UCSF
UC San Francisco Previously Published Works

Title

Discriminating nonfluent/agrammatic and logopenic PPA variants with automatically extracted morphosyntactic measures from connected speech

Permalink

https://escholarship.org/uc/item/5jf0c1cr

Journal

Cortex, 173

ISSN

0010-9452

Authors

Lukic, Sladjana
Fan, Zekai
García, Adolfo M
et al.

Publication Date

2024-04-01

DOI

10.1016/j.cortex.2023.12.013

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cortex. Author manuscript; available in PMC 2024 July 14.

Published in final edited form as:

Cortex. 2024 April ; 173: 34–48. doi:10.1016/j.cortex.2023.12.013.

Discriminating nonfluent/agrammatic and logopenic PPA variants with automatically extracted morphosyntactic measures from connected speech

Sladjana Lukic[1,2], **Zekai Fan**[3], **Adolfo M. García**[4,5,6], **Ariane E. Welch**[2], **Buddhika M. Ratnasiri**[1], **Stephen M. Wilson**[7], **Maya L. Henry**[8], **Jet Vonk**[1], **Jessica Deleon**[1], **Bruce L. Miller**[1], **Zachary Miller**[1], **Maria Luisa Mandelli**[1], **Maria Luisa Gorno-Tempini**[1]

[1]University of California San Francisco Memory and Aging Center, CA, 94158, USA

[2]Ruth S. Ammon College of Education and Health Sciences, Department of Communication Sciences and Disorders, Adelphi University, Garden City, NY 11530, USA

[3]Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[4]Global Brain Health Institute (GBHI), University of California, San Francisco, CA, 94158, USA

[5]Cognitive Neuroscience Center, Universidad de San Andrés, Buenos Aires, B1644BID, Argentina

[6]Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, 9160000, Chile

[7]School of Health and Rehabilitation Sciences, University of Queensland, Brisbane QLD 4072 Australia

[8]University of Texas at Austin Moody College of Communication, Austin, TX 78705, USA

Abstract

Morphosyntactic assessments are important for characterizing individuals with nonfluent/agrammatic variant primary progressive aphasia (nfvPPA). Yet, standard tests are subject to examiner bias and often fail to differentiate between nfvPPA and logopenic variant PPA (lvPPA). Moreover, relevant neural signatures remain underexplored. Here, we leverage natural language processing tools to automatically capture morphosyntactic disturbances and their neuroanatomical correlates in 35 individuals with nfvPPA relative to 10 healthy controls (HC) and 26 individuals with lvPPA.

Participants described a picture, and ensuing transcripts were analyzed via part-of-speech tagging to extract sentence-related features (e.g., subordinating and coordinating conjunctions), verbal-related features (e.g., tense markers), and nominal-related features (e.g., subjective and possessive pronouns). Gradient boosting machines were used to classify between groups using all features.

Corresponding Author: Sladjana Lukic, PhD, slukic@adelphi.edu, Hy Weinberg Center, Suite 119, Garden City, NY 11530.

Conflict of interests

The authors report no disclosures relevant to the manuscript.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We identified the most discriminant morphosyntactic marker via a feature importance algorithm and examined its neural correlates via voxel-based morphometry.

Individuals with nfvPPA produced fewer morphosyntactic elements than the other two groups. Such features robustly discriminated them from both individuals with lvPPA and HCs with an AUC of 0.95 and 0.82, respectively. The most discriminatory feature corresponded to subordinating conjunctions was correlated with cortical atrophy within the left posterior inferior frontal gyrus across groups (p_{FWE}<.05).

Automated morphosyntactic analysis can efficiently differentiate nfvPPA from lvPPA. Also, the most sensitive morphosyntactic markers correlate with a core atrophy region of nfvPPA. Our approach, thus, can contribute to a key challenge in PPA diagnosis.

Keywords

primary progressive aphasia; morphosyntax; subordination production; Natural Language Processing; cortical atrophy

1. Introduction

Primary progressive aphasia (PPA) is a collection of clinical syndromes characterized by progressive language deficits due to the gradual loss of neurons in the dominant left-hemisphere perisylvian regions (Gorno-Tempini et al., 2011; Mesulam, 1987). Each syndrome is typified by distinct linguistic deficit patterns and neuroanatomical disruptions. For example, nonfluent/agrammatic variant PPA (nfvPPA) often presents with predominant morphosyntactic deficits associated with inferior frontal atrophy (see Thompson & Mack, 2014 for a review), while logopenic variant PPA (lvPPA) involves anomia as well as phonological and verbal short-term memory impairments associated with temporo-parietal atrophy (e.g., Gorno-Tempini et al., 2008; Henry et al., 2016; Leyton et al., 2017; Lukic et al., 2019; Teichmann et al., 2013).

Despite their distinct profiles, these syndromes present several overlaps that complicate their clinical differentiation (Foxe et al., 2021). Evidence in support of the impaired syntactic processing includes results showing that individuals with nfvPPA perform particularly poorly on sentence comprehension and production tasks: they often lose the ability to build and/or process hierarchical syntax, such as the ability to form Wh questions or non-canonical (object-before-subject) argument order sentences, correctly inflect verbs for tense, use subject pronouns, and to produce subordination conjunctions, while at the same time retain the ability to use canonical structures and produce coordination conjunctions (e.g., Rogalski et al., 2011; Thompson, Cho, et al., 2012; Thompson et al., 2013; Wilson et al., 2010a, 2012, 2014; also see Thompson & Mack, 2014 for a review on PPA). However, individuals with lvPPA also show impaired syntactic comprehension (i.e. difficulties with complex syntactic structures, which contain embedded clauses and/or non-canonical argument order) and specifically with long, syntactically complex sentences (e.g., Wilson et al., 2012).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Similarly, impaired inflectional morphology (i.e. difficulties with verb and/or noun inflections) was observed in both variants, however, difficulties with inflecting pseudowords and low-frequency irregular words were specifically observed in lvPPA, most likely due to their core phonological or lexical impairments (Wilson et al., 2014). In addition, more difficulty with the possessive marking (e.g., “Mary’s shoe”) over the plural (e.g., “two shoes”) was reported in both PPA variants using more constrained morphosyntax generation tasks (Stockbridge et al., 2021a, 2021b). Furthermore, a recent online EEG study by Barbieri et al. (2021) examined morphosyntactic and verb argument violations across PPA variants, and reported a lack of a P600 to morphosyntactic violations (e.g., *The actors was singing in the theater.) in individuals with nfvPPA, but impaired verb argument structure production in both PPA variants (e.g., *Ryan was devouring on the couch) (also see Grossman et al., 2005; Kielar et al., 2018; Manouilidou et al., 2021; Peelle et al., 2007). Overall, these findings suggest that although morphosyntactic impairments may appear in both variants, differential impairment patterns may help characterize the linguistic profile of each PPA variant.

While the above evidence stems mainly from receptive or tightly controlled elicited sentence production tasks, a promising alternative lies in studying morphosyntax in connected speech (see Boschi et al., 2017 for a review on connected speech in neurodegenerative diseases). Quantitative analyses of connected speech have documented reduced production of syntactically well-formed sentences, and the omission and erroneous use of verbal inflectional morphology in individuals with nfvPPA, in that function words, especially pronouns and particles, are often reduced or eliminated (Faroqi-Shah et al., 2020; Themistocleous et al., 2021; Thompson et al., 1997; Thompson, Cho, et al., 2012; Thompson et al., 2013; Wilson et al., 2010b). Similarly, a recent study by Lavoie et al. (2021) also examined morphosyntactic features of connected speech in individuals with lvPPA, in comparison with healthy controls, using three different speech elicitation tasks (i.e., picture description, story narration, and semi-structured interviews). Compared to healthy controls, individuals with lvPPA produced a reduced proportion of open-class words, a higher proportion of verbs and pronouns, and a lower proportion of well-formed sentences (in line with Ash and Grossman, 2015; Thompson et al., 2013; Wilson et al., 2010a). However, a recent study reported that individuals with lvPPA produced fewer tense-inflected verbs compared to individuals with amnestic Alzheimer’s disease (AD) (Cho et al., 2022). These and other findings highlighted morphosyntactic deficits during connected speech across the PPA variants, with some deficits distinctly related to each variant.

However, capturing morphosyntactic deficits that differentiate nfvPPA from lvPPA still remains an enduring challenge in clinical practice. The traditional morphosyntactic analyses in PPA prove suboptimal or insufficiently exploited, due to coexisting methodological issues. First, elicited sentence tasks are highly constrained (non-connected) production tasks which are suboptimal because they fail to capture the actual distribution of morphosyntactic patterns, and they prioritize detection of errors over detection of differential patterns (which may reveal syndrome-specific patterns irrespective of whether errors are made). For instance, studies on other neurodegenerative disorders presenting frontostriatal disruptions and grammatical deficits, such as Parkinson’s disease, show that detection of differential morphosyntactic patterns in unfolding discourse (rather than errors proper) offers robust

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

discrimination between patients and controls, irrespective of language typology (Eyigoz et al., 2020).

Second, the approach to measurement in connected speech data is labor-intensive and requires sometimes subjective decisions by speech-language pathologists or others with linguistic training, which are subject to inter-rater variability, calling for more objective approaches. While good agreements between manual coding and automated quantification of syntactic deficits in aphasia have been reported (Fromm, MacWhinney, & Thompson, 2020), automation provides several benefits, such as being fast, replicable, and does not require extensive linguistic training. Furthermore, automatic methods were employed successfully in the automatic identification/diagnosis of individuals with amnestic AD (Bucks et al., 2000; Fraser et al., 2016; Rentoumi et al., 2014; Themistocleous et al., 2020) or individuals with PPA variants (Matias-Guiu et al., 2022; Themistocleous et al., 2021).

Third, the neuroanatomical substrates of morphosyntactic abilities in PPA are not completely clear, probably due to mixed receptive-expressive morphosyntax. For example, a recent study by Mesulam and colleagues (2021) identified a morphosyntactic cluster to be located predominantly within the left posterior Inferior Frontal Gyrus (IFG) using a composite of scores (the %Grammaticality of free narrative score thought to be more reflective of morphology compared to standardized NAT-NAVS scores related to sentence construction). Given that the neuroanatomical signature of morphosyntactic deficits in spontaneous speech has rarely been examined in individuals with PPA (Matias-Guiu et al., 2022; Mesulam et al., 2021; Wilson et al., 2014), automated analysis of different aspects of syntax (verb inflections, subject pronouns, and subordination conjunctions) in spontaneous speech individuals with nfvPPA in comparison with lvPPA is needed. In line with several current neurocognitive models (e.g., Friederici, 2011; Matchin & Hickok, 2020; Zaccarella & Friederici, 2015), where BA 44 or Pars Opercularis in the posterior IFG, which strongly connects to the temporal cortex via the dorsal pathway, appears to be particularly involved in a syntactic hierarchy (called *hierarchy-Merge*), we expect the reduction of syntactic embedding indicated through subordination conjunctions to be associated with the left BA44 integrity across individuals with PPA. In the transformational grammatical framework (Chomsky, 1995), syntax is comprised of lexical items and structure building operation, Merge, that creates structural dependencies and combines lexical items. While there are well-defined operations for structure building (aka Merge) as discussed under *the "growing tree"* rubric (MG; Friedmann et al., 2021) with the complexity of these structure-building processes being in BA44 and in a small pSTS cluster (Grodzinsky, Pieperhoff, & Thompson, 2021), there are recent debates regarding the extent to which syntax is hierarchical or could be considered *the cascade of expectations* that higher functional domains require (e-MGs; Chesi, 2021). The crucial difference between the two approaches, among other things, is that structure building operates from bottom to top in MG and top-down in e-MG and thus the former cannot make "on-line" (word by word) predictions; however, we adopt the former MG measure of syntax in the current manuscript to reflect more common approaches in the literature.

Here, we introduce a novel framework to tackle these three challenges. First, we used a semi-spontaneous production task which elicits diverse morphosyntactic patterns and can

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

thus reveal which specific morphosyntactic features differentiate between groups. Second, we leveraged Natural Language Processing (NLP) tools, automatically capturing patterns of use of these fine-grained morphosyntactic features without the need for subjective decisions and without restricting outcomes to error counts. In particular, we targeted three strategic sets of morphosyntactic features, including (1) sentence-related features such as intercausal connectors, like subordination conjunctions, which connect one dependent clause to an independent clause, creating a complex sentence, (2) verbal-related features such as verb inflections, like tense, which inflect the verb for past, present, or future, and (3) nominal-related features such as different types of pronouns, like subject pronouns (see Boschi et al., 2017 for a review).

With this approach, we conducted the first automated study of morphosyntactic patterns in individuals with nfvPPA versus lvPPA. Specifically, we examined whether these features can (i) robustly classify individuals from each group and (ii) add to standard (language and neuropsychological) assessments in discriminating nfvPPA from lvPPA. Moreover, we explored whether the most discriminative morphosyntactic features thus detected correlated with syndrome-specific atrophy patterns in nfvPPA. This way, we aim to inform the ongoing quest for reliable, scalable markers to differentiate individuals with nfvPPA from those with lvPPA.

2. Materials and Methods

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.1. Participants

The study comprised 71 participants, namely: 35 individuals with nfvPPA, 26 with lvPPA, and 10 healthy controls. All participants were recruited at the Memory and Aging Center (University of California, San Francisco). Patients were diagnosed based on a detailed medical history, comprehensive neurological, and standardized neuropsychological and language evaluations, and they were classified as nfvPPA or lvPPA according to the current international criteria (Gorno-Tempini et al., 2011). Inclusion criteria for PPA participants comprised a Mini-Mental State Examination (MMSE; Folstein et al., 1975) score ≥ 10, a Clinical Dementia Rating (CDR; Morris, 1993) score ≤ 2, an output of ≥15 words on the picture description task from the Western Aphasia Battery (WAB; Kertesz, 1982) and a structural MRI scan within one year of testing. All participants had normal or corrected-to-normal vision.

Healthy controls (HCs) were recruited through the Hillblom Healthy Aging Network (with normal neurological, cognitive, and speech-language profiles). A group of 10 HCs was selected as they underwent the WAB picture description task and a structural MRI within a year of testing. All the HC participants had a CDR score of zero, an MMSE score ≥25, and were determined to be clinically normal based on a consensus conference that reviewed the neuropsychological tests completed by the PPA. This group was matched for age, sex, and education to both PPA groups. The study was approved by the UCSF Committee on

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Human Research, and all participants provided written informed consent in accordance with the Declaration of Helsinki. The UCSF Human Research and Protection Program Institutional Review Board approved the three studies participants consented to (10-03946, 10-00619, 12-10512). Demographics, neuropsychological scores, speech-language scores, and expected significant group differences are provided in Table 1. No part of the study procedures and analysis plans were pre-registered prior to the research being conducted. However, participant selection criteria were established prior to data analysis.

2.2. Materials

2.2.1. Standardized neuropsychological and language assessment—The standardized assessment comprised widely used neuropsychological and speech-language tests, collected as part of the protocol's regular clinical visits (see Table 1 for a list of tests). Based on the current consensus criteria (Gorno-Tempini et al., 2011), this diagnostic assessment is proven to capture cognitive and speech-language disruptions across PPA variants. Comprehensive assessment of speech and language is part of the standard diagnostic evaluation of PPA and was recently summarized and described (see Europa et al., 2020). Because each PPA variant is associated with a particular pattern of deficits, the speech-language assessment covers several domains, such as semantics, morphosyntax, phonology, motor speech, and repetition. Comprehensive assessment of cognitive abilities is also standard practice in an evaluation of PPA assess to capture distinct neuropsychological patterns in the domains of memory, executive function, and visuospatial, that may emerge during the disease (Kramer et al., 2003; Staffaroni et al., 2019); also see Europe et al. 2020 for examples of tasks for assessing each domain within the neuropsychological and speech-language assessments. Legal copyright restrictions prevent public archiving of clinical assessment tests which can be obtained from the copyright holders in the cited references (see Table 1).

2.2.2. Connected speech assessment—Connected speech samples were collected with the picture description task from the Western Aphasia Battery (WAB). All participants viewed a black-and-white drawing of a picnic scene and described it following this instruction: "Tell me what you see and try to talk in sentences". Speech samples were recorded on a digital video camcorder (Wilson et al. 2010b). Speech samples were transcribed by experienced speech-language pathologists (SLP), who were blinded to the diagnosis. As in previous work (Berndt et al., 2000; Saffran et al., 1989; Wilson et al., 2010b), false starts, repetitions, revisions, fillers, and unintelligible segments were annotated and removed from analysis. This resulted in the average exclusion of 14.0 ± 14.9 words for nfvPPA patients and 21.5 ± 15.6 words for lvPPA patients ($t = -1.90$, $p = .063$). The samples included, on average, the total number of words produced for nfvPPA: 113.71 ± 102.35 and lvPPA: 129.69 ± 61.94 ($t = -0.75$, $p = .453$). This exclusion ensures that the following morphosyntactic analyses are not confounded by these unrelated word fragments or unintelligible speech.

2.2.3. Inter-rater reliability—Speech samples were transcribed by a licensed SLP supported by a trained research assistant. For reliability testing, data from 40% of participants (9 nfvPPA patients, 13 lvPPA patients, and 7 healthy controls) was transcribed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by an independent SLP. All transcribers were blind to the recordings' group. The correlation between the number of total words recorded per participant across the two raters was $r = 0.99$. Discrepant words across the two SLPs amounted to less than 6% of transcriptions (nfvPPA: 5.8%, lvPPA: 6.8%, CTRL: 4.6%). Importantly, most discrepant cases concerned omitting or substituting articles (e.g., 'a' with 'the'), or prepositions (e.g., 'in' vs 'on'), which were not included in the classification analysis below (see Table 2).

2.2.4. Connected speech analysis and feature extraction using NLP—The text transcripts were automatically tokenized for parts of speech using the spaCy 2 (Honnibal & Montani, 2017). The speech features were extracted using a pre-trained part-of-speech tagger and a dependency tagger trained on written English web text using a token-to-vector (tok2vec) model (https://spacy.io/models/en#en_core_web_lg). The part-of-speech tagger achieved a reported accuracy of 0.97 and the dependency tagger 0.92 (Honnibal & Montani, 2017). The part-of-speech tagger and the dependency parser together extracted 15 pairs of token-level part-of-speech and dependency tags. However, 6 of those pairs had 0 or close to 0 values, indicating that they were rarely produced by the groups. Therefore, we defined our morphosyntactic features as the 9 pairs that were consistently used across groups and restricted our analyses to them (see Table 2). We analyzed their relative feature importance. In addition, words in the participants' texts were matched with a corpus of stop words, which are common words that are purely functional and typically contain articles, prepositions, pronouns, and auxiliary verbs (e.g., a, and, they, and will). In order to control for the variation in the lengths of participants' speech, the extracted count of morphosyntactic features was then converted to fractions of the total number of non-stop words. In addition, the standardized assessment (combined neuropsychological and language tests; see Table 1) is used to examine the relative robustness of our automated speech features that were aimed to capture morphosyntactic aspects of the speech samples. See Supplementary Material Table 1 and Figure 1 on relationships between syndrome-specific standardized measures and morphosyntactic features fractions.

2.3. Statistical Analyses

2.3.1. Machine Learning Classifications—The morphosyntactic features were used in different classification tasks of machine learning. We used Gradient Boosting Machines (GBM), a method that reduces overfitting, combines distinct decision trees' predictions to generate final outcomes, handles many features simultaneously without manual feature selection, and is proven sensitive to sample size and a number of features (Myung, 2000; also see Eygioz et al., 2020 using morphological features). More importantly, to ensure that our results were not method-dependent, we replicated the analyses with other algorithms (Generalized Linear Models (GLM), Distributed Random Forest (DRF), and Extreme Gradient Boosting (XGBoost)). Our models are shown to be generalizable to unseen data by the metrics we report on the test set.

The area under the ROC (receiver operating characteristics) curve (AUC) was calculated to evaluate the predictive ability of learning algorithms using a 70-30% train-test split (Crimin et al., 2021; Sarawgi et al., 2020) with hyperparameters selected by automated machine learning package (H2O AutoML; LeDell & Poirier, 2020) on the training set. We reported

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the AUC over accuracy for established theoretical and empirical reasons (Huang & Ling, 2005). Accuracy (unlike AUC) may be substantially modified by even small changes in a sample's predicted probability if close to the classification boundary (Guo et al., 2021). The ROC curves display the sensitivity and specificity of the classification model using the morphosyntax features across different thresholds, with an optimal model having an AUC closer to 1 (Carter et al., 2016; Chicco, 2017). They represent the predictive power of the model without the need for a specific threshold, thus giving the flexibility of choosing one that best fits the use cases. See Supplementary methods for more details on four classifying algorithms (GBM, GLM, DRF, XGBoost) and all metrics (e.g., AUC, accuracy, precision, recall).

Relative feature importance within the models was then calculated using the H2O AutoML package to establish the discriminatory weight of each morphosyntactic variable between HCs and the two PPA variants. Feature importance was calculated by evaluating whether a feature was employed for node splitting and the extent to which it decreased squared error across trees, taking into account weights to ensure accuracy in the variance of response values within nodes (Breiman, 2001). Thus, it is represented by the coefficient magnitudes of the features (i.e., percentages as compared to the top feature). The study's sample size was established following a feature-to-sample ratio criterion of N-1 (Hua et al., 2004), which requires that the number of participants in the training set outnumbers the features fed to the classifier by at least one. Considering that our classifiers employed a total of nine features, this required a minimum of 10 participants in each classifier's training set. Considering a 70/30 split, this minimum was satisfied even by the classifier with the fewest subjects (lvPPA patients vs. HCs), which involved 25 participants for training.

2.3.2. MRI data acquisition and image processing—A 3T Trio (Siemens) scanner was used to obtain structural 3D T1-weighted images at UCSF in each participant. The T1-weighted images were acquired using an MP-RAGE sequence with the following parameters: repetition time (TR) = 2300 ms, echo time (TE) = 2.98 ms, inversion time = 900 ms, flip angle 9°, matrix size = 256×240, voxel size 1 mm^3 isotropic. Pre-processing of neuroimaging data was performed using the Computational Anatomy Toolbox (CAT12; http://dbm.neuro.uni-jena.de/cat) in Statistical Parametric Mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm/software/spm12) under Matlab 2019b. A standard voxel-based morphometry (VBM) approach was performed and included the following steps: tissue classification (GM, WM, CSF) using the unified segmentation (Ashburner & Friston, 2005), spatial normalization to a reference space using the Diffeomorphic Anatomical Registration using Exponentiated Lie Algebra (DARTEL) algorithm (Goto et al., 2013), modulation by the Jacobian determinant, and spatial smoothing using a 8mm full-width at half-maximum (FWHM) Gaussian Kernel. Prior to the imaging analyses, image quality control was performed on raw MRI acquisitions and pre-processed data. Quality parameters were generated for each subject under the module "Data Quality -> VBM data homogeneity" in the CAT12, such as the weighted overall image quality, which combines measurements of noise and spatial resolution of the image before preprocessing. All participants had a high weighted overall image quality above 90%, and, therefore, all participants were included in the statistical analyses reported below.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2.3.3. MRI data statistical analyses—To assess the relationship between morphosyntactic performances (i.e., the top morphosyntax feature 'subordinating conjunction') and gray matter volume (i.e., atrophy) across participants, we conducted two analyses: (1) a targeted restrictive ROI analysis, focusing on multiple left hemisphere brain regions known to be involved in language processing previously associated with language (Hu et al., 2023), and (2) a whole-brain voxel-wise analysis to explore global brain behavior relationships. The ROI analysis included the anterior IFG, posterior IFG, middle frontal gyrus, the anterior and posterior temporal, and the angular gyrus, consistent with delineations from Fedorenko et al. (available at https://evlab.mit.edu/funcloc/; Hu et al., 2023). We hypothesize, in line with prior neuroimaging research (Bulut, 2022, Leminen et al., 2019, and Zaccarella & Friederici, 2015), that the left posterior IFG (pars opercularis) is particularly implicated in morphosyntactic processing and the construction of syntactic hierarchies.

Given our study's relatively small sample size, we adopted an ROI approach alongside whole brain analysis to enhance the detection of localized effects, thereby increasing our statistical power. While ROI analysis inherently decrease the multiple comparison burden, they can potentially limit explorations to predetermined areas. To counter this, we conducted comprehensive whole-brain voxel-wise analysis to ensure we captured all regions potentially related to morphosyntax. This dual approach also aimed to reduce the risk of inflated correlations often criticized in neuroimaging studies (Marek et al., 2022; Vul et al., 2009).

Input of the analysis were the smoothed gray matter images, with the 'subordinating conjunction' fractions as the main predictor. Participant demographics (age, sex, handedness) and total intracranial volume were included as covariates to account for non-interest variance and brain size differences, respectively. The threshold of $p < 0.001$ and threshold-free cluster enhancement (TFCE) correction (Smith and Nichols, 2009; Salimi-Khorshidi et a., 2011) were applied for multiple comparisons, and the family-wise error rate (FWE) was $p < 0.05$. The TFCE was used to correct multiple comparisons at the cluster level within the ROI analysis in which we used the left-brain regions described above as explicit mask. For the whole-brain analysis, we allowed an uncorrected threshold and resulting statistical maps are provided in the supplementary material as suggested in Poldrack et al. 2008. To mitigate multicollinearity in our brain-behavior correlation analysis and to focus on the most discriminative features, we applied methods utilized in recent nvfPPA studies (García et al., 2022).

3. Results

3.1. Classification results

A binary GBM classifier between nfvPPA and lvPPA using all nine morphosyntactic features was very robust in discriminating between PPA variants with an AUC of 0.95. Moreover, these nine morphosyntactic features revealed a high AUC of 0.82 for individuals with nfvPPA versus healthy controls, but a near chance AUC of 0.68 for individuals with lvPPA versus healthy controls. ROC curve classification between nfvPPA and lvPPA individuals was significantly higher for the morphosyntactic features than for standardized assessment (AUC of 0.95 versus AUC of 0.83; Chi-squared = 6.45, p = 0.011). Fig. 1 shows

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

receiver operating characteristic (ROC) curves with the mean area under the curve (AUC) and confusion matrices from the GBM classification models.

3.2. Feature importance

We extracted nine morphosyntactic features from the three categories: (1) sentence-related features such as clausal connectors (subordinating conjunctions and coordinators), (2) verbal-related features such as tense (present, past, be + progressive "-ing"), and (3) nominal-related features such as different types of pronouns (subjective/possessive). For the GBM algorithm, we calculated a feature's "feature importance score" as the reduction in squared error between the node splitting on that feature and its children nodes (see Fig. 2). To describe differences in feature types between groups, we created weighted scores for each feature type. Firstly, the fraction score for each variable (e.g., pronouns) was multiplied by the "feature importance score", of the variable derived from the GBM algorithm. Secondly, the resulting values, for the variables, were summed across sentence-related, verb-related, or nominal-related features to derive each feature type weighted score. The weighted mean proportion (standard deviation) for each feature type for each group is displayed in Fig. 3. Feature importance analysis of the GBM model revealed the top morphosyntactic feature to be *subordinating conjunction* (e.g., like, while, if)] (see Figure 2 for the nine features and their relative feature importance).

3.3. Association between the top morphosyntactic feature and brain atrophy

The ROI vowel-wise analysis targeted to the predefined language areas revealed the top morphosyntax feature – subordinating conjunction to be significantly correlated with gray matter loss in the left posterior IFG pars opercularis (x=−45, y=16, z=16), peak level p_{FWE} = .032 FWE corrected and p<0.001 TFCE corrected (Fig. 4). These results align with our hypothesis and previous research suggesting the left posterior IFG's involvement in morphosyntactic processing. The whole-brain voxel-wise analysis did not show any significant association between the subordinating conjunction and gray matter volume surviving at the cluster level, even when applying a liberal uncorrected threshold (p<0.05). The focused nature of the ROI analysis has therefore provided a sensitive means to observe the hypothesized neural correlates of syntactic function despite the limitations posed by our sample size. In Fig. 2 of Supplementary Material, we provide the scatterplots to illustrate the relationship between the average value within each ROI and the subordination conjunction. In Fig. 3 and Table 3 of supplementary Material, we also provide uncorrected threshold statistical maps resulting from whole-brain exploratory analyses.

4. Discussion

The current study aimed to establish differential signatures of nfvPPA/agrammatic (relative to lvPPA) using automated morphosyntactic measures of natural speech. Assorted morphosyntactic features classify between individuals with nfvPPA and lvPPA with high accuracy, surpassing the discriminatory power of standardized assessments. The most important feature for classification was the use of subordinating conjunctions, a feature that was significantly associated with atrophy of the left posterior IFG pars opercularis across participants. These findings indicate that multiple morphosyntactic categories are impaired

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in agrammatic production and that subordination production may be a relevant target in differential diagnoses.

First, we showed that NLP automatically extracted morphosyntactic features from connected speech (elicited through the WAB “Picnic Scene” task) were robust in discriminating nfvPPA from lvPPA and healthy controls. Morphosyntactic selectivity is manifested in the distribution of three types of features, sentence-related (clausal connectors), verbal-related (tenses), and nominal-related features (pronouns) in distinguishing individuals with two different variants. This result is consistent with those studies primarily showing morphosyntax impairments in nfvPPA/agrammatic individuals using constrained production tasks (Thompson et al., 2013; Wilson et al., 2014), and/or connected speech (Thompson et al., 1997; Thompson et al., 2013; Wilson et al., 2010b). The findings contribute to a paucity of studies directly contrasting the nfvPPA and lvPPA variants using connected speech features (Matias-Guiu et al., 2022; Mesulam et al., 2021; Thompson et al., 2013) and provide direct evidence of the ability of these fine-grained morphosyntactic features to discriminate among the two PPA variants.

Another important finding regards the morphosyntactic features and the standardized screening batteries designed to assess the clinical features of frontotemporal dementia (FTD). For instance, in the current study, individuals with nfvPPA and lvPPA differ in their performance on verbal and visual memory tests and confrontation naming ability on the standardized (bedside) battery. The observed pattern of neuropsychological findings is consistent with previous reports; for example, Kramer et al. (2003) showed that verbal and visual memory were most impaired in the AD group relative to the FTD group and could successfully differentiate the two groups (Hodges et al., 1999). Similarly, recent studies showed that nfvPPA and lvPPA individuals differ on the tests of Uniform Data Set (UDS) battery (Staffaroni et al., 2021; Weintraub et al., 2018); for example, digit span forward best discriminated lvPPA from other PPAs, whereas phonemic:semantic fluency ratio was excellent in classifying nfvPPA compared to lvPPA and svPPA. However, the utility of these screening tests for discriminating between the nfvPPA and lvPPA remains unknown. Although, we demonstrated overall good classification for the standardized assessment (combined neuropsychological and language tests), the classification between nfvPPA and lvPPA was significantly higher for the morphosyntactic features than for the standardized assessment, suggesting that these automated features are clinically useful in differentiating the two PPA variants and can improve differential diagnosis when coupled with standardized assessment. Notably, because of a discrepancy between sensitivity and specificity, this tool still fails to have optimal discrimination but could be particularly important to ascertain nfvPPA diagnoses. Future studies should select theoretically-guided lexical features based on the epicenter of differential atrophy in lvPPA.

Notably, among all the features, we identified clausal connectors (particularly subordination) as the most highly ranked feature of connected speech for nfvPPA diagnosis. The causal connectors are a measure of syntactic complexity (Huddleston, & Pullum, 2006) and generally refer to two or more clauses related syntactically. However, while *coordination* (and, or, but) uses conjunctions to connect structures that link two units of the same grammatical status, such as two clauses or two noun phrases (a compound sentence),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

subordination (like, while, if) uses conjunctions to connect one dependent clause to an independent clause or to adverb/adjective/noun clauses, creating a complex sentence. Thus, the subordination finding as a top-ranked feature is in line with previous studies showing that individuals with nfvPPA produced fewer complex sentences and shorter utterances than individuals with lvPPA (e.g., Thompson et al., 2013; Wilson et al., 2010b). However, this finding is in contrast with Matias-Guiu et al. (2022), who found that such syntactic parameters are not specific to nfvPPA; this could be likely due to the differences in the way this feature is computed (the subordination index versus only subordination fractions), and/or differences between Spanish and English participants in the two studies. Follow-up studies should explore Minimalist Grammars (MGs) and expectation-based (e-MGs) approaches to syntax (Chesi, 2021), and compute these syntactic constructions in alternate ways which can be used to predict case checking, thematic role assignment, and memory tenure during syntax processing in PPA.

Another interesting finding is our brain-behavior correlation, which shows a correlation between a top feature *subordination* production and the left posterior IFG Pars Opercularis atrophy. This suggests that atrophy of the left posterior IFG impacts morphosyntax, especially aspects of hierarchical syntactic processing across PPA participants. The neural correlates of morphosyntax deficit in connected speech in PPA have not been systematically studied, with only a paucity of studies correlating connected speech features with MRI brain volumes (Mesulam et al., 2021 and Wilson et al., 2010b), CSF p-tau levels (Cho et al., 2022), and/or PET and DTI metabolism and fractional anisotropy (Matias-Guiu et al., 2022). Our brain-behavior correlation finding is consistent with these PPA studies. For instance, a recent study by Matias-Guiu et al. (2022) demonstrated that the subordination index was correlated with the metabolism of the left parietal-temporal lobe regions and frontal regions (middle and inferior frontal gyri). It is also in line with recent reviews on neuroimaging studies in healthy individuals, which implicated the left posterior IFG region in morphosyntactic processing (Bulut, 2022; Leminen et al., 2019), as well as a few intraoperative cortical stimulation mapping studies during sentence production (Chang, Kurteff, and Wilson, 2018) and comprehension (Riva et al., 2022), and stroke lesion-symptom mapping studies of sentence production (Gleichgerrcht et al., 2021). Finally, our brain-behavior evidence from PPA contributes to the current neurocognitive models, where Pars Opercularis in the posterior IFG (BA44) appears to be mainly involved in a syntactic hierarchy (Friederici, 2011; Matchin & Hickok, 2020; Zaccarella & Friederici, 2015, 2017).

It is also worth mentioning that this study has substantial clinical implications. Our findings can provide an additional tool for better discrimination between individuals with a clinical diagnosis of nfvPPA and lvPPA, where the clinical similarities between the two variants pose a considerable challenge for clinicians. Up to this point, no individual test can discriminate between nfvPPA and lvPPA (Harris et al., 2019). This study discovered subtle syndrome-specific morphosyntactic differences between variants and utilized them through a new automated testing tool to better discriminate between nfvPPA and lvPPA in conjunction with other clinical tests. Each of these tests may be useful in analyzing particular aspects of the variants, but differentiation between nfvPPA and lvPPA in clinical testing is still very challenging. Though we diagnose patients as discrete syndromes based on the original consensus criteria (Gorno-Tempini et al., 2011), overlap is common across neurobiological

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and pathophysiological levels (those presenting with mixed PPA; i.e., a proposed fourth subtype of PPA; see Mesulam et al., 2021). Given the lack of a one-to-one match between symptoms and pathology, there have been ongoing efforts towards modifications to the diagnostic criteria to improve diagnostic classification. Our study, first, may contribute to the discrimination of these syndromes based on these novel speech markers, and second, it could be further reassessed in additional cohorts with typical and atypical pathology to enable further validation.

Though, the semi-structured interview is suggested to be more sensitive in highlighting abnormalities of morphological and syntactical structure than the picture description, our study and others found that picture description and semi-structured interview tasks can be employed interchangeably for the assessment of lexico-semantic and syntactic domains in subjects with clinical diagnoses of PPA (also see Ash et al., 2013); yet, future studies should test these language markers across the two tasks. Moreover, recent research showed that Correct Information Units (CIU) are sensitive markers for PPA variants, in that nfvPPA scored lower than lvPPA or svPPA (e.g., Faroqi-Shah et al., 2020); however, these findings were limited by a small number of participants and variations in the literature on how to segment utterances into communication units. Although our study did not focus on CIUs and a number of errors, but rather on theory-driven morphosyntax features, further work could consider the additional CIU analysis with bigger sample sizes. These and other tools that enable automatic identification of nfvPPA and lvPPA may optimize individualized treatment, prognosis, and monitoring of the patients.

Though there is a long-standing interest in morphosyntactic production in PPA, research has not led to conclusive morphosyntactic patterns for the systematic classification of PPA variants (e.g., Faroqi-Shah et al., 2020; Themistocleous et al., 2021). Previous studies on PPA demonstrated morphological deficits in both nfvPPA and lvPPA, though with different degrees, using constrained production tasks (Thompson et al., 2013; Wilson et al., 2014) and connected speech (Thompson et al., 1997; Thompson et al., 2013; Wilson et al., 2010b). Moreover, while most PPA studies on connected speech compared nfvPPA and svPPA (Ash et al., 2006; Fraser et al., 2014; Meteyard & Patterson, 2009; Sajjadi et al., 2012), studies differentiating nfvPPA and lvPPA are limited and remain an enduring challenge (Matias-Guiu et al., 2022; Wilson et al., 2010b). Our findings show substantial sensitivity of morphosyntactic production in individuals with nfvPPA compared to lvPPA and healthy controls. Finally, the use of subordinating conjunctions, which was ranked as a top feature, was correlated with the integrity of crucial brain regions implicated in syntactic processing, the left inferior frontal region, particularly Pars Opercularis, that are typically compromised in nfvPPA compared to lvPPA. In the current study, we focused solely on syntax production during connected speech given that the majority of studies have focused on sentence comprehension, with fewer in the domain of sentence production. However, future studies that aim to test comprehension of naturalistic narratives that more closely resemble our natural dynamic environment will contribute to a functional subdivision of the PPA.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Limitations

We also identify potential limitations in our study. While the sample size was relatively small for machine learning analysis, the number of participants in the lvPPA and nfvPPA groups are relatively large for these diagnoses, and the application of NLP methods on connected speech compared to previous literature and the prevalence of these syndromes. Another limitation of our study is that our sample is based on specialized clinic-based recruitment, which increases selection bias and results in relatively highly educated and predominantly white participants. This sample composition may limit generalizability to the general population. Future research is required to assess whether morphosyntactic features are equally sensitive to diagnosis group classification in more educationally and racially/ethnically diverse samples as well as across those with typical and atypical pathology. Further research validating the metrics as sensitive to a given syndrome is the first necessary step before clinical implementation is possible. Recent tools, such as the TELL app (García et al., 2023), allow incorporating automated pipelines initially validated in scientific papers for ulterior use in clinical settings. For example, after its initial version (García et al., 2023), TELL has incorporated novel metrics capturing specific motor speech (García et al., 2021) and lexico-semantic (Ferrante et al., 2023) alterations. The present work paves the way for similar developments in the morphosyntactic domain. Connected speech studies often lack deep cognitive and biomarker phenotyping of participants in addition to the analysis of linguistic features. One of the strengths of our study is the analysis of brain regions in relation to connected speech features to validate observed behavioral differences against neuroimaging biomarker evidence of atrophy patterns.

5. Conclusion

Using connected speech, a highly ecologically valid approach to language assessment, and an automated natural language processing technique, we demonstrated the sensitivity of morphosyntactic production for the systematic classification of nfvPPA compared to lvPPA and healthy controls and provided evidence for its neural substrates in the left posterior IFG volume. Beyond that, we showed that the automated features are clinically useful in differentiating nfvPPA and lvPPA and have the potential to improve the differential diagnosis of PPA variants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

This work was supported by the National Institutes of Health (MLGT., NINDS R01 NS050915, NIDCD K24 DC015544; NIA U01 AG052943; B.L.M., NIA P50 AG023501, NIA P01 AG019724; SMW., NIDCD R01 DC013270). Adolfo García is supported by the National Institute On Aging of the National Institutes of Health (R01AG075775); GBHI, Alzheimer's Association, and Alzheimer's Society (Alzheimer's Association GBHI ALZ UK-22-865742); ANID (FONDECYT Regular 1210176, 1210195); Programa Interdisciplinario de Investigación Experimental en Comunicación y Cognición (PIIECC), Facultad de Humanidades, USACH; and the Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat), funded by the National Institutes of Aging of the National Institutes of Health and Fogarty International Center (FIC) under award number R01AG057234, an Alzheimer's Association grant (SG-20-725707-ReDLat), the Rainwater Foundation, and the GBHI. We are thankful to our patients and healthy volunteers for participating in the research.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Data availability

The conditions of our ethics approval do not permit public archiving of anonymized study data. Data generated by the UCSF MAC are available upon request. Data requests can be submitted through the UCSF MAC Resource Request form: http://memory.ucsf.edu/resources/data. Access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data. All requests will undergo UCSF regulated procedure, thus requiring submission of a Material Transfer Agreement (MTA) which can be found at https://icd.ucsf.edu/material-transfer-and-data-agreements. No commercial use would be approved.

References

Ash S & Grossman M (2015). Why study connected speech production. In: R.M. Willems (Ed.) Cognitive neuroscience of natural language use. Cambridge, USA: Cambridge University Press, pp. 29–58.

Ash S, Evans E, O'Shea J, Powers J, Boller A, Weinberg D, & Grossman M (2013). Differentiating primary progressive aphasias in a brief sample of connected speech. Neurology, 81(4), 329–336. 10.1212/WNL.0b013e31829c5d0e [PubMed: 23794681]

Ash S, Moore P, Antani S, McCawley G, Work M, & Grossman M (2006). Trying to tell a tale: Discourse impairments in progressive aphasia and frontotemporal dementia. Neurology, 66(9), 1405–1413. 10.1212/01.wnl.0000210435.72614.38 [PubMed: 16682675]

Barbieri E, Litcofsky KA, Walenski M, Chiappetta B, Mesulam MM, & Thompson CK (2021). Online sentence processing impairments in agrammatic and logopenic primary progressive aphasia: Evidence from ERP. Neuropsychologia, 151, 107728. 10.1016/j.neuropsychologia.2020.107728 [PubMed: 33326758]

Berndt RS, Wayland S, Rochon E, Saffran EM, & Schwartz MF (2000). Quantitative production analysis: A training manual for the analysis of aphasic sentence production. Hove, East Sussex UK: Psychology Press

Boschi V, Catricala E, Consonni M, Chesi C, Moro A, & Cappa SF (2017). Connected speech in neurodegenerative language disorders: a review. Frontiers in psychology, 8, 269. doi: 10.3389/fpsyg.2017.00269 [PubMed: 28321196]

Breiman L (2001). Random forests. Machine learning, 45, 5–32.

Bucks RS, Singh S, Cuerden JM, & Wilcock GK (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. Aphasiology, 14(1), 71–91. 10.1080/026870300401603

Bulut T (2022). Neural correlates of morphological processing: An activation likelihood estimation meta-analysis. Cortex, 151, 49–69. 10.1016/j.cortex.2022.02.010 [PubMed: 35397379]

Carter JV, Pan J, Rai SN, & Galandiuk S (2016). ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery, 159(6), 1638–1645. 10.1016/j.surg.2015.12.029 [PubMed: 26962006]

Chang EF, Kurteff G, & Wilson SM (2018). Selective interference with syntactic encoding during sentence production by direct electrocortical stimulation of the inferior frontal gyrus. Journal of cognitive neuroscience, 30(3), 411–420. 10.1162/jocn_a_01215 [PubMed: 29211650]

Chesi C (2021). Expectation-based minimalist grammars. arXiv preprint arXiv:2109.13871.

Chicco D (2017). Ten quick tips for machine learning in computational biology. BioData mining, 10(1), 35. DOI 10.1186/s13040-017-0155-3 [PubMed: 29234465]

Cho S, Cousins KAQ, Shellikeri S, Ash S, Irwin DJ, Liberman MY, & Nevler N (2022). Lexical and acoustic speech features relating to Alzheimer disease pathology. Neurology, 99(4), e313–e322. 10.1212/WNL.0000000000200581 [PubMed: 35487701]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Crimin K, Allen PJ, Abba I, Ahlberg C, Benz L, Lau H, & Florian H (2021). Identifying predictive factors of patient dropout in Alzheimer's disease clinical trials. Alzheimer's & Dementia, 17, e052361. 10.48550/arXiv.2009.00700

de la Fuente Garcia S, Ritchie CW, & Luz S. (2020). Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: a systematic review. Journal of Alzheimer's Disease, 78(4), 1547–1574. DOI: 10.3233/JAD-200888

Delis DC, Kramer JH, Kaplan E, & Ober BA (2000). California verbal learning test--. Assessment.

Dunn LM, & Dunn LM (1965). Peabody picture vocabulary test.

Europa E, Iaccarino L, Perry DC, Weis E, Welch AE, Rabinovici GD, & Henry ML (2020). Diagnostic assessment in primary progressive aphasia: An illustrative case example. American journal of speech-language pathology, 29(4), 1833–1849. [PubMed: 32910678]

Eyigoz E, Courson M, Sedeño L, Rogg K, Orozco-Arroyave JR, Nöth E, & García AM (2020). From discourse to pathology: automatic identification of Parkinson's disease patients via morphological measures across three languages. Cortex, 132, 191–205. 10.1016/j.cortex.2020.08.020 [PubMed: 32992069]

Faroqi-Shah Y, Treanor A, Ratner NB, Ficek B, Webster K, & Tsapkini K (2020). Using narratives in differential diagnosis of neurodegenerative syndromes. Journal of Communication Disorders, 85, 105994. 10.1016/j.jcomdis.2020.105994 [PubMed: 32388191]

Ferrante FJ, Migeot J, Birba A, Amoruso L, Pérez G, Hesse E, & García AM (2023). Multivariate word properties in fluency tasks reveal markers of Alzheimer's dementia. Alzheimer's & Dementia.

Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research, 12(3), 189–198. 10.1016/0022-3956(75)90026-6 [PubMed: 1202204]

Foxe D, Cheung SC, Cordato NJ, Burrell JR, Ahmed RM, Taylor-Rubin C, & Piguet O (2021). Verbal short-term memory disturbance in the primary progressive aphasias: Challenges and distinctions in a clinical setting. Brain Sciences, 11(8), 1060. 10.3390/brainsci11081060 [PubMed: 34439679]

Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, & Rochon E (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. cortex, 55, 43–60. 10.1016/j.cortex.2012.12.006 [PubMed: 23332818]

Fraser KC, Meltzer JA, & Rudzicz F (2016). Linguistic features identify Alzheimer's disease in narrative speech. Journal of Alzheimer's Disease, 49(2), 407–422. DOI: 10.3233/JAD-150520

Friederici AD (2011). The brain basis of language processing: from structure to function. Physiological reviews, 91(4), 1357–1392. 10.1152/physrev.00006.2011 [PubMed: 22013214]

Fromm D, MacWhinney B, & Thompson CK (2020). Automation of the northwestern Narrative Language analysis system. Journal of Speech, Language, and Hearing Research, 63(6), 1835–1844. 10.1044/2020_JSLHR-19-00267

García AM, Johann F, Echegoyen R, Calcaterra C, Riera P, Belloli L, & Carrillo F (2023). Toolkit to Examine Lifelike Language (TELL): An app to capture speech and language markers of neurodegeneration. Behavior Research Methods, 1–15.

García AM, Arias-Vergara T, C Vasquez-Correa J, Nöth E, Schuster M, Welch AE, & Orozco-Arroyave JR (2021). Cognitive determinants of dysarthria in Parkinson's disease: an automated machine learning approach. Movement Disorders, 36(12), 2862–2873. [PubMed: 34390508]

García AM, Welch AE, Mandelli ML, Henry ML, Lukic S, Prioris MJT, & Gorno-Tempini ML (2022). Automated detection of speech timing alterations in autopsy-confirmed nonfluent/agrammatic variant primary progressive aphasia. Neurology, 99(5), e500–e511. 10.1212/WNL.0000000000200750 [PubMed: 35914945]

Gleichgerrcht E, Roth R, Fridriksson J, den Ouden D, Delgaizo J, Stark B, & Bonilha L (2021). Neural bases of elements of syntax during speech production in patients with aphasia. Brain and Language, 222, 105025. [PubMed: 34555689]

Gorno-Tempini ML, Brambati SM, Ginex V, Ogar J, Dronkers NF, Marcone A, et al. (2008). The logopenic/phonological variant of primary progressive aphasia. Neurology 71, 1227–1234. doi: 10.1212/01.wnl.0000320506.79811.da [PubMed: 18633132]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa S, et al. (2011). Classification of primary progressive aphasia and its variants. Neurology 76, 1006–1014. 10.1212/WNL.0b013e31821103e6 [PubMed: 21325651]

Goto M, Abe O, Aoki S, Hayashi N, Miyati T, Takao H, & Ohtomo K (2013). Diffeomorphic anatomical registration through exponentiated lie algebra provides reduced effect of scanner for cortex volumetry with atlas-based method in healthy subjects. Neuroradiology, 55(7), 869–875. DOI 10.1007/s00234-013-1193-2 [PubMed: 23619702]

Goodglass H, Kaplan E, & Weintraub S (2001). BDAE: The Boston Diagnostic Aphasia Examination. Philadelphia, PA: Lippincott Williams & Wilkins.

Grossman M, Rhee J, & Moore P (2005). Sentence processing in frontotemporal dementia. Cortex, 41(6), 764–777. 10.1016/S0010-9452(08)70295-8 [PubMed: 16353366]

Guo Y, Li C, Roan C, Pakhomov S, & Cohen T (2021). Crossing the "Cookie Theft" corpus chasm: applying what BERT learns from outside data to the ADReSS challenge dementia detection task. Frontiers in Computer Science, 3, 642517. 10.3389/fcomp.2021.642517

Harris JM, Saxon JA, Jones M, Snowden JS, & Thompson JC (2019). Neuropsychological differentiation of progressive aphasic disorders. Journal of Neuropsychology, 13(2), 214–239. 10.1111/jnp.12149 [PubMed: 29424041]

Henry ML, Wilson SM, Babiak MC, Mandelli ML, Beeson PM, Miller ZA, et al. (2016). Phonological processing in primary progressive aphasia. J. Cogn. Neurosci 28, 210–222. doi: 10.1162/jocn_a_00901 [PubMed: 26544920]

Hodges JR, Patterson K, Ward R, Garrard P, Bak T, Perry R, & Gregory C (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: a comparative neuropsychological study. Neuropsychology, 13(1), 31. 10.1037/0894-4105.13.1.31 [PubMed: 10067773]

Honnibal M, & Montani I (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear, 7(1), 411–420.

Hu J, Small H, Kean H, Takahashi A, Zekelman L, Kleinman D, & Fedorenko E (2023). Precision fMRI reveals that the language-selective network supports both phrase-structure building and lexical access during language production. Cerebral Cortex, 33(8), 4384–4404. [PubMed: 36130104]

Hua J, Xiong Z, Lowey J, Suh E, & Dougherty ER (2004). Optimal number of features as a function of sample size for various classification rules. Bioinformatics, 21(8), 1509–1515. 10.1093/bioinformatics/bti17 [PubMed: 15572470]

Huang J, & Ling CX (2005). Using AUC and accuracy in evaluating learning algorithms. IEEE Transactions on knowledge and Data Engineering, 17(3), 299–310. DOI: 10.1109/TKDE.2005.50

Huddleston R, & Pullum GK (2006). Coordination and subordination. The handbook of English linguistics, 198–219.

Kielar A, Deschamps T, Jokel R, & Meltzer JA (2018). Abnormal language-related oscillatory responses in primary progressive aphasia. NeuroImage: Clinical, 18, 560–574. 10.1016/j.nicl.2018.02.028 [PubMed: 29845004]

Kertesz A (2007). Western Aphasia Battery--Revised.

Kramer JH, Jurik J, Sharon JS, Rankin KP, Rosen HJ, Johnson JK, & Miller BL (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. Cognitive and Behavioral Neurology, 16(4), 211–218. [PubMed: 14665820]

Lavoie M, Black SE, Tang-Wai DF, Graham NL, Stewart S, Leonard C, & Rochon E (2021). Description of connected speech across different elicitation tasks in the logopenic variant of primary progressive aphasia. International Journal of Language & Communication Disorders, 56(5), 1074–1085. 10.1111/1460-6984.12660 [PubMed: 34383346]

LeDell E, & Poirier S (2020, July). H2o automl: Scalable automatic machine learning. In Proceedings of the AutoML Workshop at ICML (Vol. 2020).

Leminen A, Smolka E, Dunabeitia JA, & Pliatsikas C (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). Cortex, 116, 4–44. 10.1016/j.cortex.2018.08.016 [PubMed: 30268324]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Leyton CE, Hodges JR, Piguet O, and Ballard KJ (2017). Common and divergent neural correlates of anomia in amnestic and logopenic presentations of Alzheimer's disease. Cortex 86, 45–54. doi: 10.1016/j.cortex.2016.10.019 [PubMed: 27875715]

Lukic S, Mandelli ML, Welch A, Jordan K, Shwe W, Neuhaus J, et al. (2019). Neurocognitive basis of repetition deficits in primary progressive aphasia. Brain Lang. 194, 35–45. doi: 10.1016/j.bandl.2019.04.003 [PubMed: 31055171]

Manouilidou C, Nerantzini M, Chiappetta BM, Mesulam MM, & Thompson CK (2021). What language disorders reveal about the mechanisms of morphological processing. Frontiers in psychology, 12. doi: 10.3389/fpsyg.2021.701802

Matchin W, & Hickok G (2020). The cortical organization of syntax. Cerebral Cortex, 30(3), 1481–1498. [PubMed: 31670779]

Matias-Guiu Jordi A., et al. "Identification of the main components of spontaneous speech in primary progressive aphasia and their neural underpinnings using multimodal MRI and FDG-PET imaging." Cortex 146 (2022): 141–160. 10.1016/j.cortex.2021.10.010 [PubMed: 34864342]

Mesulam M (1987). Primary progressive aphasia-differentiation from Alzheimer 's disease. Ann. Neurol 22, 533–534. [PubMed: 3324947]

Mesulam MM, Coventry CA, Rader BM, Kuang A, Sridhar J, Martersteck A, & Rogalski EJ (2021). Modularity and granularity across the language network-a primary progressive aphasia perspective. Cortex, 141, 482–496. 10.1016/j.cortex.2021.05.002 [PubMed: 34153680]

Mesulam MM, Coventry C, Bigio EH, Geula C, Thompson C, Bonakdarpour B, & Weintraub S (2021). Nosology of primary progressive aphasia and the neuropathology of language. Frontotemporal Dementias: Emerging Milestones of the 21st Century, 33–49.

Meteyard L, & Patterson K (2009). The relation between content and structure in language production: An analysis of speech errors in semantic dementia. Brain and language, 110(3), 121–134. 10.1016/j.bandl.2009.03.007 [PubMed: 19477502]

Morris JC (1997). Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. International psychogeriatrics, 9(S1), 173–176. 10.1017/S1041610297004870 [PubMed: 9447441]

Myung IJ (2000). The importance of complexity in model selection. Journal of mathematical psychology, 44(1), 190–204. 10.1006/jmps.1999.1283 [PubMed: 10733864]

Peelle JE, Cooke A, Moore P, Vesely L, & Grossman M (2007). Syntactic and thematic components of sentence processing in progressive nonfluent aphasia and nonaphasic frontotemporal dementia. Journal of Neurolinguistics, 20(6), 482–494. 10.1016/j.jneuroling.2007.04.002 [PubMed: 18978927]

Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, & Nichols TE (2008). Guidelines for reporting an fMRI study. Neuroimage, 40(2), 409–414. 10.1016/j.neuroimage.2007.11.048 [PubMed: 18191585]

Rentoumi V, Raoufian L, Ahmed S, de Jager CA, & Garrard P (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology. Journal of Alzheimer's Disease, 42(s3), S3–S17. DOI: 10.3233/JAD-140555

Riva M, Wilson SM, Cai R, Castellano A, Jordan KM, Henry RG, & Chang EF (2022). Evaluating syntactic comprehension during awake intraoperative cortical stimulation mapping. Journal of Neurosurgery, 1(aop), 1–8. 10.3171/2022.8.JNS221335

Roiser JP, Linden DE, Gorno-Tempinin ML, Moran RJ, Dickerson BC, & Grafton ST (2016). Minimum statistical standards for submissions to Neuroimage: Clinical. NeuroImage: Clinical, 12, 1045. doi: 10.1016/j.nicl.2016.08.002 [PubMed: 27995071]

Rogalski E, Cobia D, Harrison TM, Wieneke C, Thompson CK, Weintraub S, & Mesulam MM (2011). Anatomy of language impairments in primary progressive aphasia. Journal of neuroscience, 31(9), 3344–3350. 10.1523/JNEUROSCI.5544-10.2011 [PubMed: 21368046]

Saffran EM, Berndt RS, & Schwartz MF (1989). The quantitative analysis of agrammatic production: Procedure and data. Brain and language, 37(3), 440–479. 10.1016/0093-934X(89)90030-8 [PubMed: 2804622]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sajjadi SA, Patterson K, Tomek M, & Nestor PJ (2012). Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. Aphasiology, 26(6), 847–866. 10.1080/02687038.2012.654933

Salimi-Khorshidi G, Smith SM, & Nichols TE (2011). Adjusting the effect of nonstationarity in cluster-based and TFCE inference. Neuroimage, 54(3), 2006–2019. 10.1016/j.neuroimage.2010.09.088 [PubMed: 20955803]

Sarawgi U, Zulfikar W, Soliman N, & Maes P (2020). Multimodal inductive transfer learning for detection of Alzheimer's dementia and its severity. 10.48550/arXiv.2009.00700.

Smith SM, & Nichols TE (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage, 44(1), 83–98. 10.1016/j.neuroimage.2008.03.061 [PubMed: 18501637]

Staffaroni AM, Ljubenkov PA, Kornak J, Cobigo Y, Datta S, Marx G, & Rosen HJ (2019). Longitudinal multimodal imaging and clinical endpoints for frontotemporal dementia clinical trials. Brain, 142(2), 443–459. [PubMed: 30698757]

Staffaroni AM, Weintraub S, Rascovsky K, Rankin KP, Taylor J, Fields JA, & Kramer JH (2021). Uniform data set language measures for bvFTD and PPA diagnosis and monitoring. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 13(1), e12148. 10.1002/dad2.12148

Stockbridge MD, Matchin W, Walker A, Breining B, Fridriksson J, Hickok G, & Hillis AE (2021). One cat, two cats, red cat, blue cats: eliciting morphemes from individuals with primary progressive aphasia. Aphasiology, 35(12), 1611–1622. 10.1080/02687038.2020.1852167

Stockbridge MD, Walker A, Matchin W, Breining BL, Fridriksson J, Hillis AE, & Hickok G (2021). A double dissociation between plural and possessive "s": Evidence from the Morphosyntactic Generation test. Cognitive neuropsychology, 38(1), 116–123. 10.1080/02643294.2020.1833851 [PubMed: 33096962]

Teichmann M, Kas A, Boutet C, Ferrieux S, Nogues M, Samri D, & Migliaccio R (2013). Deciphering logopenic primary progressive aphasia: a clinical, imaging and biomarker investigation. Brain, 136(11), 3474–3488. 10.1093/brain/awt266 [PubMed: 24108322]

Themistocleous C, Eckerström M, & Kokkinakis D (2020). Voice quality and speech fluency distinguish individuals with mild cognitive impairment from healthy controls. Plos one, 15(7), e0236009. [PubMed: 32658934]

Themistocleous C, Ficek B, Webster K, den Ouden DB, Hillis AE, & Tsapkini K (2021). Automatic subtyping of individuals with Primary Progressive Aphasia. Journal of Alzheimer's Disease, 79(3), 1185–1194.

Themistocleous C, Webster K, Afthinos A, & Tsapkini K (2021). Part of speech production in patients with primary progressive aphasia: An analysis based on natural language processing. American journal of speech-language pathology, 30(1S), 466–480. 10.1044/2020_AJSLP-19-00114 [PubMed: 32697669]

Thompson CK, Ballard KJ, Tait ME, Weintraub S, & Mesulam M (1997). Patterns of language decline in non-fluent primary progressive aphasia. Aphasiology, 11(4–5), 297–321. 10.1080/02687039708248473

Thompson CK, Cho S, Hsu CJ, Wieneke C, Rademaker A, Weitner BB, et al. (2012) Dissociations between fluency and agrammatism in primary progressive aphasia. Aphasiology, 26, 20–43. 10.1080/02687038.2011.584691 [PubMed: 22199417]

Thompson CK, & Mack JE (2014). Grammatical impairments in PPA. Aphasiology, 28(8–9), 1018–1037. 10.1080/02687038.2014.912744 [PubMed: 25642014]

Thompson CK, Meltzer-Asscher A, Cho S, Lee J, Wieneke C, Weintraub S, & Mesulam M (2013). Syntactic and morphosyntactic processing in stroke-induced and primary progressive aphasia. Behavioural neurology, 26(1–2), 35–54. 10.3233/BEN-2012-110220 [PubMed: 22713394]

Warrington EK, & James M (1991). The visual object and space battery perception. Bury St Edmunds: Thames Valley Company.

Weintraub S, Besser L, Dodge HH, Teylan M, Ferris S, Goldstein FC, & Morris JC (2018). Version 3 of the Alzheimer Disease Centers' neuropsychological test battery in the Uniform Data Set (UDS). Alzheimer disease and associated disorders, 32(1), 10. doi: 10.1097/WAD.0000000000000223 [PubMed: 29240561]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Wilson SM, Brandt TH, Henry ML, Babiak M, Ogar JM, Salli C, Wilson L, Peralta K, Miller BL, & Gorno-Tempini ML (2014). Inflectional morphology in primary progressive aphasia: An elicited production study. Brain and Language, 136, 58–68. 10.1016/j.bandl.2014.07.001 [PubMed: 25129631]

Wilson SM, Dronkers NF, Ogar JM, Jang J, Growdon ME, Agosta F, et al. (2010a). Neural correlates of syntactic processing in the nonfluent variant of primary progressive aphasia. J. Neurosci 30, 16845–16854. doi: 10.1523/JNEUROSCI.2547-10.2010 [PubMed: 21159955]

Wilson SM, Galantucci S, Tartaglia MC, & Gorno-Tempini ML (2012). The neural basis of syntactic deficits in primary progressive aphasia. Brain and Language, 122(3), 190–198. 10.1016/j.bandl.2012.04.005 [PubMed: 22546214]

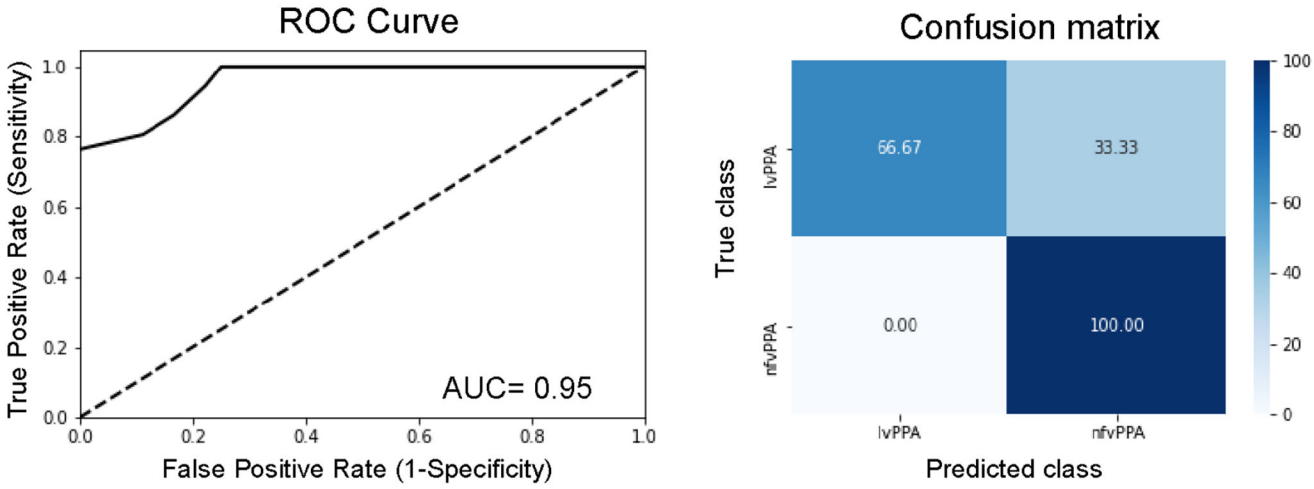
Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, et al. (2010b). Connected speech production in three variants of primary progressive aphasia. Brain 133, 2069–2088. doi: 10.1093/brain/awq129 [PubMed: 20542982]

Zaccarella E, & Friederici AD (2015). Merge in the human brain: A sub-region based functional investigation in the left pars opercularis. Frontiers in psychology, 6, 1818. 10.3389/fpsyg.2015.01818 [PubMed: 26640453]

Zaccarella E, & Friederici AD (2017). The neurobiological nature of syntactic hierarchies. Neuroscience & Biobehavioral Reviews, 81, 205–212. [PubMed: 27477445]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(a) Morphosyntax features: nfvPPA vs. lvPPA



(b) Standardized assessment: nfvPPA vs. lvPPA

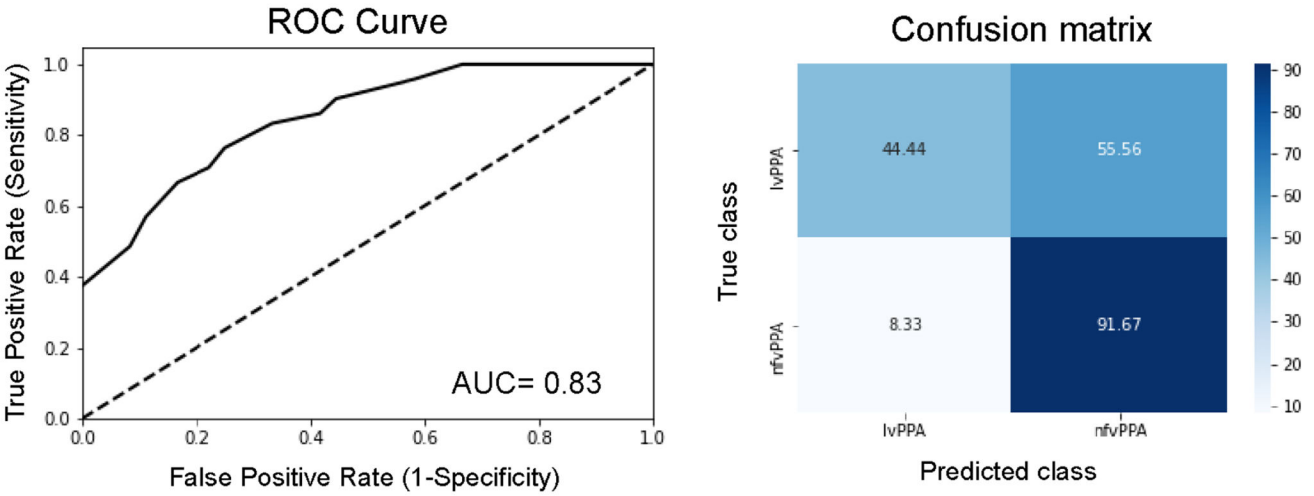
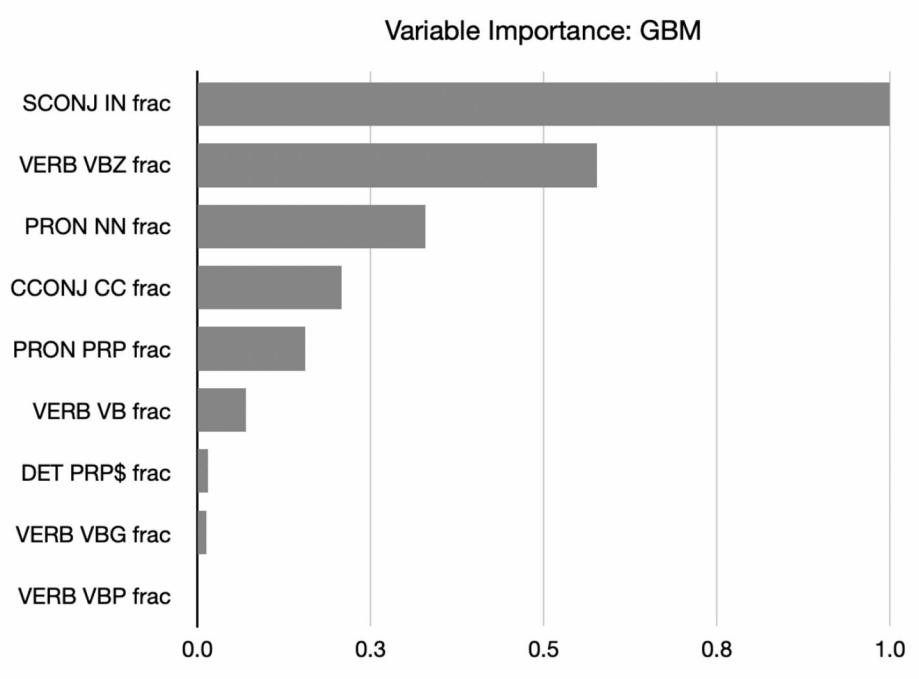


Fig. 1.

Mean receiver operating characteristic (ROC) curves with the mean area under the curve (AUC) and confusion matrix from the GBM classification model of two PPA variants. The best mean AUC resulted from GBM for nfvPPA versus lvPPA using morphosyntactic features.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(a)



(b)

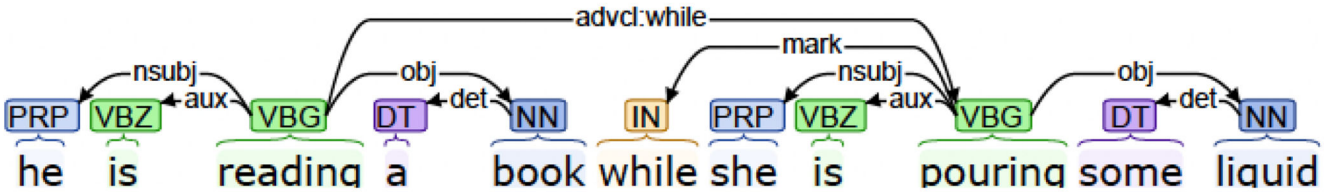


Fig. 2.

Feature importance results show *subordinating conjunction* as a top feature (a) and an example of *subordinating conjunction* with the POS tagger and the dependency tagger (b).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

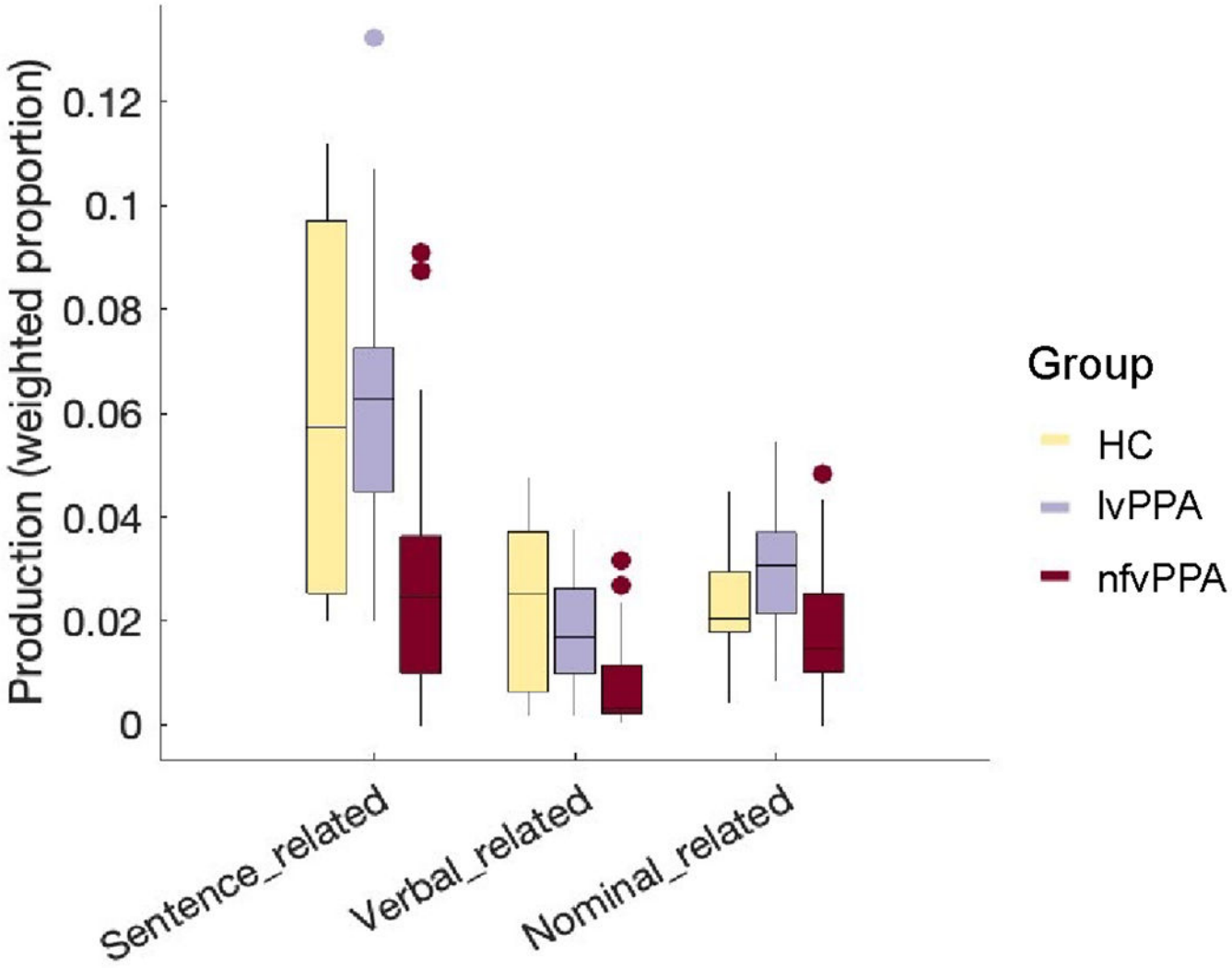


Fig. 3.
The weighted mean proportion (standard deviation) for each feature type for each group.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

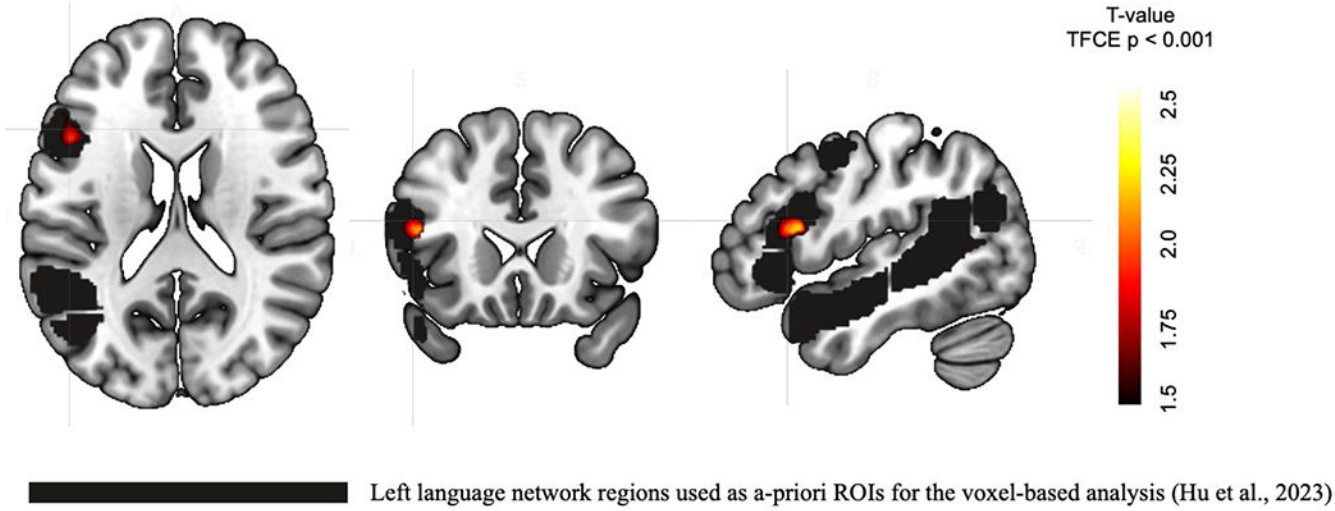


Fig. 4.

Gray matter loss correlates with the top morphosyntactic feature – *subordinating conjunction* in the left posterior IFGs pars opercularis across all participants using left language network regions a priori defined in Hu et al., 2023.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Demographic, neuropsychological, and speech-language data for individuals with nonfluent agrammatic and logopenic PPA variants (nfvPPA and lvPPA) and a group of healthy controls (HC). Group differences ($p < .05$; one-way ANOVA and Tukey HSD Test); Superscript letters indicate the group showing significant differences in pairwise TukeyHSD test: a: nfvPPA; b: lvPPA; c: HC. Sex, handedness, and ethnicity were compared with χ^2 tests; MMSE: Mini-Mental State Examination (Folstein et al., 1975); CDR: Clinical Dementia Rating (Morris, 1993); CVLT-SF: California Verbal Learning Test-UCSF version (Delis et al., 1987); WAB: Western Aphasia Battery (Kertesz, 1982); VOSP: Visual Object and Space Perception (Warrington & James, 1991); ND = normative data used for the following tests: CVLT-SF (scores extracted from a normative sample in our center with mean age: 54.7 ± 14.1 and education: 16.7 ± 2.2), PPVT: Peabody Picture Vocabulary Test (Dunn & Dunn, 1981; scores extracted from a normative sample in our center with mean age: 67.7 ± 4.1 and education: 18.0 ± 1.2), the WAB Repetition Total (scores extracted from the WAB-Revised manual), Syntax Comprehension (Wilson et al., 2010a; scores extracted from a normative sample in our center), and BDAE Sentence Comprehension (Goodglass, Kaplan, & Weintraub, 2001).

	nfvPPA	lvPPA	HC	*p*-value
Demographics				
N	35	26	10	
Age, mean (SD)	69.1 (8.1)	65.8 (8.6)	68.5 (6.2)	0.27
Education, mean (SD)	16.4 (2.4)	16.5 (2.4)	17.2 (1.4)	0.60
Sex, *n (%)* female	24 (68%)	15 (57%)	7 (70%)	0.63
Handedness, *n (%)* right	30 (86%)	19 (73%)	7 (70%)	0.52
Ethnicity, *n (%)* non-hispanics	32 (91%)	23 (88%)	10 (100%)	0.54
MMSE (max 30)	26.3 (2.6)	21.8 (5.0)ac	28.9 (2.2)	<.001
CDR Total	0.4 (0.3)	0.5 (0.2)	0.0 (0.0)ab	<.001
CDR Language	1.2 (0.6)	1.0 (0.5)	0.0 (0.0)ab	<.001
Neuropsychological assessment				
Digit Span Forwards	5.0 (1.1)	4.3 (0.8)a	7.6 (1.2)ab	<.001
Digit Span Backwards	3.7 (1.3)	3.2 (0.8)	5.4 (1.6)ab	<.001
Modified Trials (total time in seconds)	71.4 (36.4)b	95.8 (31.7)	23.5 (13.4)ab	<.001
Modified Trials (# of correct lines)	12.1 (3.7)	9.4 (4.9)ac	14.0 (0.0)	<.003
CVLT-SF Trials 1-4 (40)	24.1 (6.2)	14.0 (7.4)a	29.8 (3.4)ND	<.001
CVLT-SF 30 sec free recall (10)	6.7 (2.4)	3.4 (2.7)a	8.0 (1.1)ND	<.001
CVLT-SF 10 min free recall (10)	6.4 (2.1)	2.8 (2.8)a	7.5 (1.3)ND	<.001
Visuospatial: Benson figure copy (17)	14.9 (1.5)	14.9 (1.9)	15.3 (0.7)	0.76
Visual Memory: Benson figure recall (17)	10.5 (3.3)	6.3 (3.8)ac	10.9 (2.5)	<.001
Visual Object and Space Perception (10)	8.7 (1.5)	7.8 (1.8)a,c	9.1 (1.3)	<.001
Speech-language assessment				
Motor Speech rating (1-7)	2.1 (1.2)	0.0 (0.0)a	---	<.001
Boston (object) naming test (%)	87.6 (12.0)	66.4 (24.8)a	98.0 (3.2)ab	<.001
Phonemic (D-letter) fluency	6.4 (3.8)	8.7 (4.6)	15.3 (5.2)ab	<.001
Semantic (animal) fluency	12.3 (6.3)	9.7 (5.3)	22.0 (4.3)ab	<.001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	nfvPPA	lvPPA	HC	*p*-value
Peabody Picture Vocabulary Test (16)	14.7 (1.2)	13.8 (2.0)a	15.6 (0.5)[ND]	0.03
WAB Repetition (100)	88.5 (11.5)	73.4 (11.1)a	99.0 (1.0)[ND]	<.001
WAB Sequential Command (100)	70.9 (12.5)	66.0 (11.8)	---	0.12
Syntax Comprehension (%)	92.4 (8.6)	87.0 (11.6)a	98.6 (1.8)[ND]	0.04
Sentence Comprehension (BDAE 60-64 cards) (5)	4.0 (1.0)	3.8 (1.1)	---	0.33

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Parts of speech tags selected as features in the classification analysis.

Tag	Meaning	Example
SCONJ IN	subordinating conjunction	if, that, while, like
VERB VBZ	verb, 3rd person singular present	appears, looks
PRON NN	pronoun noun, singular or mass	something, someone
CCONJ CC	conjunction, coordinating	and, or, but
PRON PRP	pronoun personal	I, it, they, them
VERB VB	verb, base form	know, say
DET PRP$	determiner pronoun, possessive	his, her, their
VERB VBG	verb, gerund or present participle	drinking, getting
VERB VBP	verb, non-3rd person singular present	see, drink