyntactic output problem (see Table 6). The distribution of participants in the three different categories (PD-MR, PD-Other, and PD-NC) over the two clusters was unequal ($\chi^2(2) = 37.662; p < 0.001$, with Yate's correction), and the same was true if the uncategorized participants were excluded ($\chi^2(1) = 22.688; p < 0.001$, with Yate's correction). Therefore, the hierarchical clustering appears to split the patients nonrandomly into those that pattern with the PD-MR group (Group B), and those that do not (Group A). Henceforth, Group B is labeled here as "NAVS-Low," by which we aim to capture the fact that these patients show a specific pattern of impaired performance on NAVS subtests (see Section 3.3), without suggesting they should necessarily be labeled as "agrammatic." Group A is henceforth labeled as NAVS-Other.

3.3 | NAVS predictors of agrammatic speech output

The first binary LDA on group membership revealed that morphosyntactically reduced speech (i.e., membership of the PD-MR group) was predicted with 86.7% success, using the variables SPPT, SPPT-NC-C, and SCT-NC-C. Six participants were excluded from this analysis, due to missing discriminant variables. Group assignment was correct for 6/10 participants with morphosyntactic output problems (PD-MR), and for 46/50 participants who did not show reduced morphosyntax (PD-Other), so sensitivity was 60% and specificity was 92%. Adding the ASPT-G measure to the input variables did not substantially

TABLE 5 Pearson's multiple correlations table between NAVS measures, aphasia severity, and lesion size

	1	2	3	4	5	6	7	8	9	10
1 ASPT-A										
2 ASPT-G	0.810**									
3 ASPT-W	0.781**	0.707**								
4 SPPT total	0.750**	0.882**	0.734**							
5 SCT total	0.579**	0.627**	0.458**	0.760**						
6 ASPT-2-1	0.142	0.217	0.152	0.101	-0.096					
7 ASPT-3-1	0.147	0.385*	0.101	0.461**	0.197	0.508**				
8 SPPT-NC-C	-0.090	-0.097	-0.224	0.033	0.076	-0.208	-0.056			
9 SCT-NC-C	0.060	0.154	0.253*	0.187	0.113	0.037	0.095	0.273*		
10 WAB-R AQ	0.826**	0.856**	0.804**	0.889**	0.706**	0.023	0.217	-0.015	0.116	
11 Lesion size	-0.550**	-0.716**	-0.579**	-0.696**	-0.542**	-0.026	-0.237	-0.046	-0.135	-0.742**

ASPT-A = Argument Structure Production Test, Code A; ASPT-G = Argument Structure Production Test, Code G; ASPT-W = Argument Structure Production Test, Code W; ASPT-2-1 = Argument Structure Production Test, 2-1; ASPT-3-1 = Argument Structure Production Test, 3-1; NAVS = Northwestern Assessment of Verbs and Sentences; NC-C = Noncanonical minus canonical; SCT = Sentence Comprehension Test; SPPT = Sentence Production Priming Test; WAB-R = Western Aphasia Battery-Revised, aphasia quotient.

*Correlation is significant at the 0.05 level (uncorrected; two-tailed); n range = 64-71 (not all participants were able to complete all NAVS tasks).

**Correlation is significant at the 0.0009 level ($p < 0.05$ with Bonferroni correction; two-tailed).

improve classification (87.3% success), but the independently predictive variables were now identified as ASPT-G and SCT-NC-C, and only three participants had to be excluded. Together these variables correctly categorized 9/12 PD-MR participants (sensitivity 75%) and 46/51 PD-Other participants (specificity 90.2%). Independent-samples t tests between the two groups showed that the PD-MR group scored significantly lower on all NAVS-subtests (all p 's < 0.05), except for the derived measures SPPT-NC-C, SCT-NC-C, ASPT-2-1, and ASPT-3-1, where differences did not approach significance. In fact, on the SPPT, the PD-MR group scored numerically higher on production of noncanonical than on canonical structures, but this small difference was also not significant ($t(58) = -0.327, p = 0.745$).

Not surprisingly, the second binary LDA, on group membership of one of the two clusters based on NAVS performance (without our custom ASPT-G score), had high predictive success, at 93.8%. Six participants were excluded due to missing discriminant variables. Variables contributing independently to classification were ASPT-A, SPPT, SPPT-NC-C, and ASPT-2-1. Group assignment was correct for 15/17 members of Group B and for 46/48 members of Group A. Adding ASPT-G brought the classification success necessarily to 100% (by necessity, as the LDA based on clustering was circular and included only to reveal which NAVS scores were driving the clustering into two groups). The same six participants were excluded due to missing discriminant variables, and the variables that contributed independently to the classification were ASPT-A, ASPT-G, SPPT, and SPPT-NC-C. For all of the NAVS subscores, t tests between the two groups

TABLE 6 Distribution of participants based on hierarchical clustering (NAVS: Group A, Group B) and morphosyntactic deficits in descriptive speech (PD-MR, PD-Other, PD-NC; $n = 65$)

	PD-MR	PD-Other	PD-NC
Group A	2	46	0
Group B	8	4	5

NAVS = Northwestern Assessment of Verbs and Sentences; NC = not categorizable; PD-MR = morphosyntactic deficit in picture description; PD-Other = no morphosyntactic deficit; .

revealed that scores were significantly lower (greater impairment) for the NAVS-Low group (all p 's = < 0.05), with the notable exception of SCT-NC-C, for which the difference was not significant ($t(21.16) = 0.938, p = 0.36$), and SPPT-NC-C, for which the NAVS-Low group had significantly *higher* scores. The latter indicates that, relative to the difference scores in the NAVS-Other group (mean = -0.12), the disadvantage for primed noncanonical versus canonical sentences was *smaller* in the NAVS-Low group (mean = -0.03 ; $t(56.96) = -3.060, p < 0.05$).

3.4 | Voxel-wise and connectome lesion-symptom mapping

3.4.1 | Binary classifications

Voxel-based lesion-symptom mapping

Figure S1 (Supporting Information) illustrates the overall structural brain damage associated with patients who have a syntactic deficit, as identified based on hierarchical clustering of NAVS scores (NAVS-Low) and based on morphosyntactically deficient descriptive speech (PD-MR). NAVS grouping was predicted by damage to a range of left-hemisphere perisylvian regions, including inferior and middle frontal cortex, insula, precentral and postcentral gyrus, inferior parietal cortex, supramarginal gyrus, angular gyrus, putamen, superior and middle temporal gyrus, and the superior part of the temporal pole. Damage to largely the same regions is predictive of grouping based on impaired morphosyntactic production. Within these regions, though, the proportional damage that was predictive of NAVS-Low and PD-MR group classifications was very different, with larger lesioned proportions of regions being predictive of PD-MR than of NAVS-Low grouping ($t(15) = 2.901, p < 0.01$).

After regressing out the overall lesion size and correcting for multiple comparisons, the binary classification based on the hierarchical clustering of NAVS subscores was only associated with a single voxel in VLSM. Damage to this area predicted membership of the NAVS-Low group, the group most likely to have a syntactic deficit (Figure 1, red). Its location was medial to the left superior temporal gyrus (Montreal Neurological Institute (MNI) coordinates $[-34, -32, 8]$), crossed by the optic radiations, but on the edge of the inferior

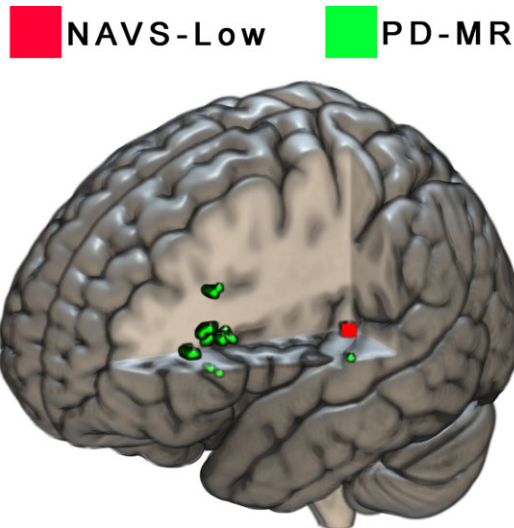


FIGURE 2 Areas associated by VLSM with syntactic deficits based on hierarchical clustering of NAVS scores (red) and based on morphosyntactic deficits in descriptive speech (green). For the purpose of clarity, the voxel associated with the NAVS grouping (red) has been enlarged and accentuated [Color figure can be viewed at wileyonlinelibrary.com]

occipito-frontal fasciculus, the inferior longitudinal fasciculus (ILF), and the internal capsule (IC), as based on an overlay of the tractography white matter atlas by Catani and Thiebaut de Schotten (2008).

The VLSM on the binary classification based on our clinical judgment of morphosyntactic problems in the descriptive speech samples yielded voxels associated with the impaired group (PD-MR) in the triangular part of IFG (Figure 1, green), in both gray and medial white matter, as well as one lateral voxel in superior temporal gyrus (MNI [−66, −17, 9]).

Connectome-based lesion-symptom mapping

Results from the CLSM analysis that was restricted to left-hemisphere dorsal/ventral-stream connections are shown in Figure 2 and Table 7. NAVS-Low and PD-MR group membership was predicted by both typically dorsal connections, for example, between IFG_tri and SMG, but also by typically ventral connectivity, for example, between IFG_orb and pITG (NAVS-Low) and between MFOG and STG (PD-MR).

3.4.2 | Overall ASPT-A, ASPT-G, SPPT, and SCT scores

Voxel-based lesion-symptom mapping

Figure S2 (Supporting Information) shows the regional damage associated with lower scores on the NAVS subscores. There was clear overlap between these regions. Around the left-hemisphere perisylvian area, lower ASPT-A scores were primarily related to damage in posterior temporal and inferior parietal cortex, whereas the ASPT-G scores were associated with a wider cortical area that includes inferior frontal cortex. SPPT scores were supported by (anterior, posterior, and medial) temporal as well as inferior frontal regions, while SCT scores were supported by more posterior temporal cortex, partly overlapping with but inferior to the region associated with ASPT-A scores.

After regressing out overall lesion size, the ASPT-A, ASPT-G, SPPT, and SCT scores were all negatively correlated with damage to a

TABLE 7 Left-hemisphere perisylvian structural connectivity predictive of agrammatic group membership. The label “dorsoventral” is given to connections between frontal and temporal hubs, as well as between regions that are considered part of separate streams

Variable	Predictive structural connectivity	Dorsal/ventral	z	
PD-MR	IFG pars opercularis—IFG pars triangularis	Dorsal	−2.927	
	IFG pars orbitalis—supramarginal gyrus	Dorsoventral	−2.407	
	Lateral fronto-orbital gyrus—supramarginal gyrus	Dorsoventral	−2.920	
	Supramarginal gyrus—angular gyrus	Dorsal	−2.814	
	Middle fronto-orbital gyrus—superior temporal gyrus	Ventral	−2.513	
	Supramarginal gyrus—pole of superior temporal gyrus	Dorsoventral	−2.677	
	Superior temporal gyrus—pole of superior temporal gyrus	Ventral	−2.694	
	Superior temporal gyrus—middle temporal gyrus	Ventral	−2.460	
	Middle temporal gyrus—inferior temporal gyrus	Ventral	−2.890	
	IFG pars orbitalis—insula	Dorsoventral	−2.649	
	Supramarginal gyrus—insula	Dorsal	−2.741	
	Supramarginal gyrus—pSTG	Dorsal	−2.717	
	NAVS-Low	IFG pars opercularis—IFG pars triangularis	Dorsal	−3.216
		IFG pars orbitalis—supramarginal gyrus	Dorsoventral	−2.904
IFG pars triangularis—supramarginal gyrus		Dorsal	−2.544	
Lateral fronto-orbital gyrus—supramarginal gyrus		Dorsoventral	−2.697	
Supramarginal gyrus—angular gyrus		Dorsal	−2.996	
Supramarginal gyrus—pole of superior temporal gyrus		Dorsoventral	−3.152	
Superior temporal gyrus—middle temporal gyrus		Ventral	−2.491	
Supramarginal gyrus—insula		Dorsal	−3.332	
Middle temporal gyrus—insula		Dorsoventral	−2.552	
IFG pars orbitalis—pSTG		Dorsoventral	−2.689	
Lateral fronto-orbital gyrus—pSTG		Ventral	−2.502	
Insula—pSTG		Dorsal	−2.794	
IFG pars triangularis—posterior middle temporal gyrus		Ventral	−2.621	
IFG pars orbitalis—posterior inferior temporal gyrus		Ventral	−2.560	

Notes. IFG = inferior frontal gyrus; NAVS-Low = membership of group with agrammatic performance pattern on the Northwestern Assessment of Verbs and Sentences; PD-MR = morphosyntactic deficit in picture description; pSTG = posterior superior temporal gyrus. Shared connections between both groupings are printed in bold.

middle to posterior region in the superior temporal gyrus, medial to Heschl's gyrus, and on the edge of the optic radiations and the ILF, about where it crosses the long segment of the arcuate fasciculus (AF) (see Figure 3 and Table 8). This is approximately the same region that was identified as associated with syntactic deficits based on the

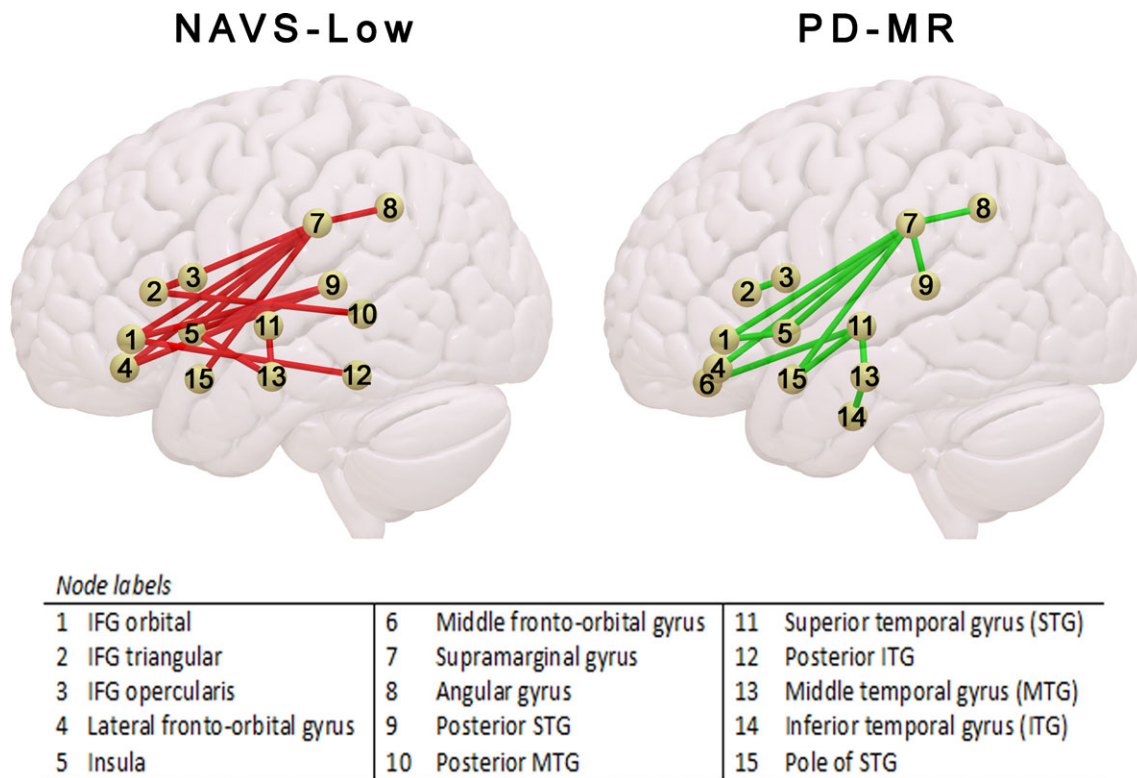


FIGURE 3 Reduced left-hemisphere perisylvian structural connectivity predictive of membership of a group of aphasic speakers with uniformly low performance on the Northwestern Assessment of Verbs and Sentences (NAVS-Low) and a group of aphasic speakers characterized as having a morphosyntactic production deficit, based on their descriptive speech (PD-MR) [Color figure can be viewed at [wileyonlinelibrary.com](#)]

NAVS hierarchical clustering data. In addition, ASPT-A scores were negatively correlated with damage to the angular gyrus, extending into middle occipital gyrus. This latter region feeds into the posterior segment of the AF (Catani & Thiebaut de Schotten, 2008).

Connectome-based lesion-symptom mapping

Results from the CLSM analysis that was restricted to left-hemisphere dorsal/ventral-stream connections are shown in Figure 4. For all four subscores, it is visually evident that both dorsal and ventral connections are predictive of performance on the NAVS. Out of 136 possible connections in this restricted analysis, ASPT-A scores were predicted by integrity of 27 connections, ASPT-G scores by 54 connections, SPPT scores by 60 connections, and SCT scores by 29 connections.

3.4.3 | ASPT three-place and two-place verb production

Voxel-based lesion-symptom mapping

No regional lesion damage was predictive of low scores on three-place or two-place verbs, after partial regression of scores on one-place verbs.

Connectome-based lesion-symptom mapping

The CLSM that was restricted to left-hemisphere dorsal and ventral stream connections showed the dorsoventral connection between MTG and insula to be predictive of poor performance on three-place verbs, after regressing out one-place verb performance. No connectivity values were predictive of two-place verb performance, after regression of one-place verb performance (see Table 9).

3.4.4 | SPPT and SCT noncanonical sentences

Voxel-based lesion-symptom mapping

No regional lesion damage was predictive of low scores on noncanonical sentences on the SPPT, after partial regression with canonical sentence performance. For the SCT, however, low performance on noncanonical sentences, after partial regression with canonical sentence performance, was associated with lesions in posterior temporal and inferior parietal cortex, as shown in Figure 5.

Connectome-based lesion-symptom mapping

The CLSM that was restricted to left-hemisphere dorsal and ventral stream connections revealed 15 connections to be predictive of low

TABLE 8 Regional structural damage negatively correlated with NAVS subscores

	Predictive regional damage	k
ASPT-A	Superior temporal gyrus (BA 41); angular gyrus, middle occipital gyrus	47
ASPT-G	Mid/post. Superior temporal gyrus, medial to Heschl's gyrus	27
SPPT	Mid/post. Superior temporal gyrus, medial to Heschl's gyrus	16
SCT	pSTG	1

Notes. k = number of 1 mm^3 voxels. ASPT-A = Argument Structure Production Test, Code A; ASPT-G = Argument Structure Production Test, Code G; pSTG = posterior superior temporal gyrus; SCT = Sentence Comprehension Test; SPPT = Sentence Production Priming Test.

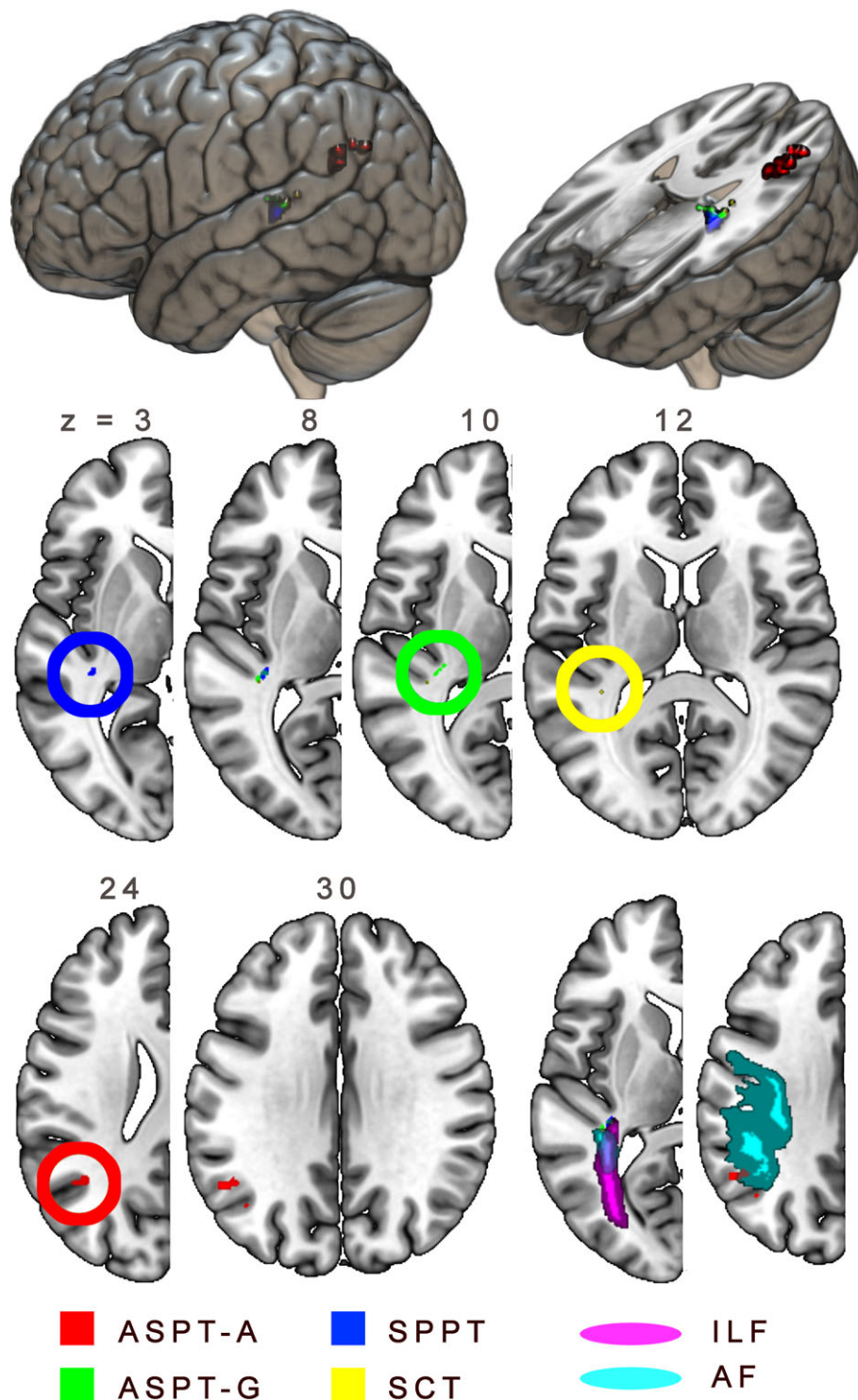


FIGURE 4 Structural lesion predictors of impaired performance on Northwestern Assessment of Verbs and Sentences (NAVS) subtests. ; AF = arcuate fasciculus; ILF = inferior longitudinal fasciculus; ILF and AF based on atlas by Catani and Thiebaut de Schotten (2008) [Color figure can be viewed at wileyonlinelibrary.com]

SPPT scores on noncanonical sentences when canonical sentence scores were partialled out. These are listed in Table 8 and shown in Figure 6. Most of these connections (12/15) are technically assigned to the ventral stream, but all of them have either inferior frontal or middle fronto-orbital nodes that connect either within Broca's area itself, or to superior, middle, and inferior temporal nodes. No connections were predictive of low SCT performance on noncanonical sentences, after partialing out canonical-sentence scores.

4 | DISCUSSION

We assessed neural correlates of agrammatic language performance as measured with subtests of the NAVS, as well as based on an informal clinical assessment of descriptive speech characteristics in a group of stroke survivors with and without aphasia. In the discussion of our results, we first focus on the behavioral results and patterns, before we turn to the lesion and connectome correlations.

TABLE 9 Left-hemisphere perisylvian structural connectivity predictive of NAVS scores on complex structures, after partial regression of scores on less complex structures. The label “dorsoventral” is given to connections between frontal and temporal hubs, as well as between regions that are considered part of separate streams

Variable	Predictive structural connectivity	Dorsal/ventral	z
ASPT 3-place	Middle temporal gyrus–insula	Dorsoventral	2.562
ASPT 2-place	None		
SPPT noncanonical	IFG pars opercularis–IFG pars triangularis left	Dorsal	2.962
	IFG pars orbitalis–IFG pars triangularis	Ventral	3.024
	IFG pars opercularis left–superior temporal gyrus	Dorsoventral	2.812
	IFG pars orbitalis–superior temporal gyrus	Ventral	2.954
	IFG pars triangularis–superior temporal gyrus	Dorsoventral	3.173
	Middle fronto-orbital gyrus–superior temporal gyrus	Ventral	3.048
	IFG pars opercularis–pole of superior temporal gyrus	Ventral	2.633
	IFG pars orbitalis–pole of superior temporal gyrus	Ventral	3.072
	IFG pars triangularis–pole of superior temporal gyrus	Ventral	2.852
	IFG pars triangularis–middle temporal gyrus	Ventral	2.865
	IFG pars opercularis–pole of middle temporal gyrus	Ventral	2.799
	Middle fronto-orbital gyrus–pole of middle temporal gyrus	Ventral	2.873
	Middle fronto-orbital gyrus–inferior temporal gyrus	Ventral	2.853
	IFG pars opercularis–posterior middle temporal gyrus	Ventral	2.645
IFG pars triangularis–posterior middle temporal gyrus	Ventral	2.745	
SCT noncanonical	None		

ASPT-2-1 = Argument Structure Production Test, 2-1; ASPT-3-1 = Argument Structure Production Test, 3-1; IFG = inferior frontal gyrus; NAVS = North-western Assessment of Verbs and Sentences; SCT = Sentence Comprehension Test; SPPT = Sentence Production Priming Test.

4.1 | Identification of syntactic production/comprehension deficits with the NAVS

Overall scores on the subtests of the NAVS correlate with one another, but also strongly with general aphasia severity, as reflected in the WAB-AQ. This reinforces the point that these overall scores should not be used to diagnose or identify aphasic speakers as having a syntax-specific deficit without further analysis. Namely, overall performance on different NAVS subtests can be affected by deficits in speech planning/production, lexical processing/representations, auditory or visual processing, general working memory, as well as reduced resources for attentional control. By contrast, the derived difference scores, in which performance on syntactically less complex structures is subtracted from more complex structures, are not correlated with overall severity or lesion size. In line with existing literature, participants as a group performed worse on production and comprehension of noncanonical sentence structures, and on production of structures with more complex verbs, in terms of argument structure. For example, Love and Oster (2002) report a similar canonicity effect in comprehension that was not specific to a particular group of speakers with expressive versus receptive aphasia and not different between speakers with mild versus severe expressive aphasia. For sentence production and comprehension, such problems with noncanonical sentences are correlated, though weakly at best, suggesting the possibility of a central deficit underlying this performance pattern at least in some patients. Verb argument structure problems, on the other hand, are separate and not correlated with canonicity problems. Instead, problems with complex verbs are correlated with overall sentence production, as captured with the total score on the SPPT and the custom ASPT-G score, which is only awarded in case the produced utterance

is grammatically correct. This may reflect the pivotal role that verbs and their argument structure play in sentence (de)construction, independent of structural differences between canonical and noncanonical word orders. NAVS subscores, complemented with derived measures that capture complexity effects, may thus provide further details on specific problems with syntactic processing in speakers with aphasia, as well as on the relation between different aspects of syntactic processing.

Our hierarchical clustering analysis revealed a group of patients who pattern together in their performance on the NAVS, particularly based on production scores reflecting sentence structure (ASPT-G, SPPT, SPPT-NC-C) as well as verb complexity (ASPT-A, ASPT-2-1). Against our expectations, this group showed relatively better primed production of noncanonical sentences than of canonical sentences. This should not be interpreted as intact performance on noncanonical structures, as overall SPPT scores were indeed significantly lower in this group, but it does show that a specific problem with noncanonical sentence production is either not part of a consistent pattern of “agrammatic” behavior, or not accurately captured by the SPPT. In practice, the primed production task can be challenging even for mildly aphasic speakers, due to its specific requirements to “flip” the arguments used with the reversible verbs in the prime sentence. We may speculate, then, that for some patients who do in fact have a grammatical deficit, it is easier to reverse the arguments after hearing a noncanonical prime, because in those cases the presented argument order is necessarily counter to an agent-first default structure. Taking a strategic approach to the SPPT challenge in which the first (prime) picture is the reference point, patients may thus “flip” the order of the arguments back to the agent-first order they prefer for that prime picture and so produce the correct structure for the second (cue) picture (essentially disregarding its contents). This may be harder to

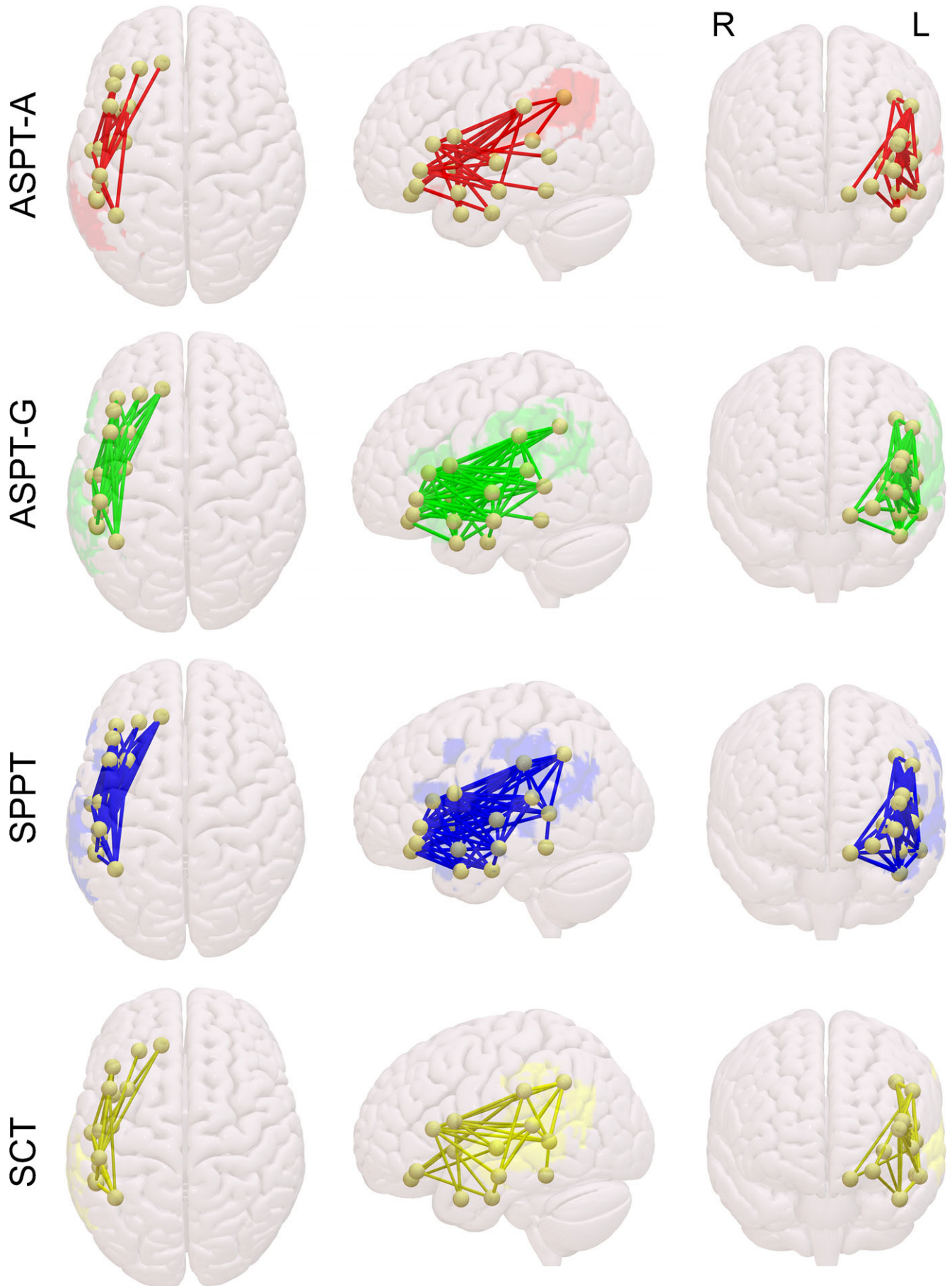


FIGURE 5 Reduced left-hemisphere perisylvian structural connectivity predictive of impaired performance on Northwestern Assessment of Verbs and Sentences (NAVS) subtests [Color figure can be viewed at wileyonlinelibrary.com]

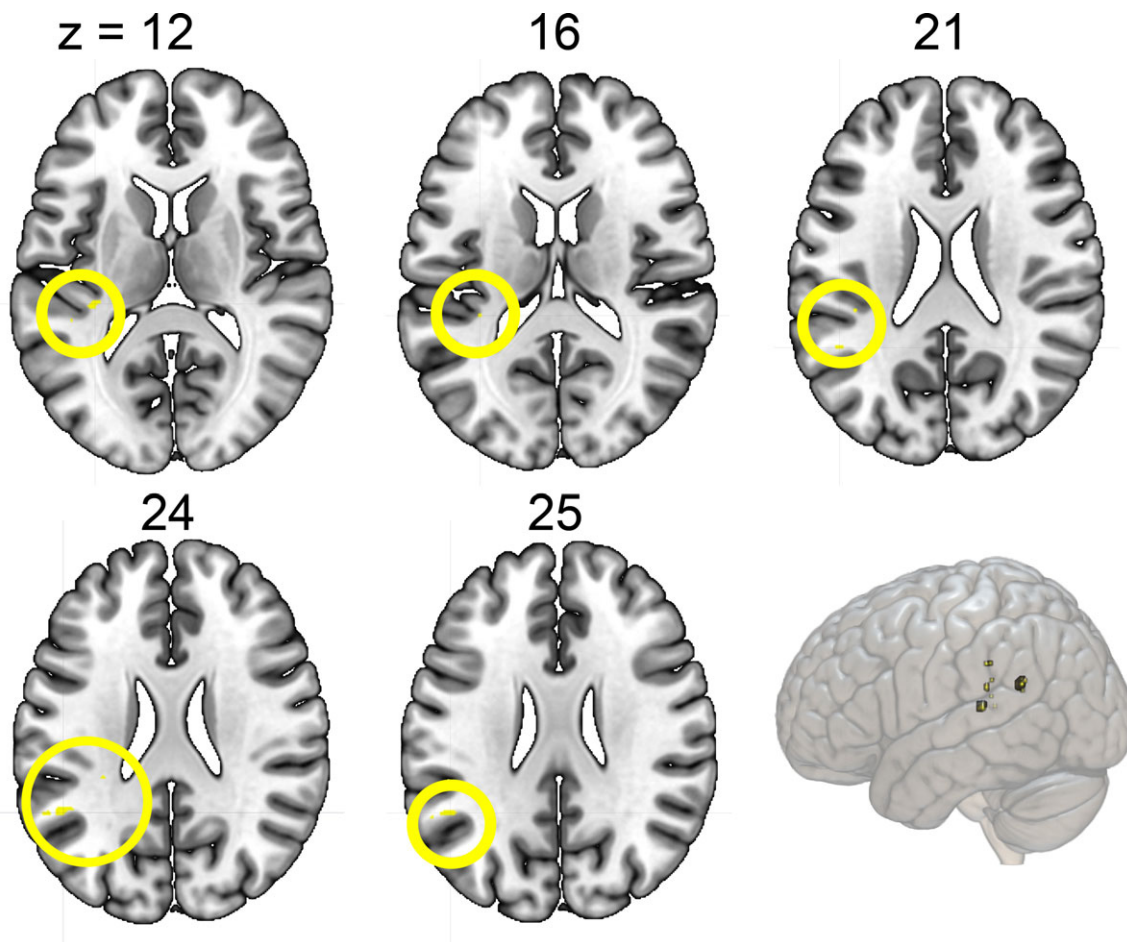


FIGURE 6 Structural lesion predictors of impaired performance on noncanonical sentences in the Sentence Comprehension Test (SCT) of the Northwestern Assessment of Verbs and Sentences (NAVS), controlling for performance on canonical sentences [Color figure can be viewed at wileyonlinelibrary.com]

do for canonical primes, as in these cases the patient has to change the argument order from agent-first to the nonpreferred agent-last (relative to the prime picture). This speculative account for how participants may approach the SPPT and perhaps other language tasks naturally requires further investigation in a dedicated study, and is beyond the scope of the current paper.

Importantly, a nonrandom proportion of these patients were also identified as having a morphosyntactic production deficit based on their narrative spontaneous speech samples (the PD-MR group). The production of morphosyntactically reduced speech itself was best predicted by low scores on the ASPT-G measure, as well as disproportionate difficulty with noncanonical sentences in comprehension (SCT-NC-C). Although there certainly was no complete overlap between the NAVS-Low and the PD-MR groups, these correlations indicate that the NAVS is indeed able to some extent to identify patients who show the typical agrammatic production pattern in their spontaneous speech. Nevertheless, the heterogeneity of results suggests individual variation, likely reflecting different functional substrates for “agrammatic” production patterns.

The identification of different subtypes of agrammatism should be the topic of further investigation, and it is our perspective that a fruitful starting point may be the difference between verb-based deficits and structure-based deficits. As a step toward the optimization of

diagnostic tools, we suggest that the NAVS ASPT-A score and the derived score of ASPT-2-1 may be used to identify verb-based deficits, whereas a negative SCT-NC-C score (greater impairment on noncanonical than canonical sentences) may be used to identify structure-based comprehension deficits. Given the unexpected patterns of performance on the SPPT-NC-C score, more work needs to be done on the isolation of structure-based production deficits, at least as regards canonical versus noncanonical sentence production. Even so, we are not ready to suggest specific cut-off scores on the ASPT-A, ASPT-2-1, and SCT-NC-C measures either, and a thorough diagnosis of individual patients should always be conducted based on the overall performance pattern of the individual, where laboratory subtests may supplement clinical impressions. Although it correlates with ASPT-A, as well as with general aphasia severity (WAB-R AQ), the added measure of ASPT-G was found to be a strong predictor of morphosyntactic problems in spontaneous (descriptive) speech as well, and also correlated with specific problems on the most complex verbs (ASPT-3-1). We would advocate the addition of the ASPT-G measure to the variables coded from the NAVS, especially if no other spontaneous speech analysis is part of the diagnostic battery. Essentially, however, such a spontaneous speech analysis should always be the basis for the overall diagnosis of agrammatic deficits, with subscores on the NAVS serving to further specify the nature of the deficit.

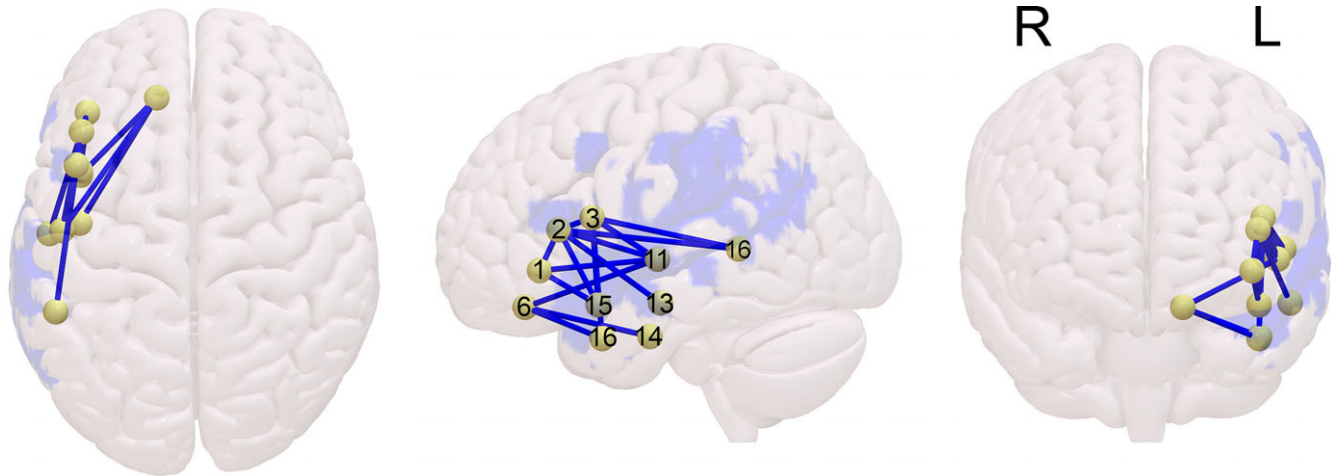


FIGURE 7 Reduced left-hemisphere perisylvian structural connectivity predictive of impaired performance on noncanonical sentences in the Sentence Production Priming Test (SPPT) of the Northwestern Assessment of Verbs and Sentences (NAVS), controlling for performance on canonical sentences [Color figure can be viewed at wileyonlinelibrary.com]

4.2 | Lesion and connectome-based predictors of agrammatic performance

In our discussion of the VLSM data, we will focus on the results of analyses from which the overall effect of lesion size was regressed out, as it is such a strong predictor of general aphasia severity and task performance. As noted (Section 2.4.5), the results from the CLSM analyses did not survive correction for overall lesion size. We therefore treat these as informative but somewhat less reliable, and center our discussion around the VLSM results.

The identification of morphosyntactic production problems in descriptive speech was most strongly predicted by structural lesions in the triangular part of the IFG, which corresponds to Brodmann's area 45 and is part of Broca's area. This is in line with previous studies ascribing a role for this area in syntactic production, though the region is anterior to the area that has been associated most directly with hierarchical sentence structuring, the opercular BA 44 (e.g., Friederici et al., 2006). It is now well accepted that the aphasic syndrome labeled "Broca's aphasia" is not caused by lesions that are restricted to Broca's area, or even necessarily by lesions that include Broca's area (Fridriksson, Bonilha, & Rorden, 2007; Mohr, 1976). However, our results do show that the symptom of agrammatic speech output as strictly characterized by morphosyntactically reduced speech is predicted by lesions to Broca's area. First of all, of course not all speakers diagnosed with Broca's aphasia present with agrammatic speech output, that is, it is not part of the diagnostic criteria. Second, this once more illustrates limits to the clinical usefulness of patient classification into broad aphasia types, if behavioral patterns are not further specified. Importantly, the observed correlation does not mean that every stroke survivor with structural damage to BA 45 will necessarily present with morphosyntactically reduced speech output. It does mean that it is more likely that they do, including in the chronic stage poststroke. The CLSM results further confirm that network integrity along both the dorsal and the ventral stream underlies the preservation of morphosyntactic production after stroke.

Notably, we emphasize that what is labeled here as "agrammatic output" should be primarily characterized more specifically as morphosyntactically reduced speech. Henseler et al. (2014) conducted a

similar analysis, based on a subjective and in their words "unspecific" (927) rating of the integrity of "syntactic structure" in German spontaneous speech. They particularly found damage to posterior temporal and parietal cortex to be predictive of poor performance on the syntactic-structure index. So, it may well be that they were capturing more broad patterns of syntactic disintegration, while our PD-MR analysis deliberately focused on production of morphosyntax.

The group of stroke survivors who patterned together in their performance on the NAVS, with low scores reflecting poor sentence and verb argument structure production, was most strongly characterized not by lesions to Broca's area, but by damage to a small medial posterior superior temporal region. We note that the coordinates reported here (MNI [−34, −32, 8]) are not far removed at all from the area reported by Rogalsky et al. (2018) as associated with agrammatic comprehension problems in a sentence-picture matching task (MNI [−36, −49, 15], converted from Talairach coordinates [−35, −48, 17]). This same region is associated with individual subscores on the NAVS that reflect both production and comprehension difficulties: *ASPT-A*, *ASPT-G*, *SPPT*, and *SCT*. We had hypothesized the overlap between the regions predicting these scores to be located primarily along the dorsal stream between IFG and pSTG, but the area we identified indeed appears to be more ventral than dorsal. The CLSM results, however, do show that the NAVS-Low grouping is also predicted by reduced connectivity in both the dorsal and the ventral stream.

These data therefore contribute to the mounting base of evidence for the vital involvement of posterior temporal and ventral-stream regions in syntactic processing. In itself, this is helpful to our understanding of neural predictors of aphasic behavioral performance. However, the next question that will need to be addressed concerns the specific underlying function that is supported by the ventral connections running through this area and/or by the proximate cortical areas that may be part of the affected region.¹ What aspects of syntax

¹Because our analysis with lesion size regressed out is statistically conservative, we take into account the possibility that Type II error may be masking some gray matter predictors of the same behavioral scores.

appear to be supported by the different components of the network, as reflected in agrammatic performance patterns?

We had hypothesized that the lesion-symptom maps for the NAVS subtests would not only overlap, but also show distinct areas predictive of performance on different tasks. For the ASPT-G, SPPT, and SCT this was not the case, in that no separate regions survived the correction for lesion size in our analyses. We did note that we expected lesion maps associated with these tasks to be distributed, rather than focal, reflecting the many potential functional substrates of overall task performance. It is not unlikely, therefore, that the absence of more task-specific regional predictors in the group analysis is due to between-subject variance in structural (and functional) damage that affects overall task performance. This is supported by both the extent and separation of task-predictive areas as revealed in the analysis without correction for lesion size (Figure S2, Supporting Information).

It does not seem to be the case that syntactic (de)construction problems can all be reduced to difficulties with processing or accessing verb argument structure. In line with our hypothesis, the ASPT-A score on the NAVS, which is particularly reflective of verb argument structure manipulations, is strongly predicted by damage to the angular gyrus. This region is more posterior and dorsal than the medial pSTG region where we found the overlap, and is placed on the border between the dorsal and ventral processing streams in a dual-stream architecture (cf., Uddin et al., 2010). Damage to this area did not predict other overall scores on NAVS subtests. Conceptually, the location of this region makes a lot of sense. If “syntactic processing” is not exclusively a function of the dorsal stream, but rather of the interplay between dorsal and ventral streams, then a neural substrate for verb argument structure access/processing in the hub that connects these two streams fits well with the idea that verbs are pivotal to sentence (de)construction. We suggest that this is also reflected in our observation that production of correct argument structure for three-place verbs on the ASPT, after partialing out one-place verbs, is predicted by the dorsoventral connectivity between middle temporal gyrus and insula. It is also in line with a host of functional neuroimaging studies that have identified posterior superior temporal and inferior parietal neural substrates for verbs and their argument structure (Bornkessel et al., 2005; Den Ouden et al., 2009; Grewe et al., 2007; Shetreet, Palti, Friedmann, & Hadar, 2007; Thompson et al., 2007, 2010). Interestingly, the region predicting ASPT-A scores overlaps with what Binder and Desai (2011) identify as an amodal region that supports semantic event knowledge. It is possible, therefore, that our results reflect the interface between semantic event knowledge and verb argument structure, rather than stored argument-structure representations themselves.

Very near the region predicting ASPT-A scores is a region that predicted specific difficulty with noncanonical sentence structures in comprehension. We suggest, therefore, that these data are in line with the idea that access to verb argument structure and an ability to process thematic role assignment by verbs, through integration with semantic knowledge, are indeed important components to the processing of sentences in which arguments are placed out of their canonical order.

However, our results showing more anterior ventral lesions to be predictive of impaired production of grammatically well-formed sentences are also in line with previous studies that show more anterior superior and middle temporal regions to support syntactic processing

and production (Den Ouden et al., 2012; Dronkers et al., 2004; Grossman et al., 2013; Henseler et al., 2014; Ivanova et al., 2016; Magnusdottir et al., 2013; Rogalsky et al., 2018). In the current data, this is reflected in lower overall scores on subtasks of the NAVS. Although these scores themselves are correlated with an agrammatic production pattern in descriptive speech, that pattern itself was not predicted by the same ventral lesion. It is possible that overall task performance on the NAVS subtests is affected by lexical-semantic difficulties, and previous studies indeed suggest that ventral-stream regions support the mapping between semantic representations and syntactic structure (Saur et al., 2008). Again, a small cluster of voxels that is predictive of specific problems with noncanonical sentences in comprehension is also in the near proximity to the ventral region predicting overall NAVS scores.

Within the perisylvian dual-stream architecture, group membership of the PD-MR and the NAVS-Low groups was predicted by reduced connectivity between an array of regions, affecting dorsal as well as ventral connections. Syntactic (de)construction abilities appear to rely on the integrity of the interplay between dorsal and ventral processing streams, and cannot be reduced to a single neural substrate. Interestingly, one notably absent result in our VLSM analysis was that for non-canonical sentence production on the SPPT. Above, we have already discussed potential problems with this task in light of counterintuitive patterns of performance by patients with other grammatical problems. The absence of VLSM results for this measure appears to be another indication that the comparison of noncanonical and canonical sentence production based on the SPPT may not yield a valid or consistent reflection of patients' abilities. By contrast, the CLSM analysis revealed that preserved noncanonical sentence production on this priming task, after partialing out canonical sentence production, was strongly reliant on the integrity of ventral-stream connections to IFG, Broca's area, including both its triangular and opercular regions.

4.3 | Limitations

As noted above, “syntactic processing” is a heterogeneous component of the language system, with many aspects that are not addressed in the current paper. Here, we have focused on aphasic task performance on a battery that is specifically designed to identify broad problems with sentence production and comprehension that can be related to verbs and canonicity of word order in English, the NAVS. It does not address specific problems aphasic speakers may have with other aspects of syntax, such as anaphora (Bos, Dragoy, Avrutin, Iskra, & Bastiaanse, 2014; Choy & Thompson, 2010; Engel, Shapiro, & Love, 2018), case marking (Bastiaanse, Jonkers, Ruigendijk, & Van Zonneveld, 2003), discourse-linked temporal references (Bastiaanse, 2013), and adjectives (Meltzer-Asscher & Thompson, 2014), to name but a few.

For most patients in the sample, who were enrolled in this study when they were already years poststroke, we do not have detailed information about their behavioral profiles before we tested them. It is therefore quite possible that some patients who did not have aphasia at the time of testing may have shown aphasic symptoms in earlier stages of their recovery. The current paper is not able to address the important question of how behavior in the chronic poststroke stage may be a function of compensatory strategies and which types of

structural damage may only lead to temporary (syntactic) problems. However, the VLSM and CLSM analyses in the current paper do reflect how neural damage predicts behavior in the chronic stage.

We have assessed to what extent patients who do not do well on subtasks of the NAVS would also be identified as having an agrammatic production pattern, based on samples of descriptive speech. Almost inherently, no task is ideal for the elicitation of natural spontaneous speech in aphasic speakers. Picture description, such as used here, may invoke a tendency to list the objects in the picture, rather than produce full sentences. Story-(re)telling, an alternative method, may make heavy demands on (working) memory as well as expose individual differences in verbosity that do not necessarily reflect language ability. In both of these methods, problems with lexical retrieval may underlie simplified sentence constructions. Samples of spontaneous conversational speech are relatively uncontrolled, with output heavily dependent on the engagement of the speaker with topic and partner. So, our basis for the clinical identification of “agrammatic output” was limited. We did deliberately adhere to a narrow characterization of agrammatic speech, namely in terms of reduced morpho-syntactic production, following the classic definition. This was also prompted by our wish to reflect what commonly happens in clinical practice, where agrammatic output is generally characterized based on speech output, rather than on extensive testing with targeted batteries or detailed linguistic analyses of aphasic speech. The fact that even the two researchers who performed the classification did not immediately agree on all samples exemplifies the elusiveness of the concept of “agrammatism” in practice. Nevertheless, Table S1 (Supporting Information) shows that the classification at the very least appears to capture a range of quantifiable linguistic characteristics of the speech samples, where the two groups clearly dissociate. We have already made the point that we do not believe agrammatism is a homogeneous syndrome. Nevertheless, like others, our group is currently working on more detailed analyses of different types of aphasic speech samples, in order to assess to what extent various aspects of syntactic production can be independently impaired, and how such impairments may have specific neural correlates.

5 | CONCLUSIONS

The individual variation observed in previous studies implies that there are more ways than one in which syntactic production or comprehension can be impaired. The complexity and multidimensionality of what is termed “syntax” also suggest it is highly unlikely that a single process, let alone a single brain region, is responsible for sentence (de)construction. Based on the present data, then, and against the background of existing literature, we conclude that lesions in Broca's area, particularly the triangular part of the IFG, do tend to lead to agrammatic speech output, as defined by omission and incorrect use of grammatical morphemes, articles, and prepositions. Lesions in pSTG and angular gyrus tend to result in difficulty with access to verb argument structure, and/or its integration with semantic knowledge, which may lead to particular problems with thematic role identification and noncanonical sentence comprehension. Lesions to medial superior temporal gyrus, overlapping with ventral stream processing routes,

are associated with generally poor performance on sentence production and comprehension tasks and difficulty with noncanonical sentences in comprehension. For syntactic production, the construction of noncanonical sentences, insofar as it could be captured accurately by the elicitation task used in the NAVS, is not so much reliant on particular areas, but crucially on ventral connections that feed into or are fed by regions in IFG, including those that typically show up as being involved in complex syntactic processing in neuroimaging studies. This is in line with a ventral stream supporting the integration of (lexical) semantic representations and syntactic structure. Fully intact syntactic (de)construction relies on the integrity of all these regions, interacting in a network configuration.

ACKNOWLEDGMENTS

This work was supported by the National Institute on Deafness and Other Communication Disorders (P50 DC014664, PI J.F.; T32 DC0014435, trainee A.B.). The authors are grateful to Kristie Kannaley and Danielle Fahey for assistance with data analysis, to Sara Sayers and Allison Croxton for assistance with data collection, and to their research participants for their time and effort.

ORCID

Dirk-Bart den Ouden  <https://orcid.org/0000-0001-6971-7606>

Svetlana Malyutina  <https://orcid.org/0000-0003-4481-122X>

Alexandra Basilakos  <https://orcid.org/0000-0003-0053-1082>

Leonardo Bonilha  <https://orcid.org/0000-0002-6153-8942>

Ezequiel Gleichgerrcht  <https://orcid.org/0000-0002-4212-4146>

Argye E. Hillis  <https://orcid.org/0000-0002-5192-1171>

Chris Rorden  <https://orcid.org/0000-0002-7554-6142>

REFERENCES

- Alajouanine, T. (1968). *L'Aphasie et le langage pathologique*. Paris: J.-B. Bailière et Fils.
- Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, 26(3), 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018>
- Badecker, W., & Caramazza, A. (1985). On considerations of method and theory governing the use of clinical categories in neurolinguistics and cognitive neuropsychology – The case against agrammatism. *Cognition*, 20(2), 97–125. [https://doi.org/10.1016/0010-0277\(85\)90049-6](https://doi.org/10.1016/0010-0277(85)90049-6)
- Bastiaanse, R. (2013). Why reference to the past is difficult for agrammatic speakers. *Clinical Linguistics & Phonetics*, 27(4), 244–263. <https://doi.org/10.3109/02699206.2012.751626>
- Bastiaanse, R., Hurkmans, J., & Links, P. (2006). The training of verb production in Broca's aphasia: A multiple-baseline across-behaviours study. *Aphasiology*, 20(2–4), 298–311. <https://doi.org/10.1080/02687030500474922>
- Bastiaanse, R., Jonkers, R., Ruijendijk, E., & Van Zonneveld, R. (2003). Gender and case in agrammatic production. *Cortex*, 39(3), 405–417. [https://doi.org/10.1016/s0010-9452\(08\)70256-9](https://doi.org/10.1016/s0010-9452(08)70256-9)
- Bastiaanse, R., Rispens, J., Ruijendijk, E., Rabadan, O. J., & Thompson, C. K. (2002). Verbs: Some properties and their consequences for agrammatic Broca's aphasia. *Journal of Neurolinguistics*, 15(3–5), 239–264. [https://doi.org/10.1016/s0911-6044\(01\)00032-x](https://doi.org/10.1016/s0911-6044(01)00032-x)
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., & Dronkers, N. F. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, 6, 448–450. <https://doi.org/10.1038/nn1050>
- Bazzini, A., Zonca, G., Craca, A., Cafforio, E., Cellamare, F., Guarnaschelli, C., ... Luzzatti, C. (2012). Rehabilitation of argument structure deficits in aphasia. *Aphasiology*, 26(12), 1440–1460. <https://doi.org/10.1080/02687038.2012.690023>

- Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1), 144–155. <https://doi.org/10.1016/j.neuroimage.2006.09.018>
- Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, 58(3), 289–308. [https://doi.org/10.1016/0010-0277\(95\)00682-6](https://doi.org/10.1016/0010-0277(95)00682-6)
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, 26(1), 221–233. <https://doi.org/10.1016/j.neuroimage.2005.01.032>
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language*, 125(1), 60–76. <https://doi.org/10.1016/j.bandl.2013.01.010>
- Bos, L. S., Dragoy, O., Avrutin, S., Iskra, E., & Bastiaanse, R. (2014). Understanding discourse-linked elements in aphasia: A threefold study in Russian. *Neuropsychologia*, 57, 20–28. <https://doi.org/10.1016/j.neuropsychologia.2014.02.017>
- Butterworth, B., & Howard, D. (1987). Paragrammatism. *Cognition*, 26, 1–37.
- Caplan, D. (1986). In defense of agrammatism. *Cognition*, 24(3), 263–276. [https://doi.org/10.1016/s0010-0277\(86\)80004-x](https://doi.org/10.1016/s0010-0277(86)80004-x)
- Caplan, D. (1987). *Neurolinguistic and linguistic aphasiology; an introduction*. Cambridge, England: Cambridge University Press.
- Caplan, D. (1991). Agrammatism is a theoretically coherent aphasic category. *Brain and Language*, 40(2), 274–281. [https://doi.org/10.1016/0093-934x\(91\)90128-n](https://doi.org/10.1016/0093-934x(91)90128-n)
- Caplan, D., Hildebrandt, N., & Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain*, 119(Pt. 3), 933–949.
- Caplan, D., Michaud, J., Hufford, R., & Makris, N. (2016). Deficit-lesion correlations in syntactic comprehension in aphasia. *Brain and Language*, 152, 14–27. <https://doi.org/10.1016/j.bandl.2015.10.005>
- Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, 3(4), 572–582.
- Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, 44(8), 1105–1132. <https://doi.org/10.1016/j.cortex.2008.05.004>
- Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology*, 26(10), 1250–1277. <https://doi.org/10.1080/02687038.2012.693584>
- Choy, J. J., & Thompson, C. K. (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, 24(5), 551–579. <https://doi.org/10.1080/02687030802634025>
- Den Ouden, D. B., Fix, S., Parrish, T. B., & Thompson, C. K. (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, 22(2), 196–215.
- Den Ouden, D. B., Hoogduin, H., Stowe, L. A., & Bastiaanse, R. (2008). Neural correlates of Dutch verb second in speech production. *Brain and Language*, 104(2), 122–131. <https://doi.org/10.1016/j.bandl.2007.05.001>
- Den Ouden, D. B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., ... Thompson, C. K. (2012). Network modulation during complex syntactic processing. *NeuroImage*, 59(1), 815–823. <https://doi.org/10.1016/j.neuroimage.2011.07.057>
- Dickerson, B., Michaud, J., Hufford, R., & Caplan, D. (2013). Deficit lesion correlation for syntactic comprehension differs as a function of task. *Procedia - Social and Behavioral Sciences*, 94, 226–227. <https://doi.org/10.1016/j.sbspro.2013.09.112>
- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1–2), 145–177. <https://doi.org/10.1016/j.cognition.2003.11.002>
- Engel, S., Shapiro, L. P., & Love, T. (2018). Proform-antecedent linking in individuals with agrammatic aphasia: A test of the intervener hypothesis. *Journal of Neurolinguistics*, 45, 79–94. <https://doi.org/10.1016/j.jneuroling.2017.10.001>
- Faria, A. V., Joel, S. E., Zhang, Y., Oishi, K., van Zijl, P. C., Miller, M. I., ... Mori, S. (2012). Atlas-based analysis of resting-state functional connectivity: Evaluation for reproducibility and multi-modal anatomy-function correlation studies. *NeuroImage*, 61(3), 613–621. <https://doi.org/10.1016/j.neuroimage.2012.03.078>
- Fridriksson, J., Bonilha, L., & Rorden, C. (2007). Severe Broca's aphasia without Broca's area damage. *Behavioural Neurology*, 18(4), 237–238.
- Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., Den Ouden, D. B., & Rorden, C. (2016). Revealing the dual streams of speech processing. *Proceedings of the National Academy of Sciences of the United States of America*, 113(52), 15108–15113. <https://doi.org/10.1073/pnas.1614038114>
- Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268. <https://doi.org/10.1016/j.tics.2012.04.001>
- Friederici, A. D. (2017). Evolution of the neural language network. *Psychonomic Bulletin & Review*, 24(1), 41–47. <https://doi.org/10.3758/s13423-016-1090-x>
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2458–2463.
- Gleichgerrcht, E., Fridriksson, J., den Ouden, D.B., Delgaizo, J., Stark, B., Hickok, G., Rorden, C., Wilmskoetter, J., Hillis, A.E. & Bonilha, L. (submitted). Neural bases of elements of syntax during speech production in patients with aphasia.
- Gleichgerrcht, E., Fridriksson, J., Rorden, C., & Bonilha, L. (2017). Connectome-based lesion-symptom mapping (CLSM): A novel approach to map neurological function. *NeuroImage: Clinical*, 16, 461–467.
- Goodglass, H. (1973). Studies on the grammar of aphasics. In H. Goodglass & S. Blumstein (Eds.), *Psycholinguistics and aphasia* (pp. 183–215). Baltimore, MD: The Johns Hopkins University Press.
- Goodglass, H. (1997). Agrammatism in aphasiology. *Clinical Neuroscience*, 4(2), 51–56.
- Goodglass, H., & Berko, J. (1960). Agrammatism and inflectional morphology in English. *Journal of Speech, Language, and Hearing Research*, 3(3), 257–267. <https://doi.org/10.1044/jshr.0303.257>
- Grande, M., Meffert, E., Schoenberger, E., Jung, S., Frauenrath, T., Huber, W., ... Heim, S. (2012). From a concept to a word in a syntactically complete sentence: An fMRI study on spontaneous language production in an overt picture description task. *NeuroImage*, 61(3), 702–714. <https://doi.org/10.1016/j.neuroimage.2012.03.087>
- Grewe, T., Bornkessel-Schlesewsky, I., Zysset, S., Wiese, R., Von Cramon, D. Y., & Schlesewsky, M. (2007). The role of the posterior superior temporal sulcus in the processing of unmarked transitivity. *NeuroImage*, 35, 343–352.
- Grodzinsky, Y. (1984). The syntactic characterization of agrammatism. *Cognition*, 16(2), 99–120. [https://doi.org/10.1016/0010-0277\(84\)90001-5](https://doi.org/10.1016/0010-0277(84)90001-5)
- Grossman, M., Powers, J., Ash, S., McMillan, C., Burkholder, L., Irwin, D., & Trojanowski, J. Q. (2013). Disruption of large-scale neural networks in non-fluent/agrammatic variant primary progressive aphasia associated with frontotemporal degeneration pathology. *Brain and Language*, 127(2), 106–120. <https://doi.org/10.1016/j.bandl.2012.10.005>
- Heeschen, C. (1985). Agrammatism versus paragrammatism: A fictitious opposition. In M. L. Kean (Ed.), *Agrammatism* (pp. 207–248). Orlando, FL: Academic Press.
- Henseler, I., Regenbrecht, F., & Obrig, H. (2014). Lesion correlates of pathologic profiles in chronic aphasia: Comparisons of syndrome-, modality- and symptom-level assessment. *Brain*, 137, 918–930. <https://doi.org/10.1093/brain/awt374>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/Nrn2113>
- Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., & Hagoort, P. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 5933–5936.
- Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language*, 89, 312–319.

- Ivanova, M. V., Isaev, D. Y., Dragoy, O. V., Akinina, Y. S., Petrushevskiy, A. G., Fedina, O. N., ... Dronkers, N. F. (2016). Diffusion-tensor imaging of major white matter tracts and their role in language processing in aphasia. *Cortex*, 85, 165–181. <https://doi.org/10.1016/j.cortex.2016.04.019>
- Kean, M. L. (Ed.). (1985). *Agrammatism*. Orlando, FL: Academic Press.
- Kertesz, A. (2007). *Western aphasia battery - revised*. San Antonio, TX: Pearson.
- Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, 74(1), 1–25. <https://doi.org/10.1006/brln.2000.2315>
- Kleist, K. (1916). Über Leitungsaplasie und grammatische Störungen. *Zeitschrift für Psychiatrie und Neurologie*, 40, 118–199.
- Kolk, H. H., & Van Grunsven, M. M. (1985). Agrammatism as a variable phenomenon. *Cognitive Neuropsychology*, 2(4), 347–384.
- Kussmaul, A. (1877). Die störungen der prache. Versuch einer pathologie der sprache. In H. V. Ziemssen (Ed.), *Handbuch der speziellen pathologie und therapie* (Vol. 12). Leipzig, Germany: F.C.W. Vogel.
- Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIFTI conversion. *Journal of Neuroscience Methods*, 264, 47–56. <https://doi.org/10.1016/j.jneumeth.2016.03.001>
- Love, T., & Oster, E. (2002). On the categorization of aphasic typologies: The SOAP (a test of syntactic complexity). *Journal of Psycholinguistic Research*, 31(5), 503–529.
- Lukic, S., Bonakdarpour, B., Den Ouden, D. B., Price, C., & Thompson, C. (2013). Neural mechanisms of verb and sentence production: A lesion-deficit study. *Procedia - Social and Behavioral Sciences*, 94(0), 34–35. <https://doi.org/10.1016/j.sbspro.2013.09.014>
- Magnusdottir, S., Fillmore, P., Den Ouden, D. B., Hjaltason, H., Rorden, C., Kjartansson, O., ... Fridriksson, J. (2013). Damage to left anterior temporal cortex predicts impairment of complex syntactic processing: A lesion-symptom mapping study. *Human Brain Mapping*, 34(10), 2715–2723. <https://doi.org/10.1002/Hbm.22096>
- Malyutina, S., Richardson, J. D., & Den Ouden, D. B. (2016). Verb argument structure in narrative speech: Mining AphasiaBank. *Seminars in Speech and Language*, 37(01), 034–047. <https://doi.org/10.1055/s-0036-1572383>
- Meltzer-Asscher, A., & Thompson, C. K. (2014). The forgotten grammatical category: Adjective use in agrammatic aphasia. *Journal of Neurolinguistics*, 30, 48–68. <https://doi.org/10.1016/j.jneuroling.2014.04.001>
- Miceli, G., Mazzucchi, A., Menn, L., & Goodglass, H. (1983). Contrasting cases of Italian agrammatic aphasia without comprehension disorder. *Brain and Language*, 19(1), 65–97. [https://doi.org/10.1016/0093-934x\(83\)90056-1](https://doi.org/10.1016/0093-934x(83)90056-1)
- Mohr, J. P. (1976). Broca's area and Broca's aphasia. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 1, pp. 201–235). New York, NY: Academic Press.
- Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., ... Mazziotta, J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, 40(2), 570–582. <https://doi.org/10.1016/j.neuroimage.2007.12.035>
- Pick, A. (1913). *Die agrammatischen sprachstörungen*. Berlin, Germany: Springer.
- Pillay, S. B., Binder, J. R., Humphries, C., Gross, W. L., & Book, D. S. (2017). Lesion localization of speech comprehension deficits in chronic aphasia. *Neurology*, 88(10), 970–975. <https://doi.org/10.1212/wnl.0000000000003683>
- Rogalsky, C., & Hickok, G. (2011). The Role of Broca's Area in Sentence Comprehension. *Journal of Cognitive Neuroscience*, 23(7), 1664–1680. <https://doi.org/10.1162/jocn.2010.21530>
- Rogalsky, C., LaCroix, A. N., Chen, K. H., Anderson, S. W., Damasio, H., Love, T., & Hickok, G. (2018). The neurobiology of agrammatic sentence comprehension: A lesion study. *Journal of Cognitive Neuroscience*, 30(2), 234–255. https://doi.org/10.1162/jocn_a_01200
- Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H. O. (2012). Age-specific CT and MRI templates for spatial normalization. *NeuroImage*, 61(4), 957–965. <https://doi.org/10.1016/j.neuroimage.2012.03.020>
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, 37(3), 440–479.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46), 18035–18040. <https://doi.org/10.1073/pnas.0805234105>
- Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology*, 8, 19–54. <https://doi.org/10.1080/02687039408248639>
- Shetreet, E., Palti, D., Friedmann, N., & Hadar, U. (2007). Cortical representation of verb processing in sentence comprehension: Number of complements, subcategorization and thematic frames. *Cerebral Cortex*, 17(8), 1958–1969.
- Thompson, C. K. (2001). Treatment of underlying forms: A linguistic specific approach for sentence production deficits in agrammatic aphasia. In R. Chapey (Ed.), *Language intervention strategies in adult aphasia* (4th ed., pp. 605–628). Baltimore, MD: Lippincott Williams & Wilkins.
- Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19(11), 1753–1767.
- Thompson, C. K., Bonakdarpour, B., & Fix, S. F. (2010). Neural mechanisms of verb argument structure processing in agrammatic aphasic and healthy age-matched listeners. *Journal of Cognitive Neuroscience*, 22(9), 1993–2011.
- Thompson, C. K., Shapiro, L. P., & Roberts, M. M. (1993). Treatment of sentence production deficits in aphasia: A linguistic-specific approach to wh-interrogative training and generalization. *Aphasiology*, 7, 111–133.
- Thothathiri, M., Kimberg, D. Y., & Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *Journal of Cognitive Neuroscience*, 24(1), 212–222. https://doi.org/10.1162/jocn_a_00118
- Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., ... Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: Function, structure and behaviour in patients with left hemisphere damage. *Brain*, 134(Pt. 2), 415–431. <https://doi.org/10.1093/brain/awq369>
- Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Reorganization of syntactic processing following left-hemisphere brain damage: Does right-hemisphere activity preserve function? *Brain*, 133(11), 3396–3408. <https://doi.org/10.1093/brain/awq262>
- Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: Evidence from functional and structural connectivity. *Cerebral Cortex*, 20(11), 2636–2646. <https://doi.org/10.1093/cercor/bhq011>
- Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., ... Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron*, 72(2), 397–403. <https://doi.org/10.1016/j.neuron.2011.09.014>
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92, 381–397. <https://doi.org/10.1016/j.neuroimage.2014.01.060>
- Yourganov, G., Fridriksson, J., Rorden, C., Gleichgerrcht, E., & Bonilha, L. (2016). Multivariate connectome-based symptom mapping in post-stroke patients: Networks supporting language and speech. *Journal of Neuroscience*, 36(25), 6668–6679. <https://doi.org/10.1523/Jneurosci.4396-15.2016>
- Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, 39(1), 14–32.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: den Ouden D-B, Malyutina S, Basilakos A, et al. Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Hum Brain Mapp*. 2019;40:2153–2173. <https://doi.org/10.1002/hbm.24514>