



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Mapping articulatory and grammatical subcomponents of fluency deficits in post-stroke aphasia

Citation for published version:

Mirman, D, Kraft, AE, Harvey, DY, Brecher, AR & Schwartz, MF 2019, 'Mapping articulatory and grammatical subcomponents of fluency deficits in post-stroke aphasia', *Cognitive, Affective, & Behavioral Neuroscience*, vol. 19, no. 5, pp. 1286–1298. <https://doi.org/10.3758/s13415-019-00729-9>

Digital Object Identifier (DOI):

[10.3758/s13415-019-00729-9](https://doi.org/10.3758/s13415-019-00729-9)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Cognitive, Affective, & Behavioral Neuroscience

Publisher Rights Statement:

This is a post-peer-review, pre-copyedit version of an article published in Cognitive, Affective, & Behavioral Neuroscience . The final authenticated version is available online at: <http://dx.doi.org/10.3758/s13415-019-00729-9>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Mapping articulatory and grammatical sub-components of fluency deficits in post-stroke aphasia

Running head: Mapping fluency deficits in aphasia

Daniel Mirman^{1,2,4}, Amanda E. Kraft³, Denise Y. Harvey^{4,5}, Adelyn R. Brecher⁴, and Myrna F. Schwartz⁴

¹ Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, USA

² Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK

³ Drexel University, Philadelphia, PA, and Advanced Technology Laboratories, Lockheed Martin, Cherry Hill, NJ, USA

⁴ Moss Rehabilitation Research Institute, Elkins Park, PA, USA

⁵ University of Pennsylvania, Department of Neurology, Philadelphia, PA, USA

RUNNING HEAD: Mapping fluency deficits in post-stroke aphasia

Corresponding author:

Daniel Mirman

Department of Psychology

University of Alabama at Birmingham

1300 University Blvd.

Birmingham, AL 35294

Phone: +1 205-934-8737

Email: dan@danmirman.org

ABSTRACT

Fluent speech production is a critical aspect of language processing and is central to aphasia diagnosis and treatment. Multiple cognitive processes and neural sub-systems must be coordinated to produce fluent narrative speech. To refine the understanding of these systems, measures that minimize the influence of other cognitive processes were defined for articulatory deficits and grammatical deficits. Articulatory deficits were measured by the proportion of phonetic errors (articulatory and prosodic) in a word repetition task in 115 participants with aphasia following left hemisphere stroke. Grammatical deficits were assessed in 46 participants based on two measures – proportion of closed class words and proportion of words in sentences – generated during semi-structured narrative speech production (telling the Cinderella story). These measures were used to identify brain regions critical for articulatory and grammatical aspects of speech production using a multivariate lesion-symptom mapping approach based on support vector regression. Phonetic error proportion was associated with damage to the postcentral gyrus and the inferior parietal lobule (particularly the supramarginal gyrus). Proportion of closed class words in narrative speech did not have consistent lesion correlates. Proportion of words in sentences was strongly associated with frontal lobe damage, particularly the inferior and middle frontal gyri. Grammatical sentence structuring relies on frontal regions, particularly the inferior and middle frontal gyri, whereas phonetic-articulatory planning and execution relies on parietal regions, particularly the postcentral and supramarginal gyri. These results clarify and extend current understanding of the functional components of the fronto-parietal speech production system.

Keywords: Fluency; Aphasia; Lesion-symptom mapping

INTRODUCTION

Fluent speech production is a critical aspect of real-world language processing and a clinically-relevant aspect of language deficits following stroke. The focus on fluency in aphasia has a long history and continues to be central to aphasia diagnosis and treatment. Common clinical measures of fluency focus on functional aspects of producing connected speech, such as spontaneous-speech rate (words per minute), average length of utterance, and the fluency sub-score of the Western Aphasia Battery (WAB; Kertesz, 1982), a clinical rating that integrates across grammatical, lexical, and speech-motor aspects of connected speech. Indeed, it has long been recognized that fluent production of connected speech is not “one thing”; rather, that it is a consequence of effective coordination of multiple sub-systems. For example, in primary progressive aphasia (PPA), speech rate (words per minute) is correlated with, but dissociable from, grammatical deficits (Thompson et al., 2012). Recent large-scale studies of individuals with post-stroke aphasia have combined principal components analysis with lesion-symptom mapping (LSM) and found two main sub-components of fluency that are dissociable both behaviorally and neuroanatomically: segment-level (phonological/phonetic) speech production deficits and sentence- or utterance-level speech production deficits (Halai, Woollams, & Lambon Ralph, 2017; Lacey, Skipper-Kallal, Xing, Fama, & Turkeltaub, 2017). These studies identified a fronto-parietal speech production system in which sentence-level deficits were associated with frontal damage and segment-level deficits associated with damage in inferior parietal damage (for a review see Mirman & Thye, 2018).

In a detailed LSM analysis of connected speech deficits in PPA, Wilson et al. (2010) found that degeneration of frontal regions was associated with speech sound distortions (phonetic errors) and syntactic deficits, whereas phonological errors were associated with posterior temporal degeneration. The anterior locus for motor speech deficits converges with LSM evidence from apraxia of speech (AOS), a disorder of articulatory planning and programming that is distinguished primarily by speech distortions and effortful articulation, but also features substitution and addition of well-formed segments (phonological errors). Two recent LSM studies found an association between AOS and damage to precentral and postcentral gyri (Basilakos, Rorden, Bonilha, Moser, & Fridriksson, 2015; Itabashi et al., 2016). Notably, Wilson et al.’s PPA study and the apraxia of speech studies quantified segmental (phonological/phonetic) errors in the context of connected speech (e.g., picture description). In contrast, LSM studies of phonological errors during single-word production (picture naming) in post-stroke aphasia have found that these errors are associated with damage to inferior parietal (supramarginal gyrus) and somatosensory (postcentral gyrus) regions (Dell, Schwartz, Nozari, Faseyitan, & Coslett, 2013; Mirman, Chen, et al., 2015; Mirman, Zhang, Wang, Coslett, & Schwartz, 2015; Schwartz, Faseyitan, Kim, & Coslett, 2012).

The picture that emerges is that speech distortion errors localize to regions (frontal and/or central) that lie anterior to the regions associated with phonological errors (anterior parietal and/or posterior temporal). This difference may reflect different stages of articulatory planning, as would be predicted by neurocomputational models of speech production such as the Hierarchical State Feedback Model (Hickok, 2012; see also Walker & Hickok, 2015) and the DIVA model (Bohland, Bullock, & Guenther, 2010; Tourville & Guenther, 2011). These models localize segment-level articulatory planning closer to the central sulcus compared to higher levels of planning, which consist of more posterior auditory syllable and lexical target components (posterior superior temporal gyrus) and more anterior motor syllable planning (inferior frontal gyrus). However, the existing data may also reflect task

differences because the speech distortions were assessed in the context of connected speech whereas the phoneme substitutions were assessed in a single-word production task. To address this, the present study examined segment-level articulatory planning and execution errors in the context of single word production (i.e., without the need to produce connected speech) and with reduced lexical-semantic demands (a word repetition task).

LSM studies of sentence-level speech production deficits have similarly used clinical measures such as WAB Fluency, mean length of utterance, and words per minute (e.g., Basilakos et al., 2014; Catani et al., 2013; Rogalski et al., 2011) or composite scores that include multiple aspects of sentence production (e.g., den Ouden et al., 2019; Mandelli et al., 2014; Rogalski et al., 2011; Wilson et al., 2010). These studies consistently implicated damage to frontal regions (particularly the inferior frontal gyrus) and the underlying white matter (particularly the anterior segment of the arcuate fasciculus, the superior longitudinal fasciculus, and the frontal aslant tract) as the key neural correlates of sentence-level production deficits. However, as a result of using clinical and composite measures, these studies did not make some important psycholinguistic distinctions. First, measures such as utterance length and speech rate (words per minute) are dependent on articulatory agility as well as sentence planning, so they integrate rather than distinguish between articulatory-motor and sentence planning deficits. Second, these measures do not distinguish between structural aspects of sentence planning (i.e., converting a holistic message representation into a sequence of words) and morphosyntactic aspects (i.e., appropriate use of function or closed-class words and bound grammatical morphemes). Structural and morphological deficits have been dissociated behaviorally, suggesting that they are also supported by distinct cognitive processes or systems (Rochon, Saffran, Berndt, & Schwartz, 2000), though (to our knowledge) the neural correlates of this distinction have not been examined with lesion-symptom mapping.

To address this gap, the present study examined two measures of grammatical processing in sentence production that distinguish between structural aspects of sentence production (proportion of words in sentences) and morphological aspects (proportion of closed class words). These measures also minimize the influence of working and short-term memory and the influence of phonological or articulatory factors because these measures are neither dependent on production of long sentences nor on a fast speech rate. This reduction of WM/STM and articulatory agility demands allows us to better isolate sentence-level planning processes, which should rely on mid-anterior frontal regions if they are similar to other types of sequential action planning (e.g., Botvinick, 2008) and/or temporo-parietal regions if they are similar to other types of event or thematic processing (Bedny, Dravida, & Saxe, 2014; Mirman, Landrigan, & Britt, 2017; Thothathiri, Kimberg, & Schwartz, 2012).

The present study had as its major goal to produce a more focused and theoretically relevant characterization of the subcomponents of fluency than has been done to date. To this end, we examined the neural correlates of both segment-level and sentence-level speech production deficits using refined measures that minimize the influence of other processes, are rooted in prior theoretical, computational, and behavioral research, and have proven utility for producing clinical dissociations (e.g., Galluzzi, Bureca, Guariglia, & Romani, 2015; Rochon et al., 2000).

The second goal of the present study was to re-evaluate prior LSM research on fluency deficits using new behavioral measures and LSM methods. There has been growing concern across the behavioral and biological sciences about increasing rigor and reproducibility, but direct replication of a LSM study – which would require large-scale behavioral and neuroimaging testing of people with acquired language (or other cognitive) deficits – is not feasible for practical or financial reasons. Insofar

as the present results converge with prior studies, they provide critical converging evidence across laboratories that helps to establish which patterns are robust enough to emerge in different participant samples and with somewhat different measures and methods. Further, instead of standard mass-univariate voxel-based lesion-symptom mapping, we used voxel-level support vector regression lesion-symptom mapping (SVR-LSM) (Zhang, Kimberg, Coslett, Schwartz, & Wang, 2014). SVR-LSM is a multivariate lesion-symptom mapping method that is particularly well-suited to studying symptoms that may have multiple causes because it can detect independent contributions from distinct brain regions.

In sum, we applied multivariate LSM methods to precisely defined behavioral deficits at the segment-level and sentence-level of production in order to provide a more precise picture of the neural system that is required for fluent language production.

METHODS

Participants and Lesion Data

The data were drawn from an ongoing large-scale study of language processing following left hemisphere stroke. Analyses of other language deficits in earlier subsets of the participants have been reported in several previous articles (Chen, Middleton, & Mirman, 2018; Mirman, Zhang, et al., 2015; Schwartz et al., 2009, 2012; Thothathiri et al., 2012), which also provide more detailed descriptions of the participants and imaging methods. The study was carried out in accordance with protocols approved by the Institutional Review Boards at the Einstein Healthcare Network and University of Pennsylvania School of Medicine. The participants were survivors of left hemisphere stroke (not bilateral or solely subcortical) who had active aphasia or had recovered from clinically significant aphasia but continued to report language deficits. All had English as their first language, were right-handed prior to stroke, were able to produce at least one correct response on the Philadelphia Naming Test (PNT; Roach, Schwartz, Martin, Grewal, & Brecher, 1996), and passed a pure tone audiometric screening to verify sufficient hearing to perform the word repetition task. Participants were tested outside the acute phase, at least 2 months post onset, with almost all tested in the chronic phase¹. The sample included a wide range of aphasia sub-types, aphasia severity based on the WAB Aphasia Quotient, and WAB Fluency sub-scores. A subset had co-occurring apraxia of speech (AOS), based on the Apraxia Battery for Adults test (Dabul, 2000). Apraxia of speech is considered a disorder of high-level speech motor programming, affecting fluency at the phonetic level of speech production. Participants with significant peripheral dysarthria were excluded from this study in order to be able to rule out peripheral speech-motor control causes of fluency deficits. Table 1 provides further clinical and demographic details.

Lesion location was assessed based on MRI or CT brain scans, following the same procedures as previous studies of this data set (or sub-sets of these data). For the MRI scans, lesions were manually segmented on each participant's T1-weighted structural image, then the structural scans and lesion maps were registered to the Montreal Neurological Institute space Colin27 template by an automated process (Avants, Schoenemann, & Gee, 2006). For the CT scans, the lesion was drawn directly onto the Colin27 template after rotating it (pitch only) to match the approximate slice plane of the participant's

¹ For each analysis, more than 85% of participants were at least 6 months post onset (99/115 for the articulatory deficits analysis, 40/46 for the grammatical deficits analysis). Excluding participants who were less than 6 months post onset did not substantively change the results, so the more inclusive analyses are reported here. Also, the critical deficit measures used in this study were not significantly (or even marginally) correlated with either age or months post onset.

scan. Both the coding of the behavioral data and the lesion drawing were done by trained individuals who were blind to the hypotheses tested here. Lesion coverage was good throughout the left MCA territory, particularly the dorsal speech production system structures of the frontal lobe and inferior parietal lobe (Figure 1).

Quantifying Articulatory Deficits

To quantify segment-level articulatory difficulties under minimally demanding speaking conditions, we measured phonetic errors in single word repetition. A similar measure was previously used to identify articulatory impairments that have consequences for *phonological* error production (Galluzzi et al., 2015; Romani & Galluzzi, 2005; Romani, Olson, Semenza, & Granà, 2002). These studies, conducted with Italian-speaking individuals with aphasia, analyzed the phonological errors of participants who scored low or high on the articulatory-deficit measure and found that only the articulatory group showed a consistent bias toward CV-structure simplification. They concluded that there are two distinct loci for phonological error production: articulatory deficit and failed lexical-phonological retrieval (for a related account see Goldrick & Rapp, 2007). Phonetic errors are a key hallmark of apraxia of speech, but more traditional measures of apraxia of speech impose additional demands (such as complex sequential processing), so phonetic errors in word repetition provides a more focused characterization of articulatory deficits. This background establishes PEP as a measure of articulatory deficit that is relevant to both linguistic theory and clinical classification.

Using participants' data from the 175-item *Philadelphia Repetition Test (PRT)* (Dell, Martin, & Schwartz, 2007) we identified responses containing phonetic errors, defined as initial sound struggle, intrusive schwa, impaired prosody, episodic issues with nasal resonance, distorted articulation, or sound combinations that violated English phonotactics. Such phonetic errors were counted regardless of whether they occurred in the context of a correct target word attempt or an incorrect response. The following were *not* counted as phonetic errors: (1) dialectal or regional accent influences that create stretched vowels, or shortened diphthongs; (2) responses spoken with upward intonation, as if to confirm what was heard; (3) undistorted sound substitutions (in order to isolate articulatory distortion from phonological errors). For each participant, the number of target words repeated with one or more phonetic errors was divided by the total number of target words (175) to arrive at a phonetic error proportion (PEP).

Two Speech Pathologists were trained by jointly coding all 175 PRT items from 10 participants. For each participant, the coders began by listening to the first 10-20 responses to get an idea of the speaker's individual speaking characteristics. Following training, inter-coder reliability was assessed using 30 participant PRTs. An independent investigator who was not involved in the coding selected a subset of 35 responses (20%) from each of 30 participant PRTs. Each of the two Speech Pathologists applied the phonetic coding scheme independently to the same 30 subsets of 35 responses. Inter-rater agreement on this subset of 1050 trials from 30 participants was 94.3%. Once coding reliability was established, the remaining responses from these 30 participants and full PRT responses for the remaining participants were divided between the two Speech Pathologists and independently coded.

Quantifying Grammatical Deficits

Our measures of grammatical deficits derived from the Quantitative Production Analysis (QPA; Rochon et al., 2000; Saffran, Berndt, & Schwartz, 1989), a standardized procedure for eliciting and coding narrative speech production in aphasia. The narratives were produced under instructions to

retell the story of “Cinderella” or another familiar fairy tale, resulting in a narrative speech sample that included at least 150 words after removal of repetitions and filled pauses. The speech samples were transcribed and coded by a Speech Pathologist or Research Assistant specifically trained to perform transcription and QPA coding following guidelines based on published QPA procedures, adapted to accommodate fluent as well as nonfluent participants. During training, coders achieved approximately 90% agreement on their transcription and coding of utterance boundaries, utterance content, and grammatical structure. Each narrative was segmented to highlight word groupings that constitute a sentence (minimally, a noun/pronoun followed by a verb) versus unstructured groupings and isolated words. Multiple grammatical measures were derived, from which we selected two that capture different aspects of grammatical processing while minimizing the influence of articulatory agility and WM/STM demands. Both measures were proportions of total words produced, thus also controlling for the total number of words produced by the participant. In the QPA coding system, only the final occurrence of a repeated word is included (except where repetition is used for emphasis), so these measures should not be distorted by word repetition or perseveration (see Appendix A, section III.G of Saffran et al., 1989).

The first measure was proportion of words in the sample that were closed class, that is, function words such as pronouns, determiners, conjunctions, and prepositions. Along with bound suffixes such as plural /-s/ and past tense /-d/, closed class words are morpho-syntactic elements of grammar, deriving function and meaning from the surrounding open class content (nouns and verbs). We did not measure production of bound morphemes because these are relatively difficult to produce and so may be omitted either as a result of grammatical deficits or as a result of articulatory deficits. Closed class words are less difficult to articulate, so we can be more confident that their omission is tied to grammatical deficits.

The second measure of grammatical deficits was proportion of words falling within sentence boundaries. This measure captures each individual’s ability to produce the absolute minimum of English sentence structure – a noun and a verb – without requiring the individual to produce long or complex sentences. Thus, proportion of words in sentences provides a measure of the structural aspects of grammar that minimizes WM/STM constraints. Other measures of sentence structure, such as production of sentences with embedded clauses, require that participants produce relatively long sentences. Such long and complex sentences require working memory and short-term memory to maintain the sentence structure and articulatory ability to produce a long utterance. In contrast, producing a minimal two-word noun-verb sentence minimizes those demands. To further isolate the contribution of sentence structure deficits, follow-up analyses controlled for effects of syntactic processing in comprehension (reversible sentence comprehension: Thothathiri et al., 2012), single word retrieval (picture naming accuracy from the PNT), and articulatory ability (PEP).

Support vector regression lesion-symptom mapping

Lesion-symptom mapping analyses were performed using support vector regression (SVR-LSM) (Zhang et al., 2014). SVR-LSM leverages a multivariate machine learning algorithm to discover lesion-behavior relationships. Compared to standard mass-univariate voxel-based lesion-symptom mapping methods, SVR-LSM is better able to capture independent contributions of multiple brain regions to performance and is less sensitive to differences in statistical power that arise from differences in proportion of participants with lesions in each voxel. These advantages are particularly important for the present study because fluency appears to depend on multiple cognitive processes that may have distinct neural bases. As a standard pre-processing step for SVR-LSM, each participant’s voxel-wise lesion vector

was normalized by dividing each voxel's binary lesion status value by the square root of the total lesion volume. This also serves as a control for the impact of lesion volume, referred to as "direct total lesion volume control" (Zhang et al., 2014). SVR-LSM requires setting two free parameters (cost and gamma), which must be done in a principled way that is independent of the researchers' hypotheses. For each of the three main analyses, these parameters were selected based on 5-fold cross-validation to maximize prediction accuracy (as recommended in Zhang et al., 2014). Follow-up analyses used the same parameters as the corresponding main analysis.

In addition to controlling for lesion volume, the SVR-LSM analyses only included voxels with sufficient lesion involvement (Sperber & Karnath, 2017). The minimum lesion involvement recommended by Sperber and Karnath is 5% of the overall sample, though many studies use a 10% threshold. For the articulatory deficit (PEP) analysis (N=115), only voxels where at least 10 participants had lesions were included in the analysis. For the grammatical deficits analyses, the sample was substantially smaller (N=46), so the minimum lesion involvement was reduced to 5 participants. SVR-LSM produces a voxel-wise map of raw regression β values. Statistical significance for the β values was calculated using a permutation test (2000 permutations) and corrected at false discovery rate (FDR; Genovese, Lazar, & Nichols, 2002; Zhang et al., 2014) $q < 0.05$. The final results include only voxels that passed the FDR threshold, were in the top 5% of raw β values, and comprised clusters larger than 50 voxels².

RESULTS

Articulatory Deficits: PEP

Phonetic error proportion (PEP) was the primary behavioral measure of articulatory deficits, with $M = 0.10$, $SD = 0.12$, and $range = 0.00 - 0.545$ (the full distribution of this measure is shown in the left panel of Figure 2). PEP was fairly strongly correlated with aphasia severity (WAB AQ: $r = -0.532$, $p < 0.0001$) and WAB Fluency scores ($r = -0.508$, $p < 0.0001$). This sample of 115 participants included 32 participants with AOS, as defined by the Dabul Apraxia of Speech Battery (Dabul, 2000). Since speech sound distortions weight heavily in the diagnosis of AOS, it is unsurprising that participants diagnosed with AOS had significantly higher PEP (logistic regression: Estimate = 1.65, SE = 0.049, $z = 33.61$, $p < 0.0001$; AOS: $M = 0.224$, 95% CI = 0.213-0.235; no AOS: $M = 0.052$, 95% CI = 0.049-0.056). Even after controlling for WAB AQ, PEP was significantly associated with WAB Fluency and AOS (both $p < 0.05$). There was also a fairly strong correlation between PEP scores and phonological errors, as measured by the proportion of nonwords on the Philadelphia Naming Test ($r = 0.54$, $p < 0.0001$). This, too, was expected, given the well-known co-occurrence of phonetic and phonological errors in AOS³.

SVR-LSM (cost = 1, gamma = 1.5; prediction accuracy: $r = 0.481$, $p < 0.001$) of phonetic error proportion revealed that this measure of articulatory deficits was most strongly associated with damage to the mid-posterior portion of the dorsal speech stream (Figure 3A, red voxels): the postcentral gyrus

² There is no agreed-upon method of multiple comparisons correction for multivariate LSM and standard permutation-based family-wise error rate control would be redundant because the p -values themselves are calculated by permutation. FDR is currently the most appropriate available method. For more discussion see Mirman (2018; <http://mindingthebrain.blogspot.com/2018/04/correcting-for-multiple-comparisons-in.html>).

³ Naming responses by participants with AOS were scored leniently: allowing one phoneme substitution, omission, or addition for a target-related response. For example, producing "girapp" for GIRAFFE would be considered a correct response. As a result, phonetic errors of the sort that count in the PEP score would not, in the context of picture naming, contribute to the count of nonword errors on the Philadelphia Naming Test.

and the inferior parietal lobule (primarily supramarginal gyrus). This result is very similar to a prior mass-univariate VLSM analysis of the tendency to produce phonological errors in picture naming (Schwartz et al., 2012). As noted above, we calculated phonological errors in the manner of the prior study (proportion nonword responses on the PNT) for the same 115 participants included in this PEP analysis and mapped the association between these errors and lesion location using SVR-LSM. The results revealed that the lesion correlates for phonetic errors in word repetition (PEP) and phonological errors in picture naming were almost perfectly overlapping (Figure 3A, blue voxels correspond to phonological errors in picture naming, purple voxels correspond to the overlap). SVR-LSM of PEP controlling for phonological error proportion left no voxels that survived FDR correction. That is, phonetic errors in word repetition and nonword errors in picture naming seem to have largely the same lesion correlates: the mid-posterior portion of the dorsal speech system. Two possible reasons, not mutually exclusive, are that many phonological errors may be phonetically inspired (Galluzzi et al., 2015; Romani & Galluzzi, 2005), or that the neural networks for the phonological and phonetic encoding of speech segments may overlap too strongly to be disentangled with present methods.

Grammatical Deficits: Proportion of Closed Class Words

The overall mean proportion of words in a participant's sample that were closed class was 0.48 ($SD = 0.098$, $range = 0.18 - 0.61$; the full distribution of this measure is shown in the middle panel of Figure 2). The proportion of closed class words was significantly correlated with aphasia severity (WAB AQ: $r = 0.417$, $p < 0.01$), WAB Fluency score ($r = 0.497$, $p < 0.001$), and words per minute during the narrative production task ($r = 0.500$, $p < 0.001$). After controlling for aphasia severity, proportion of closed class words was still significantly associated with WAB Fluency and words per minute (both $p < 0.05$).

SVR-LSM (cost = 1, gamma = 1.5; prediction accuracy: $r = 0.406$, $p < 0.01$) of proportion of closed class words revealed no voxels that survived FDR correction, indicating that we were not able to detect a consistent lesion pattern that was associated with morpho-syntactic deficits measured by production of closed class words.

Grammatical Deficits: Proportion of Words in Sentences

The overall mean proportion of words in a participant's sample that were within sentence boundaries was 0.70, with variation between individuals covering essentially the full range ($SD = 0.30$, $range = 0.03 - 1.0$; the full distribution of this measure is shown in the right panel of Figure 2). The proportion of words in sentences was significantly correlated with aphasia severity (WAB AQ: $r = 0.485$, $p < 0.001$), WAB Fluency score ($r = 0.571$, $p < 0.001$), and words per minute during the narrative production task ($r = 0.484$, $p < 0.001$). After controlling for aphasia severity, proportion of words in sentences was still significantly associated with WAB Fluency and words per minute (both $p < 0.05$).

Figure 3B shows the SVR-LSM (cost = 1, gamma = 1; prediction accuracy: $r = 0.646$, $p < 0.001$) results for proportion of words in sentences, which revealed that this measure of grammatical deficits was primarily associated with frontal lobe damage. In particular, reduced proportion of words in sentences was most strongly associated with damage to the middle frontal gyrus and inferior frontal gyrus (primarily pars triangularis, with substantial involvement of pars orbitalis, and very little extension into pars opercularis). There was also a smaller cluster of voxels in the postcentral gyrus and inferior parietal lobule.

Following prior reports that fluency deficits are associated with damage to frontal white matter (Basilakos et al., 2014; Bonilha & Fridriksson, 2009; Catani et al., 2013; Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007; Fridriksson, Guo, Fillmore, Holland, & Rorden, 2013; Halai et al., 2017; Wilson et al., 2010), we tested overlap with the frontal aslant tract and the anterior segment of the arcuate fasciculus but found only minimal overlap with these white matter tracts. The white matter tract regions were defined using the atlas developed by Catani and colleagues⁴ (Catani, Dell'Acqua, Bizzi, et al., 2012; Catani, Dell'Acqua, Vergani, et al., 2012; Catani & de Schotten, 2012) using a 50% threshold (i.e., regions that corresponded to those white matter tracts in at least 50% of neurologically typical participants). For both the anterior segment of the arcuate fasciculus and the frontal aslant tract, there was less than 5% overlap between the white matter tract and the region identified by SVR-LSM (calculated either as a proportion of the white matter tract or as a proportion of the SVR-LSM region).

As described above, we chose proportion of words in sentences as a measure of grammatical deficits because it reflects the ability to structure sentences and is minimally influenced by other cognitive and articulatory abilities. Nevertheless, articulatory and word retrieval deficits may contribute to reduced proportion of words in sentences. To address this, we conducted follow-up analyses of proportion of words in sentences controlling for PEP as a measure of articulatory deficits, and controlling for picture naming accuracy (from the PNT) as a measure of word retrieval ability. The SVR-LSM results did not change substantially: after FDR correction, the remaining voxels were primarily in inferior and middle frontal gyri, with some extension in the precentral and postcentral gyri.

To distinguish sentence structure processing deficits in production from related deficits in comprehension, we conducted a follow-up analysis of proportion of words in sentences controlling for comprehension of reversible sentences. Reversible sentence comprehension scores came from a sentence-to-picture matching task (Breedin & Saffran, 1999) in which sentences such as “The man serves the woman” were presented with a picture of a man serving a woman (target) and a picture of a woman serving a man (distractor). Correct responses in this task require correctly assigning thematic roles in the sentences (i.e., identifying the actor/agent and the recipient/patient of the action). Prior studies have found that impaired comprehension of such sentences is associated with damage to the temporo-parietal cortex (Race, Ochfeld, Leigh, & Hillis, 2012; Thothathiri et al., 2012). Our follow-up analysis showed that controlling for comprehension of reversible sentences did not change the pattern of lesion correlates of proportion of words in sentence: the primary lesion correlates were still in the inferior and middle frontal gyri, though with greater extension into the superior frontal gyrus, and the precentral and postcentral gyri.

The results of all of the SVR-LSM analyses are summarized in Table 2.

Alternative Analyses

Because SVR-LSM is a relatively new analysis technique, best practices in its implementation are still under investigation. A recent reimplementaion (DeMarco & Turkeltaub, 2018) offers alternative methods of lesion volume control and multiple comparisons correction. We re-analyzed the data using this implementation, controlling for lesion volume by regressing it out of both behavioral data and raw lesion data (“Regress on Both”) and correcting for multiple comparisons using cluster size correction based on 10,000 permutations. The results were broadly the same as reported above, with only slight differences in some of the follow-up analyses. PEP was associated with a significant cluster in the postcentral gyrus and inferior parietal lobule, which was also significant after controlling for

⁴ Available from the Natbrainlab: <http://www.natbrainlab.co.uk/atlas-maps>.

phonological (nonword) errors in picture naming. Analysis of proportion of closed class words identified no significant clusters after correction. Proportion of words in sentences was associated with a significant cluster in the frontal lobe, which remained significant after controlling for picture naming accuracy (PNT) and reversible sentence comprehension, but not after controlling for PEP.

Cluster-level correction may not be a good strategy for SVR-LSM because it tends to produce single over-large clusters (Mirman et al., 2018), which will tend to overestimate the contribution of a single region and underestimate the contributions of other regions. Therefore, we also recomputed FDR-corrected thresholds using voxel-wise p-values from the 10,000 permutations, but those results were substantially less conservative than our initial analyses. Finally, in order to eliminate incorrect estimation of $p=0$ for some voxels in our original 2000-permutation analyses, we “regularized” those p-values (for a discussion of regularization see (Donnelly & Verkuilen, 2017)) by shifting them away from that floor level by either a fixed amount (0.5/2000) or by random sampling. This produced FDR thresholds that were virtually identical to the original calculation, suggesting that the presence of $p=0$ values did not substantially distort the FDR calculation. In sum, alternative strategies for lesion volume correction and multiple comparisons correction produced substantively the same patterns of results, indicating that the observed patterns are relatively robust with respect to details of the analysis strategy.

DISCUSSION

Fluent speech production requires the rapid coordination of a wide range of cognitive processes, including executive functions and working/short-term memory, grammatical knowledge, semantic memory, and articulatory planning and execution. There is broad agreement among prior studies that fluent speech production is supported by fronto-parietal brain regions. The present study provides a finer-grain investigation of the sub-components of this system by using behavioral measures of articulatory and grammatical deficits that minimize the influence of other cognitive processes, and by using a multivariate lesion-symptom mapping method (SVR-LSM), which can more effectively detect contributions of multiple brain regions.

Articulatory deficit was measured by production of phonetic distortion errors (such as distorted articulation, intrusive schwa, impaired prosody, and violations of English phonotactics) produced in a word repetition task. The single word repetition task minimizes lexical-semantic and higher-level cognitive demands, and evaluating phonetic distortions allowed us to evaluate articulatory planning and execution aspects. Individuals with significant peripheral articulatory deficits (dysarthria) were excluded, so this measure effectively isolated central articulatory planning and execution deficits. The SVR-LSM results revealed that these deficits were associated with damage to the postcentral gyrus and the inferior parietal lobule, overlapping very strongly with lesion correlates of production of phonological errors in picture naming. This finding differs from other studies that suggested a more anterior locus of articulatory deficits in apraxia of speech (Basilakos et al., 2015; Itabashi et al., 2016) or based on assessments of errors in connected speech (e.g., Wilson et al., 2010). Assessing articulatory deficits in the context of connected speech may introduce frontal lobe sentence-level contributions to fluent speech production. Similarly, although apraxia of speech is characterized by articulatory deficits, it strongly co-occurs with non-fluent forms of aphasia, so clinical definitions of AOS do not isolate articulatory deficits. A recent study found that single-word repetition errors were associated with posterior peri-Sylvian damage whereas articulatory deficits in spontaneous speech were associated

anterior insula and inferior frontal damage (Ripamonti et al., 2018). That is, LSM results for articulatory deficits in prior studies may have been shifted anteriorly by contributions from other, co-occurring fluency deficits. By measuring articulatory deficits in a word repetition task, the present study minimized such effects and more precisely identified the postcentral and inferior parietal lesion correlates of articulatory deficits as a subcomponent of apraxia of speech.

Grammatical deficits were measured by proportion of closed class words (morpho-syntax) and proportion of words in sentences (sentence structure) produced in the context of a semi-structured narrative (telling the Cinderella story). These measures minimize executive function and WM/STM demands because they do not require production of long sentences or complex grammatical structures and they minimize articulatory demands because they do not require a fast speech rate or production of difficult bound morphemes. Thus, these measures effectively isolate morpho-syntactic and sentence structuring aspects of grammatical deficits. Proportion of closed class words did not have consistent lesion correlates. The cross-validation accuracy during SVR-LSM parameter optimization was statistically significant, indicating that there was some systematic relationship between lesion pattern and proportion of closed class words; however, no voxels survived FDR correction for multiple comparisons, so it is not clear *which* aspects of the lesion pattern were associated with this grammatical deficit measure. Proportion of words in sentences was strongly associated with frontal lobe damage, particularly the inferior and middle frontal gyri. This pattern remained after controlling for articulatory deficits, word retrieval ability, and reversible sentence comprehension. This pattern converges with a study of PPA that dissociated fluency from grammatical deficits and found that fluency deficits (MLU) were associated with middle frontal gyrus degeneration whereas grammatical deficits were associated with distributed fronto-parietal degeneration (Rogalski et al., 2011).

Unlike some prior studies, we did not find that sentence-level fluency deficits were associated with frontal white matter damage. Two factors may have contributed to this difference between our results and previous studies. First, some of those prior studies used broader measures of fluent production such as mean (or median) length of utterance (Catani et al., 2013), WAB Fluency (Basilakos et al., 2014; Fridriksson et al., 2013), and a composite “speech quanta” score (Halai et al., 2017). We used proportion of words in sentences in order to more precisely isolate the role of sentence structure deficits from other articulatory and cognitive deficits. The prior findings that frontal white matter damage is associated with fluency deficits may reflect that impaired integration or coordination of multiple sub-systems contributes to utterance lengths, WAB Fluency scores, and other such broad measures of fluency. Second, the prior studies used standard, mass-univariate VLSM, which is susceptible to mislocalization of grey matter effects into the underlying white matter (Mah, Husain, Rees, & Nachev, 2014), though that mislocalization may be mitigated by common LSM best-practices (Sperber & Karnath, 2017). SVR-LSM is less susceptible to this artifact (Zhang et al., 2014), while maintaining sensitivity to effects of white matter damage (Griffis, Nenert, Allendorfer, & Szaflarski, 2017; Mirman, Zhang, et al., 2015).

Building on prior findings, the present results provide a clearer picture of the neural system that supports spoken language production and how it is impaired following left hemisphere stroke (Figure 4). This system has three main components. The first is sentence structuring: fluent narrative speech production requires the ability to structure grammatically complete descriptions of events or propositional statements involving, minimally, noun plus verb. This ability seems to critically involve the left inferior and middle frontal gyri and aligns with the neural basis of other types of sequential action planning (e.g., Botvinick, 2008). The second component is populating the sentence structure with actual

words; that is, semantically-driven word retrieval. This component was not examined in the present study, but the results of numerous previous studies converge to identify the anterior temporal lobe as the critical region for semantically-driven word retrieval (Chen et al., 2018; Lambon Ralph, McClelland, Patterson, Galton, & Hodges, 2001; Mesulam et al., 2013; Schwartz et al., 2009; Walker et al., 2011), with possible additional roles for the uncinate fasciculus (Catani et al., 2013) and inferior frontal regions (Chen et al., 2018; Harvey & Schnur, 2015; Mirman & Graziano, 2013; Schnur et al., 2009). The third component is phonetic and phonological planning and execution of those words, which relies primarily on inferior parietal regions. Motor systems are, no doubt, also important for articulatory motor control, but the phonetic and phonological aspects of articulation appear to rely more heavily on the action planning systems of inferior parietal lobule and the postcentral gyrus system, presumably for online somatosensory and auditory monitoring of articulation (Hickok, 2012; Rauschecker, 2011; Rogalsky et al., 2015).

Fluent speech production also requires working memory and related executive functions in order to formulate a coherent narrative, maintain it, and update the speaker's place in that narrative. Working memory and executive functions can be impaired in aphasia (Halai et al., 2017; Lacey et al., 2017), which could produce a different sort of impairment in narrative speech production (and comprehension). In the present study, we have sought to identify distinct components of the speech production system so we used measures that had minimal demands on other components, but fluent speech production requires coordination among all of these components. Further, it is likely that these components interact during speech production. For example, articulatory deficits may influence lexical selection (speaker may prefer words that are easier to articulate), grammatical production (omission of bound morphemes), and sentence structure (speaker may prefer shorter sentences). Understanding of the system's components should be complemented with understanding of how those components interact.

Clinical measures of fluent language production, such as the WAB Fluency sub-score, mean (or median) length of utterance, words per minute, etc., typically do not distinguish between different sources of non-fluent speech production. As a result, they can reflect deficits in any of these components, or multiple components (Caplan, 2012). The present results show that these components have quite distinct neural substrates. In addition to providing new insights into the neural basis of spoken language production, these results suggest that diagnosis and treatment of post-stroke non-fluent language deficits should carefully consider whether the non-fluency is arising from deficits at grammatical, lexical-semantic, or phonological/phonetic levels. Both finer-grained behavioral assessment and consideration of precise lesion location can help to localize the deficit. Similarly, treatments that improve grammatical processing are unlikely to improve phonological retrieval and encoding, and vice versa. Therefore, treatment research and selection should distinguish between sources of non-fluency.

ACKNOWLEDGEMENTS

We thank Danielle Brown, Casey Ferrara, and Claire Healy for their help with coding and preparing the data for analysis. We also thank Marco Catani and Stephanie Forkel for providing their maps of the frontal aslant tract, John Medaglia for his help with the white matter overlap analyses, and Melissa Thye for her help with data visualization. This research was funded by NIH grants R01DC010805 to DM and R01DC000191 to MFS, and supported by the Moss Rehabilitation Research Institute, Drexel University, and the University of Alabama at Birmingham.

OPEN PRACTICES STATEMENT

Due to privacy restrictions, none of the data for the studies reported here are available. The SVR-LSM analyses were implemented using the SVR-LSM toolbox (<https://github.com/dmirman/SVR-LSM>) and GUI implementation (<https://github.com/atdemarco/svrlsmgui>), both of which are publicly available. None of the experiments was preregistered.

REFERENCES

- Avants, B. B., Schoenemann, P. T., & Gee, J. C. (2006). Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Medical Image Analysis, 10*, 397–412.
- Basilakos, A., Fillmore, P. T., Rorden, C., Guo, D., Bonilha, L., & Fridriksson, J. (2014). Regional White Matter Damage Predicts Speech Fluency in Chronic Post-Stroke Aphasia. *Frontiers in Human Neuroscience, 8*(October), 1–9. <http://doi.org/10.3389/fnhum.2014.00845>
- Basilakos, A., Rorden, C., Bonilha, L., Moser, D., & Fridriksson, J. (2015). Patterns of Poststroke Brain Damage That Predict Speech Production Errors in Apraxia of Speech and Aphasia Dissociate. *Stroke, 46*(6), 1561–1566. <http://doi.org/10.1161/STROKEAHA.115.009211>
- Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective & Behavioral Neuroscience, 14*(3), 891–901. <http://doi.org/10.3758/s13415-013-0217-z>
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22*(7), 1504–29. <http://doi.org/10.1162/jocn.2009.21306>
- Bonilha, L., & Fridriksson, J. (2009). Subcortical damage and white matter disconnection associated with non-fluent speech. *Brain, 132*(6), 2–3. <http://doi.org/10.1093/brain/awn200>
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences, 12*(5), 201–208. <http://doi.org/10.1016/j.tics.2008.02.009>
- Breedin, S., & Saffran, E. M. (1999). Sentence processing in the face of semantic loss: a case study. *Journal of Experimental Psychology: General, 128*(4), 547–562.
- Caplan, D. (2012). Aphasic Syndromes. In K. M. Heilman & E. Valenstein (Eds.), *Clinical Neuropsychology* (Fifth Edit, pp. 22–41). New York, NY, USA: Oxford University Press.
- Catani, M., & de Schotten, M. T. (2012). *Atlas of Human Brain Connections*. Oxford University Press.
- Catani, M., Dell'Acqua, F., Bizzi, A., Forkel, S. J., Williams, S. C., Simmons, A., ... Thiebaut de Schotten, M. (2012). Beyond cortical localization in clinico-anatomical correlation. *Cortex, 48*(10), 1262–1287. <http://doi.org/10.1016/j.cortex.2012.07.001>
- Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., ... Thiebaut de Schotten, M. (2012). Short frontal lobe connections of the human brain. *Cortex, 48*(2), 273–291. <http://doi.org/10.1016/j.cortex.2011.12.001>
- Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., ... Rogalski, E. J. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain, 136*(8), 2619–2628. <http://doi.org/10.1093/brain/awt163>
- Chen, Q., Middleton, E., & Mirman, D. (2018). Words fail: Lesion-symptom mapping of errors of omission in post-stroke aphasia. *Journal of Neuropsychology*. <http://doi.org/10.1111/jnp.12148>
- Dabul, B. L. (2000). *Apraxia Battery for Adults (ABA-2)* (Second Ed.). Austin, TX: Pro-Ed.
- Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language, 56*(4), 490–520. <http://doi.org/10.1016/j.jml.2006.05.007>

- Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O. K., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition*, *128*(3), 380–396. <http://doi.org/10.1016/j.cognition.2013.05.007>
- DeMarco, A. T., & Turkeltaub, P. E. (2018). A multivariate lesion symptom mapping toolbox and examination of lesion-volume biases and correction methods in lesion-symptom mapping. *Human Brain Mapping*, *39*(11), 4169–4182. <http://doi.org/10.1002/hbm.24289>
- den Ouden, D.-B., Malyutina, S., Basilakos, A., Bonilha, L., Gleichgerrcht, E., Yourganov, G., ... Fridriksson, J. (2019). Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Human Brain Mapping*. <http://doi.org/10.1002/hbm.24514>
- Donnelly, S., & Verkuilen, J. (2017). Empirical logit analysis is not logistic regression. *Journal of Memory and Language*, *94*, 28–42. <http://doi.org/10.1016/j.jml.2016.10.005>
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432–1441. <http://doi.org/10.1093/brain/awm042>
- Fridriksson, J., Guo, D., Fillmore, P. T., Holland, A., & Rorden, C. (2013). Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. *Brain*, *136*(11), 3451–3460. <http://doi.org/10.1093/brain/awt267>
- Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, *71*, 64–83. <http://doi.org/10.1016/j.neuropsychologia.2015.03.007>
- Genovese, C. R., Lazar, N. A., & Nichols, T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878. <http://doi.org/10.1006/nimg.2001.1037>
- Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, *102*(2), 219–260. <http://doi.org/10.1016/j.cognition.2005.12.010>
- Griffis, J. C., Nenert, R., Allendorfer, J. B., & Szaflarski, J. P. (2017). Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. *NeuroImage: Clinical*, *14*, 552–565. <http://doi.org/10.1016/j.nicl.2017.02.019>
- Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*, *86*, 275–289. <http://doi.org/10.1016/j.cortex.2016.04.016>
- Harvey, D. Y., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex*, *67*, 37–58. <http://doi.org/10.1016/j.cortex.2015.03.004>
- Hickok, G. S. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135–145. <http://doi.org/10.1038/nrn3158>
- Itabashi, R., Nishio, Y., Kataoka, Y., Yazawa, Y., Furui, E., Matsuda, M., & Mori, E. (2016). Damage to the left precentral gyrus is associated with apraxia of speech in acute stroke. *Stroke*, *47*(1), 31–36. <http://doi.org/10.1161/STROKEAHA.115.010402>
- Kertesz, A. (1982). *Western Aphasia Battery*. New York: Grune & Stratton.
- Lacey, E. H., Skipper-Kallal, L. M., Xing, S., Fama, M. E., & Turkeltaub, P. E. (2017). Mapping Common Aphasia Assessments to Underlying Cognitive Processes and Their Neural Substrates. *Neurorehabilitation and Neural Repair*, *31*(5), 442–450. <http://doi.org/10.1177/1545968316688797>
- Lambon Ralph, M. A., McClelland, J. L., Patterson, K. E., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, *13*(3), 341–356.

- Mah, Y.-H., Husain, M., Rees, G., & Nachev, P. (2014). Human brain lesion-deficit inference remapped. *Brain*, *137*(9), 2522–2531. <http://doi.org/10.1093/brain/awu164>
- Mandelli, M. L., Caverzasi, E., Binney, R. J., Henry, M. L., Lobach, I., Block, N., ... Gorno-Tempini, M. L. (2014). Frontal White Matter Tracts Sustaining Speech Production in Primary Progressive Aphasia. *Journal of Neuroscience*, *34*(29), 9754–9767. <http://doi.org/10.1523/JNEUROSCI.3464-13.2014>
- Mesulam, M.-M., Wieneke, C. A., Hurley, R. S., Rademaker, A., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2013). Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain*, *136*(Pt 2), 601–618. <http://doi.org/10.1093/brain/aws336>
- Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural Organization of Spoken Language Revealed by Lesion-Symptom Mapping. *Nature Communications*, *6*(6762), 1–9. <http://doi.org/10.1038/ncomms7762>
- Mirman, D., & Graziano, K. M. (2013). The Neural Basis of Inhibitory Effects of Semantic and Phonological Neighbors in Spoken Word Production. *Journal of Cognitive Neuroscience*, *25*(9), 1504–1516. http://doi.org/10.1162/jocn_a_00408
- Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and Thematic Semantic Systems. *Psychological Bulletin*, *143*(5), Manuscript in press. <http://doi.org/10.1037/bul0000092>
- Mirman, D., Landrigan, J.-F., Kokolis, S., Verillo, S., Ferrara, C., & Pustina, D. (2018). Corrections for multiple comparisons in voxel-based lesion-symptom mapping. *Neuropsychologia*, *115*, 112–123. <http://doi.org/10.1016/j.neuropsychologia.2017.08.025>
- Mirman, D., & Thye, M. (2018). Uncovering the Neuroanatomy of Core Language Systems Using Lesion-Symptom Mapping. *Current Directions in Psychological Science*.
- Mirman, D., Zhang, Y., Wang, Z., Coslett, H. B., & Schwartz, M. F. (2015). The ins and outs of meaning: Behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. *Neuropsychologia*, *76*, 208–219. <http://doi.org/10.1016/j.neuropsychologia.2015.02.014>
- Race, D. S., Ochfeld, E., Leigh, R., & Hillis, A. E. (2012). Lesion analysis of cortical regions associated with the comprehension of Nonreversible and Reversible yes/no questions. *Neuropsychologia*, *50*(8), 1946–1953. <http://doi.org/10.1016/j.neuropsychologia.2012.04.019>
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, *271*(1–2), 16–25. <http://doi.org/10.1016/j.heares.2010.09.001>
- Ripamonti, E., Frustaci, M., Zonca, G., Aggujaro, S., Molteni, F., & Luzzatti, C. (2018). Disentangling phonological and articulatory processing: A neuroanatomical study in aphasia. *Neuropsychologia*, *121*, 175–185. <http://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2018.10.015>
- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. R. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, *24*, 121–133.
- Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, *72*, 193–218. <http://doi.org/10.1006/brln.1999.2285>
- Rogalski, E. J., Cobia, D., Harrison, T. M., Wieneke, C. a, Thompson, C. K., Weintraub, S., & Mesulam, M.-M. (2011). Anatomy of language impairments in primary progressive aphasia. *The Journal of Neuroscience*, *31*(9), 3344–50. <http://doi.org/10.1523/JNEUROSCI.5544-10.2011>
- Rogalsky, C., Poppa, T., Chen, K. H., Anderson, S. W., Damasio, H., Love, T., & Hickok, G. S. (2015). Speech repetition as a window on the neurobiology of auditory-motor integration for speech: A voxel-based lesion symptom mapping study. *Neuropsychologia*, *71*, 18–27. <http://doi.org/10.1016/j.neuropsychologia.2015.03.012>
- Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology*, *22*(7), 817–850. <http://doi.org/10.1080/02643290442000365>

- Romani, C., Olson, A., Semenza, C., & Granà, A. (2002). Patterns of phonological errors as a function of a phonological versus an articulatory locus of impairment. *Cortex*, *38*(4), 541–567. [http://doi.org/10.1016/S0010-9452\(08\)70022-4](http://doi.org/10.1016/S0010-9452(08)70022-4)
- 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.
- Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, *106*(1), 322–327.
- Schwartz, M. F., Faseyitan, O. K., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, *135*(12), 3799–3814. <http://doi.org/10.1093/brain/aws300>
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O. K., Brecher, A. R., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, *132*(12), 3411–3427. <http://doi.org/10.1093/brain/awp284>
- Sperber, C., & Karnath, H.-O. (2017). Impact of correction factors in human brain lesion-behavior inference. *Human Brain Mapping*, *38*(3), 1692–1701. <http://doi.org/10.1002/hbm.23490>
- Thompson, C. K., Cho, S., Hsu, C.-J., Wieneke, C., Rademaker, A., Weitner, B. B., ... Weintraub, S. (2012). Dissociations between fluency and agrammatism in primary progressive aphasia. *Aphasiology*, *26*(1), 20–43. <http://doi.org/10.1080/02687038.2011.584691>
- 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. http://doi.org/10.1162/jocn_a_00118
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. <http://doi.org/10.1080/01690960903498424>
- Walker, G. M., & Hickok, G. S. (2015). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic Bulletin & Review*. <http://doi.org/10.3758/s13423-015-0903-7>
- Walker, G. M., Schwartz, M. F., Kimberg, D. Y., Faseyitan, O. K., Brecher, A. R., Dell, G. S., & Coslett, H. B. (2011). Support for anterior temporal involvement in semantic error production in aphasia: New evidence from VLSM. *Brain and Language*, *117*(3), 110–122.
- Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., ... Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069–2088. <http://doi.org/10.1093/brain/awq129>
- Zhang, Y., Kimberg, D. Y., Coslett, H. B., Schwartz, M. F., & Wang, Z. (2014). Multivariate lesion-symptom mapping using support vector regression. *Human Brain Mapping*, *35*(12), 5861–5876. <http://doi.org/10.1002/hbm.22590>

TABLE 1: Background information about participants in the main analyses.

	Analysis of articulatory deficits	Analysis of grammatical deficits
N	115	46
Gender	50 Female, 65 Male	19 Female, 27 Male
Scan type	65 MRI, 50 CT	30 MRI, 16 CT
Median age [IQR] (range)	58 [50-65] (26-79)	55 [48-65] (30-75)
Median months post onset [IQR] (range)	19 [7-60] (2-381)	27.5 [10-81] (2-240)
Median WAB AQ [IQR] (range)	74.5 [57.2-88.7] (25.2-97.9)	73.4 [64.0-84.5] (41.4-99.3)
Median WAB Fluency [IQR] (range)	7 [4-9] (1-10)	5.5 [4-8] (2-10)
Aphasia sub-types	46 Anomic 33 Broca 22 Conduction 9 Wernicke 4 TCM 1 Global	20 Anomic 17 Broca 6 Conduction 2 Wernicke 1 TCM

Abbreviations: IQR = Inter-Quartile Range; WAB = Western Aphasia Battery; AQ = Aphasia Quotient; TCM = Transcortical Motor.

TABLE 2: Number of voxels in each region of interest for each analysis where any voxels survived FDR correction. All regions refer to left hemisphere only.

ROI	Words In Sent.	Controlling for PEP	Controlling for PNT	Controlling for Rev. Sentence Comp.	Phonetic Errors (PRT)	Nonword Errors (PNT)
IFG: Pars Triangularis	7325	4696	6959	8734	-	-
IFG: Pars Orbitalis	1311	258	1265	2257	-	-
IFG: Pars Opercularis	50	14	22	24	-	-
MFG	6536	3780	4461	10469	-	-
SFG	1	-	81	2476	-	-
Supp Motor Area	-	-	-	163	-	-
Precentral Gyrus	418	442	1105	5779	98	66
Postcentral Gyrus	1032	324	256	1604	9935	2013
Inferior Parietal Lobule (incl. SMG and AG)	1491	157	66	653	5565	109
Rolandic Operculum	-	-	-	-	990	-
STG	89	135	-	-	989	4
Heschl's Gyrus	-	-	-	-	115	-
Insula	69	30	132	110	-	-
Caudate	41	-	94	-	-	-
Middle Occipital Gyrus	-	-	41	123	1	-
Temporal Pole: Superior	-	7	-	44	-	-
Superior Parietal Gyrus	-	-	-	-	54	-

Abbreviations: IFG = inferior frontal gyrus; MFG = middle frontal gyrus; SFG = superior frontal gyrus; SMG = supramarginal gyrus; AG = angular gyrus; PRT = Philadelphia Repetition Test; PNT = Philadelphia Naming Test.

FIGURE 1. Top: Lesion overlap for all 115 participants in articulatory deficits analyses, thresholded to include only voxels that were lesioned in at least 10 participants (max = 65 participants). **Bottom:** Lesion overlap for all 46 participants in grammatical deficits analyses, thresholded to include only voxels that were lesioned in at least 5 participants (max = 33 participants).

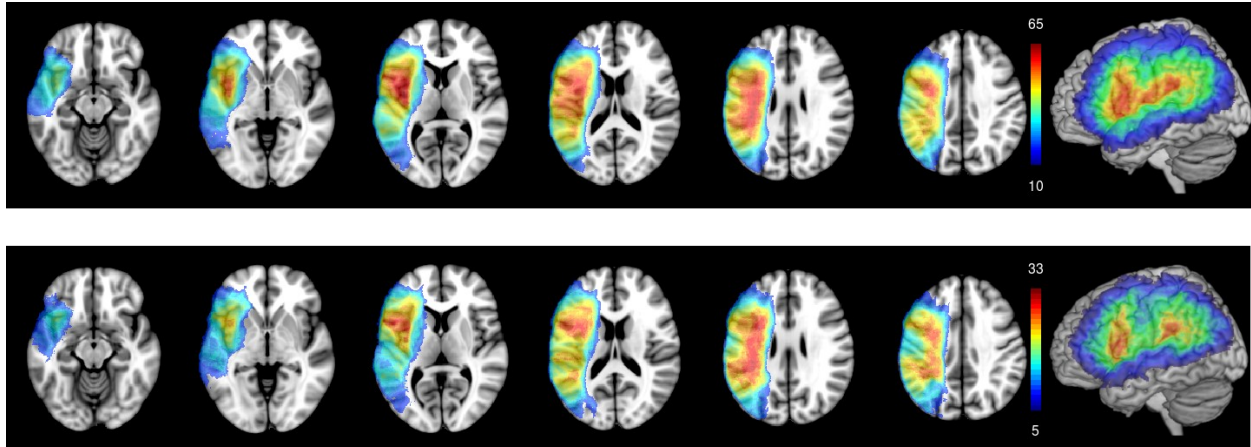


FIGURE 2. Distributions of behavioral scores for primary analyses. **Left:** Articulatory deficits score: phonetic error proportion (PEP). **Middle:** Grammatical deficits score: proportion of closed class words. **Right:** Grammatical deficits score: proportion of words in sentences.

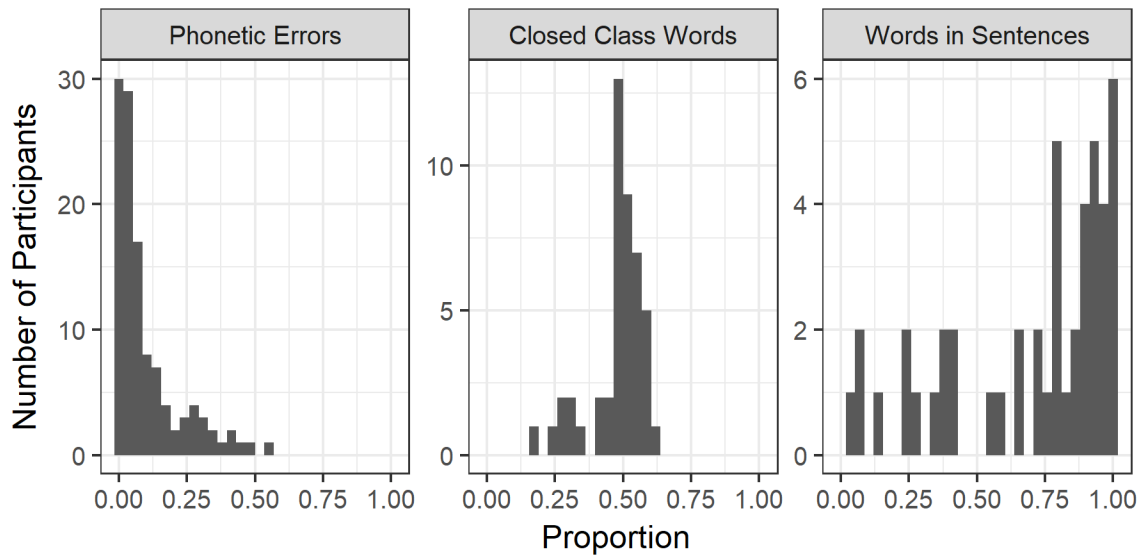


FIGURE 3. (A, top row) Results of SVR-LSM for phonetic error proportions in word repetition (red), nonword errors in picture naming (blue), and their overlap (purple). **(B, bottom row)** Results of SVR-LSM for proportion of words in sentences.

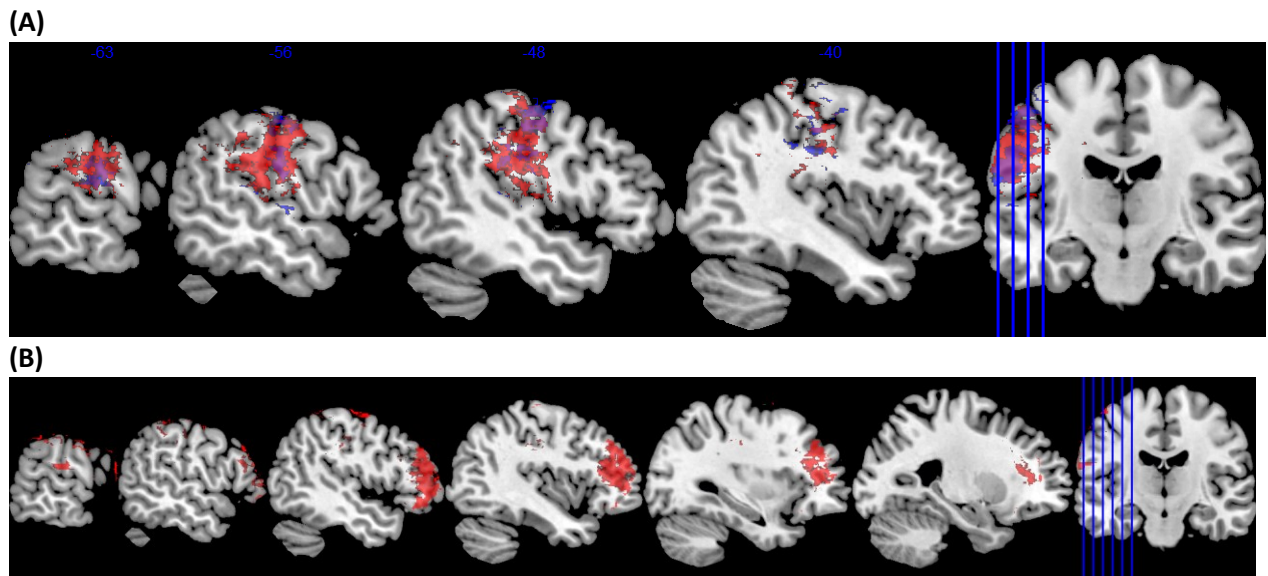


FIGURE 4. Schematic diagram of the three neural components of the fluent speech production system: frontal sentence structure component (red), anterior temporal word retrieval component (green), and parietal phonological planning and execution component (blue).

