

# UCSF

## UC San Francisco Previously Published Works

### Title

Clinical dimensions along the non-fluent variant primary progressive aphasia spectrum

### Permalink

<https://escholarship.org/uc/item/6qk57600>

### Journal

Brain, 147(4)

### ISSN

0006-8950

### Authors

Illán-Gala, Ignacio  
Lorca-Puls, Diego L  
Tee, Boon Lead  
[et al.](#)

### Publication Date

2024-04-04

### DOI

10.1093/brain/awad396

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed



# Clinical dimensions along the non-fluent variant primary progressive aphasia spectrum

Ignacio Illán-Gala,<sup>1,2,3</sup> Diego L. Lorca-Puls,<sup>4,5</sup> Boon Lead Tee,<sup>4</sup> Zoe Ezzes,<sup>4</sup> Jessica de Leon,<sup>4</sup> Zachary A. Miller,<sup>4</sup> Sara Rubio-Guerra,<sup>1</sup> Miguel Santos-Santos,<sup>1</sup> David Gómez-Andrés,<sup>6</sup> Lea T. Grinberg,<sup>3,4,7</sup> Salvatore Spina,<sup>4</sup> Joel H. Kramer,<sup>4</sup> Lisa D. Wauters,<sup>8</sup> Maya L. Henry,<sup>8</sup> Adam L. Boxer,<sup>4</sup> Howard J. Rosen,<sup>4</sup> Bruce L. Miller,<sup>4</sup> William W. Seeley,<sup>4</sup> Maria Luisa Mandelli<sup>4</sup> and Maria Luisa Gorno-Tempini<sup>4</sup>

It is debated whether primary progressive apraxia of speech (PPAOS) and progressive agrammatic aphasia (PAA) belong to the same clinical spectrum, traditionally termed non-fluent/agrammatic variant primary progressive aphasia (nfvPPA), or exist as two completely distinct syndromic entities with specific pathologic/prognostic correlates. We analysed speech, language and disease severity features in a comprehensive cohort of patients with progressive motor speech impairment and/or agrammatism to ascertain evidence of naturally occurring, clinically meaningful non-overlapping syndromic entities (e.g. PPAOS and PAA) in our data. We also assessed if data-driven latent clinical dimensions with aetiologic/prognostic value could be identified.

We included 98 participants, 43 of whom had an autopsy-confirmed neuropathological diagnosis. Speech pathologists assessed motor speech features indicative of dysarthria and apraxia of speech (AOS). Quantitative expressive/receptive agrammatism measures were obtained and compared with healthy controls. Baseline and longitudinal disease severity was evaluated using the Clinical Dementia Rating Sum of Boxes (CDR-SB). We investigated the data's clustering tendency and cluster stability to form robust symptom clusters and employed principal component analysis to extract data-driven latent clinical dimensions (LCD). The longitudinal CDR-SB change was estimated using linear mixed-effects models. Of the participants included in this study, 93 conformed to previously reported clinical profiles (75 with AOS and agrammatism, 12 PPAOS and six PAA). The remaining five participants were characterized by non-fluent speech, executive dysfunction and dysarthria without apraxia of speech or frank agrammatism. No baseline clinical features differentiated between frontotemporal lobar degeneration neuropathological subgroups. The Hopkins statistic demonstrated a low cluster tendency in the entire sample (0.45 with values near 0.5 indicating random data). Cluster stability analyses showed that only two robust subgroups (differing in agrammatism, executive dysfunction and overall disease severity) could be identified. Three data-driven components accounted for 71% of the variance [(i) severity-agrammatism; (ii) prominent AOS; and (iii) prominent dysarthria]. None of these data-driven LCDs allowed an accurate prediction of neuropathology. The severity-agrammatism component was an independent predictor of a faster CDR-SB increase in all the participants. Higher dysarthria severity, reduced words per minute and expressive and receptive agrammatism severity at baseline independently predicted accelerated disease progression. Our findings indicate that PPAOS and PAA, rather than exist as completely distinct syndromic entities, constitute a clinical continuum. In our cohort, splitting the nfvPPA spectrum into separate clinical phenotypes did not improve clinical-pathological correlations, stressing the need for new biological markers and consensus regarding updated terminology and clinical classification.

Received May 02, 2023. Revised October 21, 2023. Accepted November 05, 2023. Advance access publication November 21, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

- 1 Sant Pau Memory Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Biomedical Research Institute Sant Pau, Universitat Autònoma de Barcelona, 08025, Barcelona, Spain
- 2 Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas (CIBERNED), Madrid, 28029, Spain
- 3 Global Brain Health Institute, University of California, San Francisco, CA 94143, USA
- 4 Memory and Aging Center, Department of Neurology, UCSF Weill Institute for Neurosciences, University of California, San Francisco, CA 94158, USA
- 5 Sección de Neurología, Departamento de Especialidades, Facultad de Medicina, Universidad de Concepción, Concepción, 4070001, Chile
- 6 Vall d'Hebron Institut de Recerca (VHIR), Hospital Universitari Vall d'Hebron, 08035, Barcelona, Spain
- 7 Department of Pathology, University of California San Francisco, San Francisco, CA 94143, USA
- 8 Department of Communication Sciences and Disorders, University of Texas, Austin, TX 78712-0114, USA

Correspondence to: Dr Ignacio Illán-Gala  
 Sant Pau Memory Unit, Department of Neurology  
 Hospital de la Santa Creu i Sant Pau  
 Biomedical Research Institute Sant Pau  
 Universitat Autònoma de Barcelona, 08025, Barcelona, Spain  
 E-mail: iillan@santpau.cat

**Keywords:** apraxia of speech; dysarthria; primary progressive aphasia; corticobasal degeneration; progressive supranuclear palsy; magnetic resonance imaging

## Introduction

The non-fluent/agrammatic variant of primary progressive aphasia (nfvPPA) is a neurodegenerative clinical syndrome currently defined by effortful speech [mainly caused by apraxia of speech (AOS)], and varying degrees of expressive agrammatism.<sup>1</sup> Effortful speech and expressive agrammatism in the setting of relatively spared single-word comprehension and object knowledge represent the two core features of nfvPPA and can appear combined or in isolation.<sup>2</sup> But, in addition to AOS and agrammatic production errors, a wide range of changes in motor speech and grammar abilities can be noted in nfvPPA, and these features frequently overlap with other cognitive, speech and language deficits, like dysarthria or executive dysfunction.<sup>3</sup> Importantly, previous studies have documented the existence of distinguishable speech-language profiles in patients with nfvPPA arising from the differential expression of AOS versus expressive agrammatism.<sup>4–7</sup> For example, the term ‘progressive agrammatic aphasia’ or ‘agrammatic variant PPA’<sup>5,8</sup> has been used to refer to patients with expressive agrammatism as the most prominent presenting symptom in the absence of AOS. In contrast, patients exhibiting AOS as the most salient clinical feature without clear concomitant aphasia (including expressive agrammatism) have been ascribed the term ‘primary progressive apraxia of speech’ (PPAOS).<sup>6</sup> Some authors suggest that PPAOS and progressive agrammatic aphasia (PAA) should be considered separate syndromic entities because they may have specific aetiological and prognostic features. However, existing evidence does not clearly indicate whether these phenotypes represent a clinical continuum or exist as distinct syndromic entities with specific anatomical, pathological and prognostic correlates. It remains unclear whether syndromic definitions or data-driven approaches will significantly improve clinical-pathological correlations and disease progression.<sup>9</sup> Finally, to date, the most successful speech therapy approach to nfvPPA rehabilitation<sup>10</sup> targets both motor speech and grammar symptoms. This approach is justified by evidence from longitudinal studies showing that, even in cases where AOS and agrammatism initially present in isolation, they eventually co-occur as the disease progresses.

From a pathological perspective, large autopsy-proven studies have shown that patients presenting with both AOS and agrammatism, PPAOS and PAA are typically associated with progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD).<sup>11</sup> But, less frequently, these phenotypes are also the clinical presentation of Pick’s disease,<sup>12</sup> or other frontotemporal lobar degeneration (FTLD) subtypes, like the phosphorylated 43-kDa TAR DNA-binding protein inclusions (FTLD-TDP) subtypes.<sup>13</sup> Current syndromic classifications have proven helpful in discriminating between patients with FTLD versus Alzheimer’s disease,<sup>2,14,15</sup> but have failed to predict FTLD subtypes (i.e. PSP versus CBD) or longitudinal decline at the single-subject level.<sup>11</sup> Refining the classification of patients within the nfvPPA-spectrum could improve clinical-pathological correlations<sup>16</sup> and the prediction of progression rate, both of which are essential steps to designing successful trials targeting molecular mechanisms of pathology.<sup>17</sup>

Data-driven classifications allow the recognition of individual differences within phenotypes and may provide valuable insights for the refinement of diagnostic labels and the development of precision medicine approaches.<sup>18</sup> Unfortunately, previous studies exploring data-driven subtypes of nfvPPA were limited by their relatively small sample size and lack of longitudinal and autopsy data.<sup>19,20</sup> In this study, we aimed to characterize the speech, language and cognitive features among patients with a broad spectrum of non-fluent speech and/or language impairments and explore the existence of data-driven latent clinical dimensions (LCD) with aetiological and/or prognostic value.

We hypothesize that PAA and PPAOS represent a clinical continuum encompassing several endophenotypes with prognostic implications that could be useful to improve clinical-pathological correlations.

## Material and methods

### Participant selection

We searched the Memory and Aging Center of the University of California, San Francisco (MAC-UCSF) database to identify

participants that satisfied the following inclusion criteria: (i) completion of a comprehensive speech-language assessment (including connected speech samples)<sup>2</sup>; (ii) met at least one of the two core criteria of nfvPPA (namely, AOS or expressive agrammatism) at the moment of their speech and language assessment; and (iii) a clinical syndrome dominated by speech or language symptoms. Importantly, participants in this study were not required to exhibit prominent aphasia at the time of diagnosis. Consequently, participants not meeting the core diagnostic criteria for PPA<sup>2,21</sup> may have satisfied previous definitions of PPAOS.<sup>6</sup> In addition, some patients, who were considered to have mild expressive agrammatism at the moment of their clinical assessment, may not strictly meet the operational definition of expressive agrammatism used in this study (see later further details). We applied the same criteria as in Lorca-Puls *et al.*<sup>22</sup> to exclude participants at an advanced disease stage [i.e. those with a Mini-Mental State Examination (MMSE)<sup>23</sup> score  $\leq 10$  or, alternatively, a Clinical Dementia Rating (CDR)<sup>24</sup> global score  $\geq 3$ ] because their inclusion could bias the results. We also excluded participants meeting diagnostic criteria for the logopenic or semantic variants of PPA, or with prominent behavioural or motor symptoms meeting diagnostic criteria for the behavioural variant of frontotemporal dementia (FTD), PSP or CBD.<sup>25,26</sup> The diagnosis of all participants was imaging-supported and vascular disease was ruled out. Ninety-eight participants recruited at MAC-UCSF between September 1999 and January 2021 were included in the study. In participants with multiple clinical assessments, we selected the first visit with a complete speech-language assessment because clinical features of FTLN-related syndromes tend to converge with disease progression.<sup>3</sup>

## Genetic and neuropathological assessment

A neuropathological evaluation was available in a subgroup of the participants ( $n = 43$ ). Neuropathological diagnoses and genetic analyses were performed as previously described.<sup>27,28</sup> Briefly, participants were classified post-mortem into FTLN major molecular classes and subtypes, and genetic screening was conducted for mutations known to cause autosomal dominant FTLN or Alzheimer's disease in the participants giving their informed consent.<sup>29</sup>

## Clinical assessment

### General measures of cognitive and functional impairment

All participants underwent a complete clinical history, physical examination and neuropsychological evaluation. We used a previously reported battery of neuropsychological tests to assess major cognitive domains. In addition, all participants underwent a comprehensive standardized speech and language assessment, as previously described.<sup>30–32</sup> The CDR [available for 95 (97%) of the participants at baseline] and the CDR Staging Instrument PLUS National Alzheimer's Coordinating Center Behavior and Language Domains [CDR plus NACC FTLN, available for 84 (86%) of the participants at baseline] were adopted as general measures of disease severity.<sup>24,33</sup> The MMSE was also recorded at baseline and used as a general measure of cognitive impairment.<sup>23</sup> We also report the estimated age at symptom onset, sex, years of education and age at diagnosis. In participants with more than one visit, the longitudinal decline was characterized using the CDR sum-of-boxes (CDR-SB) when available. In longitudinal analyses, the CDR-SB was prioritized over the CDR plus NACC FTLN sum-of-boxes (CDR plus

NACC FTLN-SB) because the former resulted in fewer missing values.

### Assessing apraxia of speech and dysarthria through a structured motor speech evaluation

To elicit a wide range of motor speech behaviours, the participant was asked, as part of our motor speech evaluation (MSE),<sup>34</sup> to complete a collection of tasks such as vowel prolongation, alternating motion rate, sequential motion rate, multiple repetitions of monosyllabic words, multiple repetitions of multisyllabic words, repetition of words of increasing length, and the reading of a brief, phonetically balanced paragraph. Based on the observed motor speech ability of the patient, a certified speech-language pathologist assigned a clinical severity rating for AOS and separately for dysarthria on a scale from 0 (within normal limits) to 7 (profound). A list of deviant motor speech characteristics used to perceptually judge the presence and severity of AOS and/or dysarthria is provided in [Supplementary Table 1](#). Of note, inter-rater reliability for two independent raters (Z.E. and L.D.W., both of whom are certified speech-language pathologists) was established in a subset of 15 patients (15% of the sample) that were quasi-randomly selected according to diagnostic classification and severity. Relative to the first independent rater, this analysis yielded an intraclass correlation coefficient (ICC) of 0.86 for AOS and 0.81 for dysarthria. As for the second independent rater, ICCs were 0.85 for AOS and 0.77 for dysarthria. According to the guidelines for interpreting ICCs provided by Cicchetti, these results indicate excellent inter-rater agreement.<sup>35</sup>

### Assessing expressive agrammatism and speech fluency through a picture description task

To further characterize grammar processing abilities, we enriched the standard speech and language protocol with additional quantitative measures of syntactic ability and speech fluency by analysing connected speech samples of the participants. To assess expressive agrammatism, the participant was prompted to describe a visual scene in as much detail as possible by using sentences. Audio-recorded connected speech samples describing the 'picnic' scene from the Western Aphasia Battery<sup>36</sup> were sent to [www.saltsoftware.com](http://www.saltsoftware.com) for transcription, coding and analysis. By running the coded transcripts through the Systematic Analysis of Language Transcripts (SALT) software, a set of measures was generated, of which we selected exclusively those capturing the accuracy and complexity of the sentences produced by the patient (consistent with the definition of expressive agrammatism) and the number of words per minute as a general measure of speech fluency/rate. Four SALT-derived measurements were assessed: (i) % utterances with omission and/or commission errors (%UtWErrors), a measure of morphosyntactic accuracy that captures missing function/content words, omitted bound morphemes, inappropriate word choice, and/or incorrect morphosyntactic form; (ii) mean length of utterance (MLU), a measure of morphosyntactic complexity that captures mean sentence length in words; (iii) subordination index (SI), a measure of morphosyntactic complexity that captures the ratio of the total number of clauses to the total number of utterances; and (iv) words per minute (WPM) as a general measure of speech fluency. Only complete (not abandoned or interrupted), intelligible (without any unintelligible segments) and verbal utterances (that contained at least one verbalized word) contributed to the calculation of these quantitative, continuous measures.

### Assessing receptive agrammatism through a sentence comprehension task

To assess receptive agrammatism, the participant was prompted to complete either of two auditory sentence-to-picture matching tasks that are part of our comprehensive speech-language assessment battery. The first task involved a representative range of sentence types, varying in both length and complexity, taken from the Curtiss Yamada Comprehensive Language Evaluation (CYCLE), as previously reported.<sup>37</sup> Each patient was instructed to match the meaning of an auditorily-presented sentence with the corresponding line drawing in a three- or four-picture array. The second task was loosely based on the first and has been previously described.<sup>38</sup> To reduce the number of participants with missing data for receptive agrammatism, we also considered the syntax comprehension score from our bedside neuropsychological battery. This score includes five sentences from the Boston Diagnostic Aphasia Examination, ranges from 0 (worst) to 5, and was moderately correlated with the other sentence comprehension tasks ( $\rho = 0.650$ ,  $P < 0.001$ ).

### Assessing other relevant cognitive, behavioural and motor aspects

Because speech and language impairment could impact neuropsychological performance in tests with a high verbal load, we distinguish between verbal and non-verbal tests when defining the different cognitive domains. Hence, memory was assessed with the delayed recognition of words from the California Verbal Learning Test<sup>39</sup> and the delayed recall of the Benson figure.<sup>40</sup> Visuospatial ability was assessed with the number location subtest of the visual object space perception battery and the copy of the Benson figure.<sup>40</sup> Verbal measures of executive dysfunction included reverse digit span and the Stroop test,<sup>41</sup> while non-verbal measures included the design fluency subscale of the Delis-Kaplan Executive Function Scale and the number of correct trials in 1 min from the modified Trail-making Test.<sup>42,43</sup> Naming and word comprehension were assessed with the 15-item version of the Boston Naming Test,<sup>44</sup> and the Peabody Picture Vocabulary Test, respectively. Repetition was assessed with the Western Aphasia Battery (WAB). The caregiver-distress scores of the Neuropsychiatric Inventory (NPI) were recorded to characterize behavioural changes.<sup>45</sup> Motor changes were characterized with the Progressive Supranuclear Palsy Rating Scale (PSPRS, available in 58 participants)<sup>46</sup> and the Unified Parkinson's Disease Rating Scale part III (UPDRS-III, available in 47 participants).<sup>47</sup> We also recorded the existence of oculomotor dysfunction (at least mild slowing of vertical saccades) and postural instability (at least three to five steps in the pull test).

### Definition of speech, language and cognitive domains

To create a single index of expressive grammar ability, the %UtWErrors, MLU and SI scores were combined after their normalization. Similarly, scores from the two sentence comprehension tasks included in our speech-language assessment battery were combined with the syntax comprehension score from the bedside neuropsychological battery to create an index of receptive grammar ability. In addition, cognitive variables measuring non-overlapping neuropsychological domains were grouped based on a *priori* knowledge from the authors and their observed correlation. These domains include: (i) verbal measures of executive function

(reverse digit span and the Stroop test); (ii) non-verbal measures of executive function (design fluency subscale of the Delis-Kaplan Executive Function Scale, and the number of correct trials in 1 min at the modified Trail-making Test); (iii) verbal and non-verbal measures of memory (the California Verbal Learning Test and the delayed recall of the Benson figure); and (iv) non-verbal measures of visuospatial function (the number location subtest of the visual object space perception battery and the copy of the Benton figure). The Z-scores of variables within the same domain were averaged to maximize the number of participants with at least one measurement for each clinical domain. By grouping variables taxing similar neuropsychological domains, we avoided both data imputations and the exclusion of participants with missing data in a few variables.

### Definition of impaired performance

MSE ratings for AOS or dysarthria  $>0$ , as judged by a certified speech-language pathologist, were considered impaired. However, the performance of an individual patient on the picture description and sentence comprehension tasks cannot be labelled as 'impaired' without reference to a normative sample of neurologically intact controls. For these tasks, we used a previously validated Bayesian method to determine if a patient's task performance fell within the impaired range (while covarying out the effects of age, sex and years of education).<sup>48</sup> SALT-derived measures of expressive agrammatism (%UtWErrors, MLU and SI) and WPM were considered individually relative to a group of 18 neurologically intact controls. For receptive agrammatism, we used for comparison a group of 10 neurologically intact controls who completed the first sentence-to-picture matching task and another group of 26 neurologically intact controls who completed the second sentence-to-picture matching task, both from our comprehensive speech-language assessment battery. Detailed characteristics of these controls can be found in [Supplementary Table 2](#). For the syntax comprehension score and individual neuropsychological measurements, we resorted to previously published normative data to calculate age- and education-adjusted Z-scores.<sup>49</sup> Consistent with previous studies, Z-scores  $< -1.5$  were considered abnormal.<sup>49</sup> The output of these analyses allowed us to obtain a conservative estimate of the frequency of occurrence of expressive and receptive agrammatism across patients. For each clinical domain, a patient was labelled as having impaired performance if one or more of the scores assigned to that domain were abnormal (e.g. we considered that a patient had evidence of expressive agrammatism if either %UtWErrors, MLU or SI were impaired). We used Venn diagrams to illustrate the frequency and partial overlap of distinct deficits.

### Classification of participants into clinically defined subgroups based on major speech and language characteristics

Participants with AOS but without expressive agrammatism were classified as PPAOS, while participants with expressive agrammatism but without AOS were classified as PAA, according to previous work.<sup>5,6</sup> Participants with both AOS and expressive or receptive agrammatism meeting current consensus criteria<sup>2</sup> were classified as AOS + agrammatism. Five participants characterized by reduced speech fluency and dysarthria in the absence of AOS and quantitative evidence of agrammatism remained unclassifiable and were therefore assigned to a fourth group named 'non/fluent dysarthric' (nf-Dysarthric). However, we highlight that all five participants in

this group were reported to exhibit subtle features consistent with the presence of a mild form of expressive agrammatism as clinically rated, thus meeting diagnostic criteria for nfvPPA.

### Assessing clustering tendency

To characterize patterns of performance across participants, we employed a hierarchical heat map—a visualization technique renowned for its efficacy in discerning clusters of participants and distinct clinical features, especially within the realm of heterogeneous diseases.<sup>50</sup> This heat map displays a colour-coded matrix representing individual data juxtaposed with clinical features. The participants and the clinical features are organized using hierarchical clustering, ensuring closely related entities are juxtaposed for clearer pattern recognition. For the generation of this heat map, our initial step involved computing an Euclidean distance matrix encompassing all the pivotal speech, language and cognitive metrics outlined in this research. Subsequently, leveraging the ‘complete’ linkage method for hierarchical clustering, we processed the data using the ‘hclust’ function from the ‘stats’ package (version 4.1.1). To evaluate the clusterability of the dataset, we represented the similarity matrix, which is a visual method of how naturally the data fall into distinct groups.

We also quantified the overall clustering tendency of raw (non-binary) data with the Hopkins statistic (‘factoextra’ package, v1.0.7),<sup>37</sup> to determine if participants within the nfvPPA-spectrum tended to cluster into distinct naturally occurring, clinically meaningful subgroups (i.e. a non-random structure). The Hopkins statistic represents the probability that a given dataset is generated by a uniform or random data distribution<sup>51</sup> (values close to 0.5 indicate a low clustering tendency), while values far >0.5 and close to 1 suggest the existence of robust clusters.

### K-means clustering and cluster stability

To evaluate the robustness and reliability of multiple k-means clustering, we used the bootstrap resampling technique (i.e. sampling with replacement) from the ‘fpc’ package in R. Specifically, we generated 100 bootstrap samples from a subsample of our dataset without missing values [ $n = 90$  (92% of the total sample)] (refer to the ‘Principal component analysis’ section for further details), where each sample was drawn with replacement. For each resampled dataset, the k-means clustering algorithm was applied to determine a pre-specified number of clusters, with the aim of comparing these clusters to those derived from the original dataset. The stability of the clustering solution was quantified using the Jaccard coefficient, which measures the similarity between two sets. In this context, it assessed the degree of overlap between clusters identified in the original data and those from the bootstrap samples. The value of the Jaccard coefficient ranges between 0 and 1, where 0 indicates no similarity between the resampled and the original dataset and 1, an absolute overlap. Additionally, we monitored the frequency with which each cluster dissolved (completely disappeared) or was recovered (identified with a composition closely resembling its structure in the original dataset) across the bootstrap samples.<sup>52</sup>

### Principal component analysis

We then performed a principal component analysis (PCA) as a data-driven approach to extract new LCDs. Because PCA and k-means clustering cannot be performed in samples containing missing data, we considered a subsample of participants without missing

values for k-means clustering and PCA analyses to avoid data imputation as a source of bias. This subsample included 90 participants (92% of the total sample). Importantly, we employed PCA despite the observed low clustering tendency of the dataset because the existence (or lack) of naturally occurring, clinically meaningful symptom clusters in a dataset does not equate to the existence (or lack) of underlying clinical dimensions with aetiological/prognostic value and vice versa. Based on Bartlett’s test of sphericity, the Kaiser-Meyer-Olkin value, and *a priori* knowledge from the researchers, the following relevant features were entered into the PCA: AOS, dysarthria, speech fluency (WPM), expressive agrammatism index, receptive agrammatism index, the non-verbal and verbal executive function indexes, and CDR-SB. By selecting these variables, the Bartlett’s test of sphericity [ $\chi^2(10) = 50.7$ ,  $P < 0.001$ ] and the Kaiser-Meyer-Olkin value (0.71) indicated that the sample size and the correlation between the selected variables were suitable for PCA analysis.<sup>53,54</sup> Conversely, when including a higher number of features, the Bartlett’s test of sphericity and the Kaiser-Meyer-Olkin discouraged PCA analyses. Three components were selected based on Cattell’s criteria.<sup>55</sup> To explore the potential value of data-driven LCDs to differentiate underlying aetiology in our patient sample, we explored if the individual loadings of the extracted components (i.e. LCDs) allowed the discrimination between neuropathological subtypes in the subgroup of participants with autopsy data.

### Longitudinal analyses

Next, we investigated longitudinal decline with linear mixed-effects models. In longitudinal studies, linear mixed-effects models have proven to be powerful tools for identifying variables where baseline values are associated with different rates of change in clinical decline.<sup>56</sup> We fitted a linear mixed-effects model controlling for age, sex and the most relevant clinical features to estimate clinical decline over time (as measured by CDR-SB or CDR plus NACC FTLD-SB) across our patient sample. We also fitted a linear mixed-effects model controlling for age, sex and individual loadings for LCDs derived from the PCA. All linear mixed-effects models included a random patient-specific intercept and a random patient-specific slope. These random effects account for patient heterogeneity in baseline CDR-SB (or CDR plus NACC FTLD-SB) and its rate of change that is not explained by the predictors in the model. A term for clinical feature  $\times$  time interaction was used to study the association between the baseline clinical feature and CDR-SB (or CDR plus NACC FTLD-SB) over time. As in similar previous studies,<sup>57</sup> all linear mixed models were designed with a compound symmetry covariance structure (owing to the relative homogeneity in the covariance of effects). Of note, we obtained the same results when linear mixed models were fitted with unstructured covariance.

### Standard protocol approval, registration and patient consent

The study was approved by the Institutional Review Board of UCSF and was conducted following the Declaration of Helsinki. All participants gave their written informed consent to participate in the study.

### Statistical analysis

We compared the raw clinical measurements between clinical subgroups in the whole sample using t-test or Kruskal-Wallis test for continuous variables and the chi-square test for categorical data.

We also compared the main clinical features between the groups of participants with and without autopsy data to verify that participants without autopsy data were clinically similar to those with a definitive neuropathological diagnosis. Next, in the subgroup of participants with an autopsy-proven diagnosis, we compared the clinical features between neuropathological categories with at least three participants (namely, PSP, CBD, Pick's disease and FTLD-TDP type B). To support the definition of speech, language and cognitive domains, we explored the relationship between speech, language and cognitive measures with Spearman's correlation (refer to the 'Definition of speech, language and cognitive domains' section for further details). Statistical significance for all tests was set at 5% ( $\alpha = 0.05$ ), all statistical tests were two-sided, and all group comparisons were corrected for multiple comparisons using the false discovery rate. All statistical analyses were conducted in R v4.1.1.

## Results

### Baseline characteristics of the sample

Table 1 shows the demographics and main clinical, genetic and neuropathological features of the 98 participants included in this study. The mean (standard deviation, SD) age at diagnosis was 68.0 (7.1) and 63 participants (64%) were female. The mean (SD) time from estimated symptom onset to diagnosis was 4.3 (2.0) years and almost 50% of the participants had a global score on the CDR plus NACC FTLD of 0.5 (indicating a mild/prodromal stage). Only six participants had a global score on the CDR plus NACC FTLD of 2 (Supplementary Table 3). A definitive neuropathological diagnosis was available in 43 participants (44% of the sample). CBD and PSP were the two most frequently observed neuropathological diagnoses [ $n = 17$  (40%) and  $n = 11$  (26%), respectively]. However, 15 (34%) had other neuropathological diagnoses (Table 1). A pathogenic mutation in the GRN gene was found in five (5%) of the participants (three of whom had autopsy confirmation of FTLD-TDP type A). Of note, participants with autopsy did not differ in terms of age at symptom onset, age at MRI, sex distribution, years of education, handedness distribution, disease severity, and NPI total score from participants without an autopsy-proven diagnosis (Supplementary Tables 5 and 6).

### Frequency of the main clinical features

As expected, the core features of nfvPPA were observed in a high proportion of participants (Supplementary Table 7). AOS and expressive agrammatism were present in 89% and 70% of the participants, respectively. Reduced speech fluency (WPM) and reduced phonemic (letter) fluency were also frequently observed (89% and 84%, respectively). Overall, Fig. 1 illustrates the significant overlap between the main speech, language and cognitive features of participants within the nfvPPA-spectrum. Motor speech deficits (either AOS or dysarthria) were noted in 97% of the participants (Fig. 1A), while agrammatism (either expressive or receptive) was observed in 85% (Fig. 1B). Notably, deficits in executive function were also detected in 66% of the participants when using non-verbal tests (Supplementary Table 7).

### Frequency of previously reported clinical phenotypes

Across the nfvPPA-spectrum, 75 participants (77%) had both AOS and agrammatism (either expressive or receptive), 12 participants (12%) had AOS without agrammatism (PPAOS), while six

participants (6%) had expressive agrammatism without AOS (PAA) (Table 1). Five participants (5%) did not exhibit AOS or quantitative evidence of agrammatism, as operationally defined in this study. Instead, they were characterized by the presence of diminished speech fluency, dysarthria and executive dysfunction (Supplementary Table 7). These participants were assigned to a fourth group named 'non/fluent dysarthric' (nf-Dysarthric). Notably, all these phenotypes did not differ regarding UPDRS-III and PSPRS total scores or the frequency of oculomotor dysfunction or postural instability (Supplementary Table 4).

### Comparison of clinical and neuropathological features between phenotypes

As shown in Table 1, the AOS + agrammatism, PPAOS, PPA and nf-Dysarthric subgroups did not differ in terms of age at diagnosis and estimated time from symptom onset to diagnosis. However, participants in the PPAOS group achieved better general cognitive performance (Table 1 and Supplementary Table 8). Crucially, the AOS + agrammatism, PPAOS, PAA and nf-Dysarthric subgroups were comparable regarding their neuropathological correlates (Table 1). Similarly, none of the baseline clinical features discriminated between FTLD subgroups after accounting for multiple comparisons (Supplementary Table 9).

### Cluster tendency analysis

Next, we explored the existence of naturally occurring, clinically meaningful symptom clusters across our patient group. The visual inspection of the clustered heat map and the similarity matrix failed to reveal non-overlapping clusters of participants based on their major clinical features (Fig. 2 and Supplementary Fig. 1). As shown in Fig. 2, despite the absence of subgroups with clear-cut boundaries, participants clustered in the first arm of the tree map (in the upper part of the heat map) appeared to be less impaired in measures of expressive and receptive agrammatism, executive function and verbal fluency than the participants included within the second arm of the heat map (lower part of the heat map). The low clustering tendency was also supported by the Hopkins statistic (0.454,  $P < 0.001$ ). We also confirmed that neuropathological subtypes were similarly distributed along the nfvPPA-spectrum rather than associated with specific clinical phenotypes, as illustrated in Fig. 2.

### Cluster stability analysis

We then assessed the cluster stability of the k-means clustering solution. We tested the clustering stability for multiple pre-specified numbers of clusters. When we pre-specified the existence of two clusters, cluster formation showed notably consistent stability. But, while the first cluster was recovered with high fidelity in 98% of the bootstrap samples, the second cluster's exact composition was consistent with the original solution in 82% of the resamples. This indicates that while both clusters were reliably present, the second cluster showed more variation in its composition across bootstrap samples relative to the original solution. Clustering was mainly driven by severity (Supplementary Tables 10 and 11). The first cluster group was characterized by reduced verbal fluency, agrammatism (both expressive and receptive), executive dysfunction and higher disease severity (as measured by the CDR-SB), and the second group was characterized by the opposite pattern. Notably, solutions showed extreme instability when we pre-specified more than two clusters (Supplementary Table 12).

Table 1 Demographic and clinical characteristics of participants

Characteristic	n	All participants, n = 98	Clinical subgroups				P-value
			nfvPPA, n = 75	PPAOS, n = 12	Non-fluent dysarthric, n = 5	PAA, n = 6	
Age at diagnosis, years	98	68.0 (7.1)	68.7 (7.0)	64.2 (8.7)	68.9 (4.8)	66.2 (5.7)	0.5
Age at estimated symptom onset, years	83	63.1 (7.1)	63.7 (7.1)	58.1 (7.8)	65.6 (4.7)	61.7 (6.0)	0.5
Time from estimated symptom onset to diagnosis, years	83	4.3 (2.0)	4.4 (1.9)	3.3 (1.9)	3.3 (1.7)	4.5 (2.5)	0.5
Biological sex	98	–	–	–	–	–	0.9
Female	–	63/98 (64%)	49/75 (65%)	8/12 (67%)	3/5 (60%)	3/6 (50%)	–
Male	–	35/98 (36%)	26/75 (35%)	4/12 (33%)	2/5 (40%)	3/6 (50%)	–
Years of education	98	16.2 (3.2)	16.3 (3.1)	15.2 (2.8)	15.0 (2.0)	18.0 (5.2)	0.7
Handness	98	–	–	–	–	–	0.8
Right-handed	–	86/98 (88%)	64/75 (85%)	12/12 (100%)	5/5 (100%)	5/6 (83%)	–
Left-handed	–	11/98 (11%)	10/75 (13%)	0/12 (0%)	0/5 (0%)	1/6 (17%)	–
Ambidextrous	–	1/98 (1.0%)	1/75 (1.3%)	0/12 (0%)	0/5 (0%)	0/6 (0%)	–
MMSE/30	92	25.6 (4.1)	25.3 (4.0)	29.0 (0.9)	27.0 (2.3)	22.8 (6.4)	0.02
CDR® Sum of Boxes	95	1.9 (1.8)	1.9 (1.9)	1.0 (1.1)	2.7 (1.5)	2.5 (1.9)	0.5
CDR® plus NACC FTLD-SB	84	3.6 (2.4)	3.6 (2.6)	2.6 (1.3)	4.5 (2.0)	4.5 (2.0)	0.5
Global score CDR® plus NACC FTLD	84	–	–	–	–	–	0.9
0.5	–	40/84 (48%)	32/64 (50%)	5/10 (50%)	1/4 (25%)	2/6 (33%)	–
1	–	38/84 (45%)	26/64 (41%)	5/10 (50%)	3/4 (75%)	4/6 (67%)	–
2	–	6/84 (7.1%)	6/64 (9.4%)	0/10 (0%)	0/4 (0%)	0/6 (0%)	–
NPI total score	88	14.9 (13.0)	14.9 (13.3)	10.7 (10.6)	22.2 (10.3)	17.0 (16.9)	0.5
UPDRS-III, total score	47	14.7 (11.3)	14.4 (11.3)	2.0 (1.4)	18.0 (3.0)	19.8 (13.8)	0.5
PSPRS, total score	57	12.4 (10.1)	12.8 (10.2)	6.9 (5.8)	22.0 (18.4)	14.0 (10.2)	0.5
Mutation	98	–	–	–	–	–	0.9
No mutation	–	79/98 (81%)	59/75 (79%)	9/12 (75%)	5/5 (100%)	6/6 (100%)	–
Not screened	–	16/98 (16%)	13/75 (17%)	3/12 (25%)	0/5 (0%)	0/6 (0%)	–
GRN	–	3/98 (3.1%)	3/75 (4.0%)	0/12 (0%)	0/5 (0%)	0/6 (0%)	–
Neuropathological diagnosis	43	–	–	–	–	–	0.7
Corticobasal degeneration	–	17/43 (40%)	14/33 (42%)	1/3 (33%)	0/3 (0%)	2/4 (50%)	–
Progressive supranuclear palsy	–	11/43 (26%)	6/33 (18%)	1/3 (33%)	3/3 (100%)	1/4 (25%)	–
Pick's disease	–	7/43 (16%)	5/33 (15%)	1/3 (33%)	0/3 (0%)	1/4 (25%)	–
FTLD-TDP type A	–	4/43 (9.3%)	4/33 (12%)	0/3 (0%)	0/3 (0%)	0/4 (0%)	–
Other pathologies	–	4/43 (9.3%)	4/33 (12%)	0/3 (0%)	0/3 (0%)	0/4 (0%)	–

Unless otherwise indicated, results are expressed as mean (standard deviation, SD). The P-values are adjusted for multiple comparisons (false discovery rate, FDR) and statistically significant P-values are highlighted in bold. The 'non-fluent dysarthric' subgroup included five participants with reduced speech fluency, executive dysfunction and dysarthria but without apraxia of speech or quantitative evidence of expressive agrammatism. 'Other pathologies' included one participant with an unclassifiable tauopathy, one participant with Alzheimer's disease, one with mixed Alzheimer's disease and corticobasal degeneration, and one with FTLD-TDP type B. CDR = Clinical Dementia Rating; FTLD = frontotemporal lobar degeneration; MMSE = Mini-Mental State Examination; nfvPPA = non-fluent/agrammatic variant primary progressive aphasia; PPAOS = primary progressive apraxia of speech; PAA = primary agrammatic aphasia; NPI = neuropsychiatric inventory; PSPRS = progressive supranuclear palsy rating scale; UPDRS-III = Unified Parkinson's Disease Rating Scale part III.

## Principal component analysis

Because classifying participants into discrete clinical entities (i.e. syndromes) based on major speech and language characteristics was not useful for identifying the underlying pathology in our sample, we explored if PCA could unveil LCDs with aetiologic value. Three LCDs explained 72% of the variance in the dataset. As shown in Fig. 3, the first component explained 37% of the variance and reflected reduced speech fluency (WPM), executive dysfunction (both verbal and non-verbal), agrammatism (both expressive and receptive) and overall cognitive and functional impairment (as measured by the CDR-SB). The second component explained 19% of the variance and was defined by prominent AOS with dysarthria and lesser cognitive and functional impairment. The third component explained 12% of the variance and was mainly associated with dysarthria but not AOS.

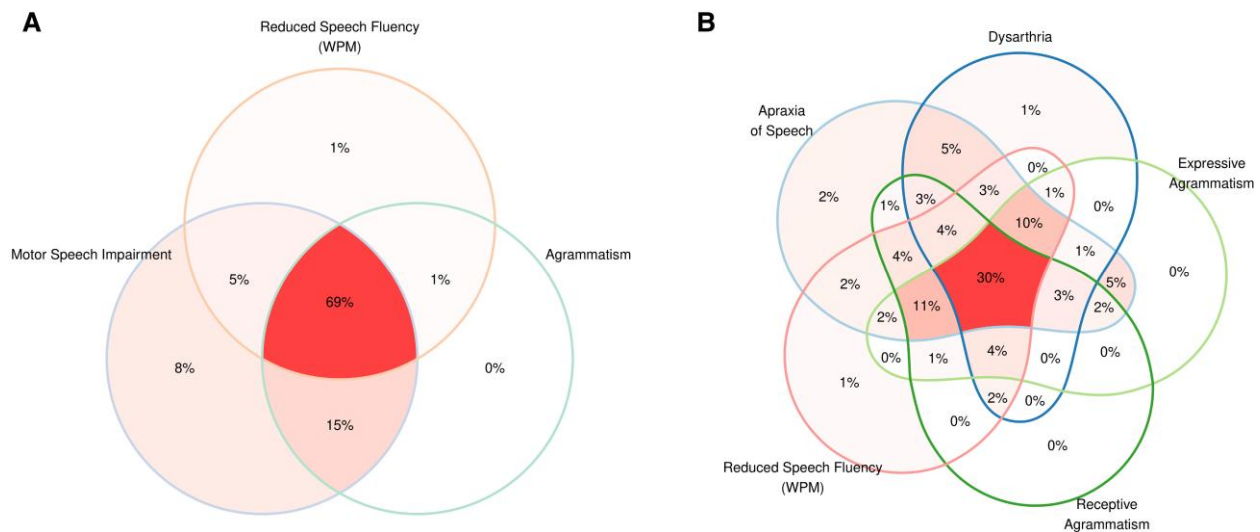
Next, we explored the neuropathological correlates of the three LCDs (from the PCA) in the subgroup of participants with autopsy

data. As shown in Fig. 4, the first and second dimensions did not yield significant differences between the neuropathological subgroups. However, the LCD reflecting prominent dysarthria was significantly reduced in participants with PSP when compared to participants with CBD, Pick's disease or FTLD-TDP type A.

## Longitudinal analyses

When considering baseline raw data, higher scores for dysarthria, reduced speech fluency (WPM) and higher expressive and receptive agrammatism were associated with an increased rate of clinical decline as measured with both the CDR-SB and the CDR plus NACC FTLD-SB (Supplementary Table 13). When considering the components derived from the PCA (Fig. 5A and Supplementary Table 14), only the LCD characterized by higher disease severity and worst agrammatism was associated with an increased rate of clinical decline (Fig. 5B). Of note, the LCD characterized by prominent AOS was associated with lower disease severity at baseline, supporting





**Figure 1 Frequency of main clinical features.** Frequency of main clinical features in the 98 participants included in this study. Venn diagram illustrating the overlap between motor speech impairment, agrammatism (either expressive or receptive) and reduced speech fluency (A), apraxia of speech, dysarthria, expressive and receptive agrammatism, and reduced speech fluency (words per minute, WPM) (B). Only one participant was not found to have impaired performance according to the thresholds defined in this study in any of the main clinical features included in this figure. This participant was diagnosed with non-fluent/agrammatic variant primary progressive aphasia (nfvPPA) because the treating physician noted mild expressive agrammatism.

the view that participants with relatively isolated AOS are typically diagnosed at an earlier disease stage than the rest of the participants.

## Discussion

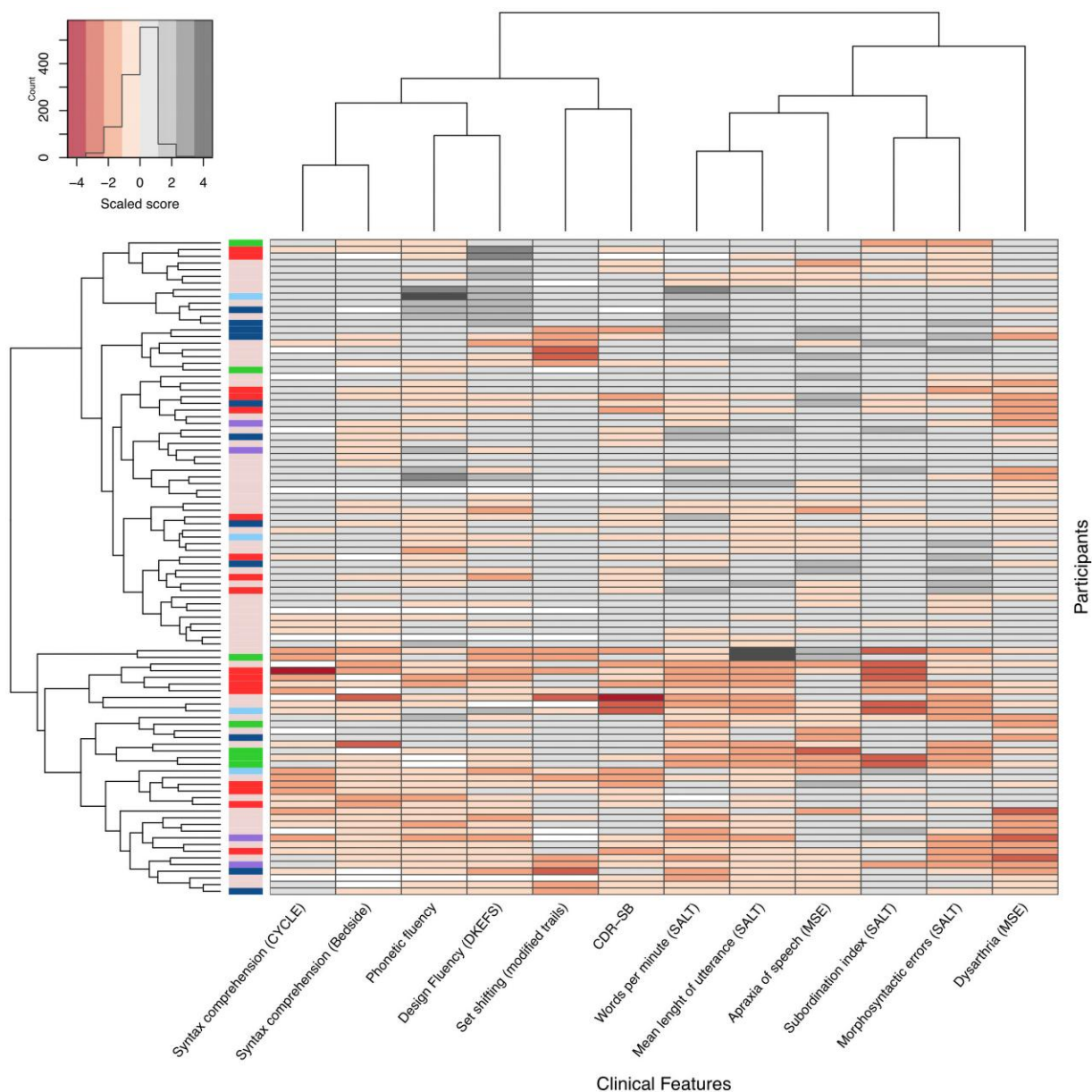
In this study, we leveraged a large cohort of participants within the nfvPPA-spectrum to examine the presence of robust endophenotypes and their potential to enhance clinical-pathological correlations and disease progression. Our findings indicate that this clinical entity embodies a continuous spectrum of substantially overlapping cognitive, speech and language characteristics. Although data-driven exploration of LCDs did not prove beneficial in predicting the underlying pathology, the clinical dimension characterized by reduced speech fluency (WPM), agrammatism and increased cognitive and functional impairments emerged as a significant predictor of faster rate of decline.

A noteworthy aspect of the present study involves the objective quantification of expressive and receptive agrammatism in an extensive cohort of participants within the nfvPPA-spectrum. Our results suggest that agrammatism plays an important prognostic role, but the multifaceted nature of grammatical processing renders it challenging to assess in routine clinical practice. The absence of standardized tests for evaluating expressive and receptive agrammatism and discrepancies in defining the threshold for 'significant' agrammatism may account for the variation in reported prevalence observed in prior studies. In our sample, up to 85% of participants demonstrated varying degrees of expressive or receptive agrammatism. Concurrently, motor speech deficits, such as AOS or dysarthria, were nearly universal. Consistent with previous studies, we also identified participants with motor speech deficits but without objective evidence of expressive agrammatism (as defined in this study). Some of these participants exhibited marked AOS, while others were characterized by reduced speech fluency, executive dysfunction and dysarthria. This latter group was predominantly observed in patients with PSP, with 3/3

(100%) of participants in this group having autopsy-confirmed PSP. We recognize that subtle differences in the subjective evaluation of AOS and dysarthria (Supplementary Table 1) may influence participant classification; for example, it is unclear if some of these participants could have been diagnosed with the prosodic subtype AOS by other research groups. Further studies are therefore needed to investigate the degree to which the classification of AOS (including potential subtypes) and dysarthria, based on expert-driven auditory-perceptual assessments, can be replicated across various research centres and countries.

Another critical finding of our study is that participants within the nfvPPA-spectrum cannot be robustly clustered into multiple clear-cut clinical syndromes with sharp boundaries but rather fall along a clinical spectrum with substantial overlap behaviourally and pathologically. When we assessed cluster stability, we were only able to robustly assign the participants to two putative groups: one group characterized by reduced verbal fluency, agrammatism (both expressive and receptive), executive dysfunction and higher disease severity (as measured by the CDR-SB), and a second group showing the opposite pattern. Notably, the clinical profile of these two groups was strongly aligned with loadings of the PCA-derived LCD characterized by reduced verbal fluency, higher agrammatism and disease severity. This LCD was an important prognostic factor in the context of the nfvPPA clinical continuum. Critically, these findings support the view that reduced speech fluency, executive dysfunction and general functional impairment should be considered in tandem with agrammatism to stratify participants based on their expected progression rate. Such stratification may reduce the heterogeneity of participants in future trials and reduce sample size requirements.<sup>58</sup>

The association between faster disease progression and participants presenting with marked agrammatism is also relevant for interpreting previous studies. For example, in keeping with our results, it has been reported that patients with prominent expressive agrammatism in the absence of AOS (referred to as 'primary agrammatic aphasia', PAA) display faster disease progression and more

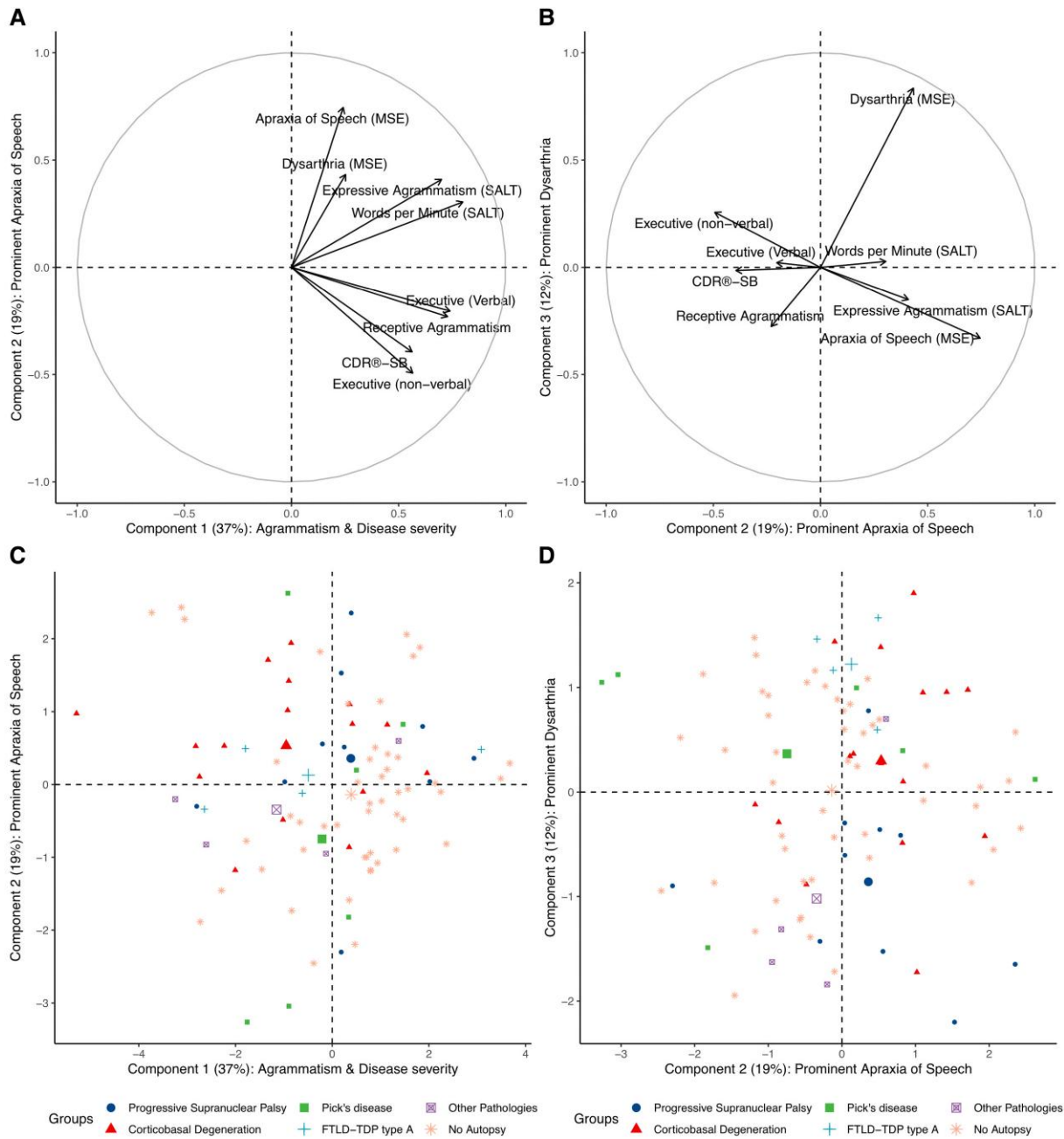


**Figure 2 Clustered heat map of all participants.** The clustered heat map illustrates the lack of robust clinical clusters within the nfvPPA-spectrum. The first unlabelled column relates to the neuropathological data of each participant [pink = no autopsy available; red = corticobasal degeneration (CBD); dark blue = progressive supranuclear palsy (PSP); green = Pick's disease; light blue = frontotemporal lobar degeneration characterized by phosphorylated 43-kDa TAR DNA-binding protein inclusions (FTLD-TDP) type A; purple = other pathologies]. Each labelled column represents its corresponding clinical feature, and each row represents a participant. The scores for all clinical features have been scaled to allow their comparison. The participants and the variables have been ordered based on similarity, as illustrated by dendrograms on the top (variables) and left (participants). CDR-SB = Clinical Dementia Rating Sum of Boxes; CYCLE = Curtiss Yamada Comprehensive Language Evaluation; DKEFS = Delis-Kaplan Executive Function Scale; MSE = motor speech evaluation; SALT = Systematic Analysis of Language Transcripts.

widespread neurodegeneration during follow-up.<sup>5</sup> Conversely, participants in our sample classified as PPAOS (AOS in the absence of expressive agrammatism) showed milder disease severity than the rest of the participants. This finding suggests that participants with relatively isolated AOS may be diagnosed at an earlier disease stage than those with relatively isolated agrammatism. This is an important observation because disease progression has been shown to accelerate with increasing disease severity in FTLD.<sup>59</sup> Thus, patients diagnosed at a more advanced disease stage are expected to deteriorate faster than those diagnosed at an earlier disease stage. Accordingly, our results support the view that patients

with significant agrammatism at diagnosis may be at a more advanced disease stage and are thus expected to deteriorate faster. Of note, validated staging tools, such as the CDR® plus NACC FTLD-SB, are more likely to yield reproducible results than the widely adopted 'estimated time from symptom onset', which is a retrospective measure influenced by recall bias, patient anosognosia and other patient- and clinician-related factors.<sup>33</sup>

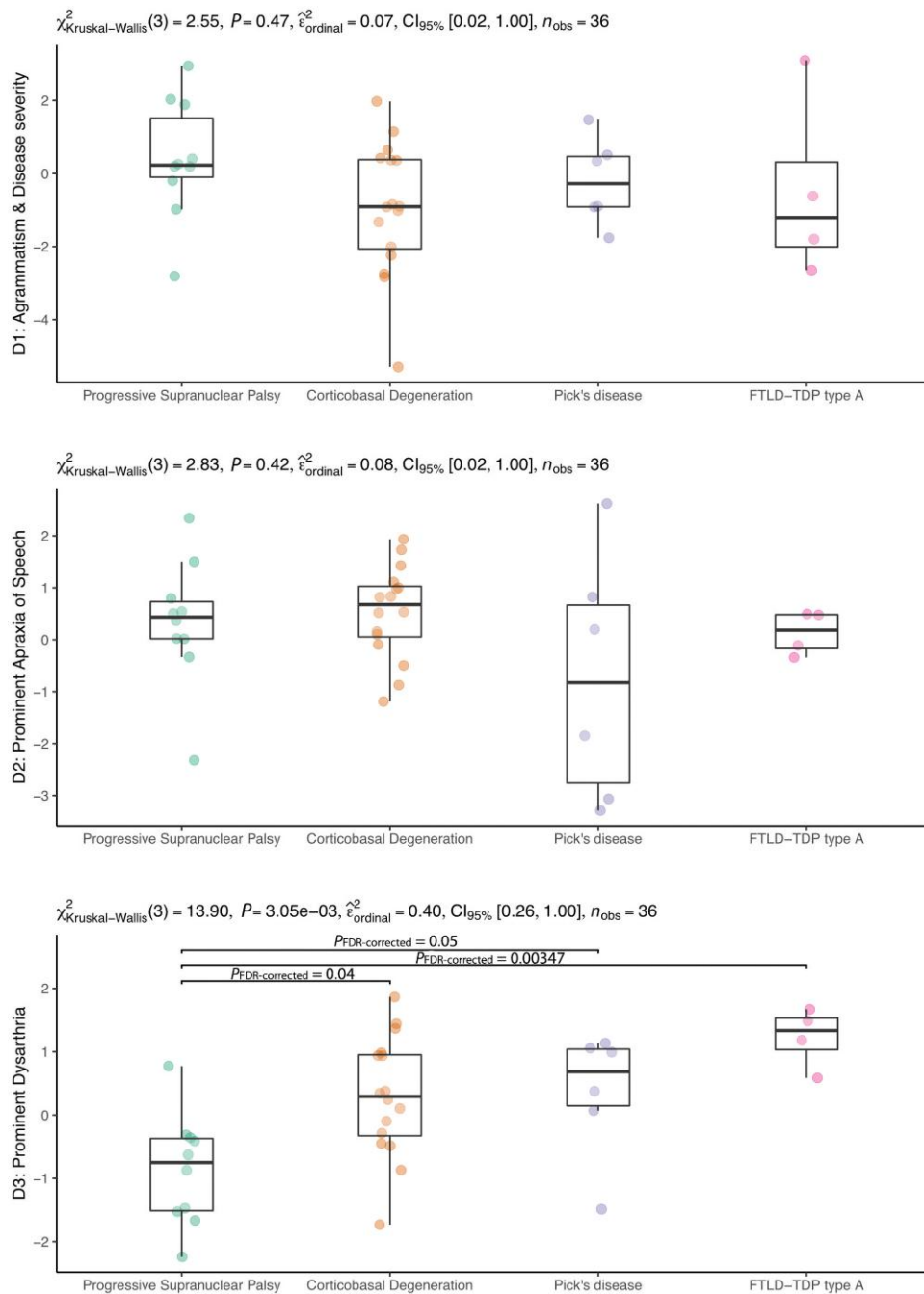
Consistent with the lack of robust clustering tendency in our large and representative sample, we showed that classifying participants into supposedly clear-cut clinical phenotypes (AOS + agrammatism, PPAOS, PAA, etc.) failed to predict underlying



**Figure 3 Characterization of latent clinical dimensions derived from principal component analysis.** In all panels, the x- and y-axes represent one of the three latent clinical dimensions (or ‘clinical components’) derived from principal component analysis (PCA). The percentage of variance explained by each component is shown in parentheses. In **A** and **B**, the value at the x- and y-axes represent the standardized coefficient (relative weight) of the components derived from PCA (higher values indicating a more substantial contribution to a given component). Each variable included in the PCA is represented by an arrow whose orthogonal projection to the x- and the y-axes indicates its standardized coefficient for the corresponding component. **(A)** The relative contribution of each variable to Component 1 (characterized by agrammatism, reduced verbal fluency and higher disease severity) and Component 2 (represented by prominent apraxia of speech and less disease severity and executive dysfunction). **(B)** The relative contribution of each variable to Component 2 and Component 3 (characterized by prominent dysarthria with less apraxia of speech). **(C and D)** The individual loadings for each component (lower individual loadings representing more impaired performance). Each participant’s neuropathological diagnosis (if available) is shown in the legend of **C** and **D**. For each neuropathological category, the centroid of the ellipse encompassing each group is represented by the biggest symbol. Notably, the distribution of individual loadings did not reveal clusters of participants but rather a widespread distribution. **(C and D)** The x- and y-axes represent the individual loading for one of the three components derived from PCA in **C** and **D**. CDR-SB = Clinical Dementia Rating sum-of-boxes; CYCLE = Curtiss Yamada Comprehensive Language Evaluation; DKEFS = Delis-Kaplan Executive Function Scale; FTLN-TDP type A = frontotemporal lobar degeneration characterized by phosphorylated 43-kDa TAR DNA-binding protein inclusions; LCD = latent clinical dimension; MSE = motor speech evaluation; SALT = Systematic Analysis of Language Transcripts.

pathology, as did the two robust data-driven subgroups derived using a k-means clustering approach. This finding prompted the investigation of data-driven latent clinical components to improve

clinical-pathological correlations and prediction of disease progression. Our data-driven approach in the nfvPPA-spectrum contrasts with subjective expert-based perceptual classifications and

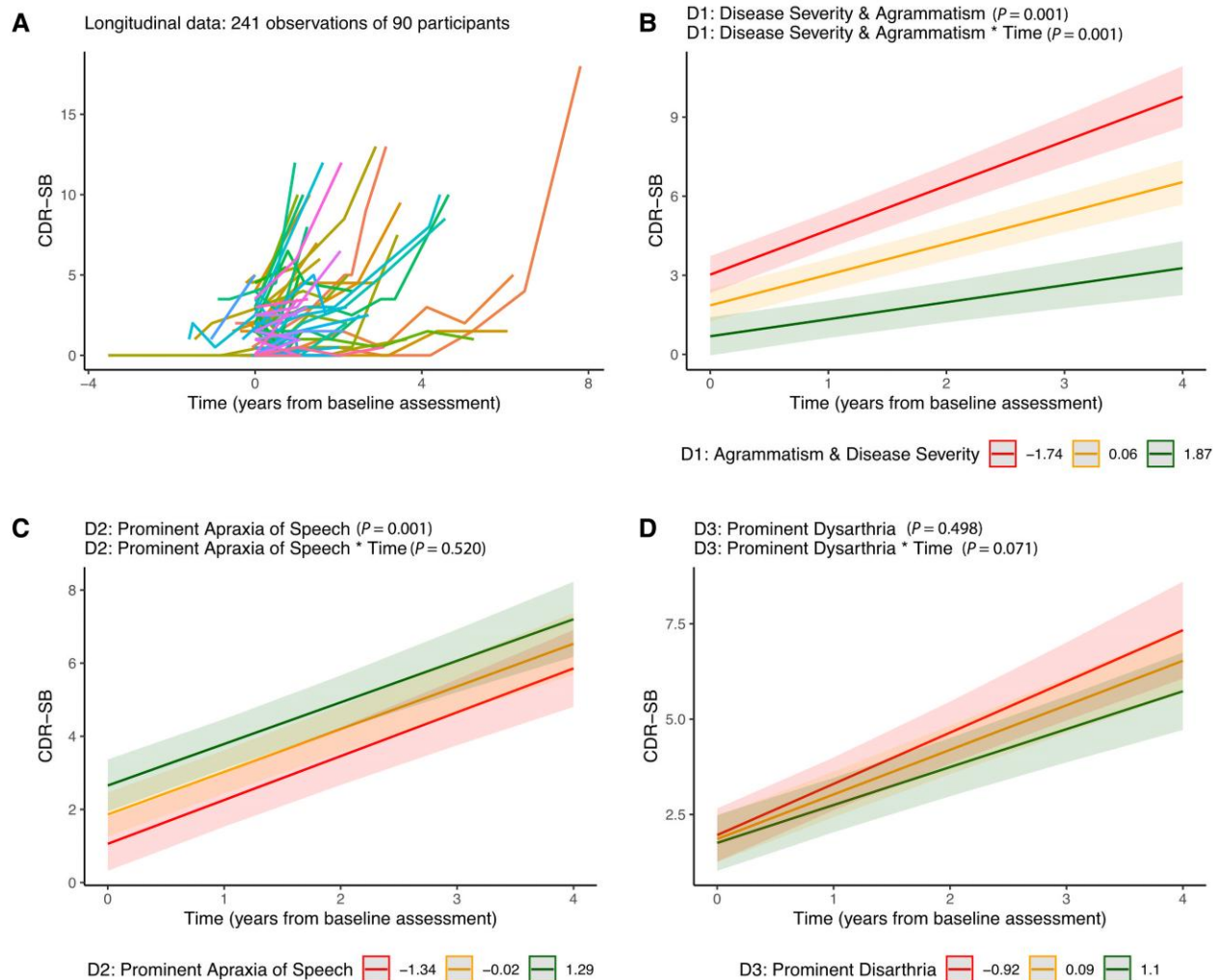


**Figure 4 Neuropathological correlates of data-driven latent clinical dimensions.** Subject-specific loading for each latent clinical component was available in 90 participants from the total sample and 36 participants with neuropathological diagnoses. D1 = latent clinical dimension 1; D2 = latent clinical dimension 2; D3 = latent clinical dimension 3.

provides valuable information to interpret previous studies.<sup>9</sup> We found that the LCD characterized by prominent dysarthria distinguished participants with underlying PSP from those with other neuropathological subtypes. This finding is in line with previous studies from our group.<sup>14</sup> But we also observed substantial overlap between clinical features, which, alone, did not allow the prediction of underlying aetiology at the single-subject level. It should be noted that segregating participants according to their clinical profile for treatment purposes may still be useful, as different patients may benefit from some speech therapy interventions depending on their

characteristic deficits.<sup>60</sup> However, most current approaches include strategies for both motor speech and grammatical deficits that are likely to co-occur during the course of the disease.

As expected, the nfvPPA-spectrum was associated with multiple neuropathological diagnoses. Consistent with previous studies, 4R tauopathies (particularly PSP) were frequently observed in participants with AOS but not agrammatism. However, both the clinical subgroups and the LCDs defined in this study did not allow the robust discrimination of neuropathological subtypes. Thus, findings of the present study discourage the use of clinical phenotypes falling within



**Figure 5** Baseline predictors of faster longitudinal decline. (A) Spaghetti plot representing longitudinal data. (B–D) CDR-SB estimates were obtained from linear mixed-effects models as a function of individual loading for Components 1, 2 and 3. For illustrative purposes, we show the CDR-SB estimates for each tercile. Error bars represent 95% confidence intervals and lower values for each latent clinical component represent more impaired performance. D1 = component 1; D2 = component 2; D3 = component 3; CDR-SB = Clinical Dementia Rating Sum of Boxes.

the nfVPPA-spectrum to predict underlying neuropathology. Moving forward, it may be more fruitful to direct efforts toward unravelling the potential of biomarkers in elucidating the underlying neuropathology and the longitudinal course of progression in patients within the nfVPPA-spectrum. Notably, future studies should assess the potential of emerging biofluid and imaging biomarkers to improve clinical-pathological correlations.<sup>61,62</sup>

Taken together, our results do not support the view that the nfVPPA-spectrum can be robustly split into clear-cut syndromic entities (i.e. PPAOS versus PAA) or that these supposedly clear-cut clinical phenotypes allow discrimination between FTLD pathological subtypes. A current challenge to the field is that FTLD-related syndromes are largely overlapping and many patients may meet diagnostic criteria for more than one syndrome.<sup>3</sup> For example, the distribution of different neuropathological entities in our cohort is very similar to that reported in a cohort of PPAOS patients (Supplementary Fig. 2), suggesting that these two diagnostic labels are largely overlapping.<sup>9</sup> In addition, participants within the nfVPPA-spectrum may also display motor symptoms or signs (mild oculomotor dysfunction or postural instability) and could also meet updated diagnostic criteria for PSP or CBD.<sup>63,64</sup>

More work is thus needed to implement multidimensional classification schemes that incorporate imaging and fluid biomarkers in an attempt to advance toward precision medicine approaches to pharmacological treatment.<sup>3,7,65</sup>

### Reinterpreting non-fluent/agrammatic variant primary progressive aphasia as a spectrum disorder

Our recent examination of brain-behaviour relationships in the nfVPPA-spectrum revealed that the neural correlates of AOS and expressive agrammatism lie next to each other in the left posterior inferior frontal lobe, explaining why these two symptoms do not always co-occur.<sup>22</sup> But, given the spatial proximity of their neural substrates, patients presenting with relatively isolated AOS or expressive agrammatism are expected to represent the exception rather than the rule. Notably, the phenotypic overlap within the nfVPPA-spectrum is only expected to increase with disease progression, thereby blurring diagnostic boundaries.

Consistent with Lorca-Puls et al.,<sup>22</sup> this study illustrates that patients within the nfVPPA-spectrum fall along a clinical spectrum with substantial overlap. However, this does not negate the

presence of phenotypic variation within nfvPPA; on the contrary, we embrace this heterogeneity while at the same time highlighting the fact that substantial phenotypic overlap also exists. To reconcile both phenomena, we propose that nfvPPA is best conceptualized as a spectrum disorder that exhibits graded distinctions but not sharp boundaries. Importantly, given that the primary symptomatology at diagnosis may still be useful to treatment approaches and inform prognosis, we suggest that a two-level diagnostic scheme is most appropriate, where the first level establishes whether the behavioural phenotype falls within the nfvPPA spectrum and then the second level records other prominent clinical or biomarker features with potential prognostic or aetiological value (e.g. general measures of disease severity, agrammatism or midbrain atrophy). This diagnostic scheme effectively conveys that multiple speech-language profiles belong to the same clinical spectrum without neglecting their graded distinctions and unnecessarily adding an extra layer of complexity by introducing other diagnostic labels.

### Limitations

This study has some limitations. For example, a definitive neuropathological diagnosis was only available in a relatively small number of participants ( $n = 43$ ). Yet, to our knowledge, our study still represents the largest series of cases with autopsy data, amounting to nearly half of the total sample. In addition, the measures we derived for expressive agrammatism (from a quantitative linguistic analysis of connected speech samples) may miss some cases with milder/subtler forms of expressive agrammatism that trained experts could detect. Our study may also have incorporated participants with neuropsychiatric or motor symptoms exceeding the customary threshold for PPA diagnosis. However, in all cases, the treating physician considered that speech and/or language deficits dominated the clinical picture. While this approach enabled the inclusion of a broader spectrum of patients, it might render the PPA phenotype in our sample ‘less pure’ (albeit more representative). Conversely, our sample was characterized by a fairly large proportion of prodromal and mild cases compared to other multicentre cohorts, suggesting that our inclusion criteria did not bias the patient sample towards participants in advanced stages of the disease.<sup>58</sup> Finally, we acknowledge that the presence or absence of AOS and dysarthria was exclusively ascertained on the basis of an expert-dependent auditory-perceptual assessment of motor speech ability. Future studies should investigate the potential of more refined quantitative measures derived from expert-independent automated speech analyses.

### Conclusion

In our cohort, splitting the nfvPPA-spectrum into different clinical phenotypes based on its major clinical features does not improve clinical-pathological correlations, stressing the need for new biological markers and consensus regarding updated terminology and clinical classification.

### Data availability

The conditions of our ethics approval do not permit the public archiving of anonymized study data. Data requests can be submitted here: <https://memory.ucsf.edu/research-trials/professional/open-science>. Following a UCSF-regulated procedure, access will be granted to designated individuals in line with ethical guidelines

on reusing sensitive data. This would require the submission of a Material Transfer Agreement. Commercial use will not be approved.

### Acknowledgements

We are indebted to the patients and their relatives for their generous assistance with our research.

### Funding

I.I.-G. is a senior Atlantic Fellow for Equity in Brain Health at the Global Brain Health Institute (GBHI), and receives funding from the GBHI, the Alzheimer’s Association, and the Alzheimer Society (GBHI ALZ UK-21-720973 and AACSF-21-850193). I.I.-G. was also supported by the Juan Rodés Contract (JR20/0018) and PI21/00791 from Instituto de Salud Carlos III. D.L.L.-P. is supported with funding from the Chilean National Agency for Research and Development (ANID SUBVENCIÓN A LA INSTALACIÓN EN LA ACADEMIA 85220006). D.L.L.-P. was also supported by a postdoctoral fellowship from the Chilean National Agency for Research and Development (ANID BECAS-CHILE 74200073). This work was supported by the National Institutes of Health (L.T.G., NIA K24 AG053435; S.S., K08AG052648; M.L.H., NIDCD R01DC016291; A.L.B., NS092089, R01AG038791, U19AG063911; H.J.R., AG045333, AG056749, AG032306, AG045390; B.L.M., NIA P50 AG023501, NIA P01 AG019724; M.L.G.-T., NINDS R01 NS050915, NIDCD K24 DC015544, NIA U01 AG052943). The UCSF Neurodegenerative Disease Brain Bank receives funding support from NIH grants P30 AG062422, P01 AG019724, U01 AG057195, and U19 AG063911, as well as the Rainwater Charitable Foundation and the Bluefield Project to Cure FTD.

### Competing interests

I.I.-G. reported receiving personal fees from Nutricia, Esteve, UCB, and Neuraxpharm Spain outside the submitted work. J.d.L. reported receiving grants from the Alzheimer’s Association outside the submitted work. L.T.G. reported receiving grants from the National Institutes of Health (NIH), Rainwater Charitable Foundation, and Weill Neuroscience Institute outside the submitted work. S.S. reports consulting fees from Techspert.io, Precision Xtract, and Acel Health outside the submitted work. A.L.B. reported receiving grants from NIH and grants from Rainwater Charitable Foundation during the conduct of the study; receiving grants from NIH, the Tau Research Consortium, the Association for Frontotemporal Degeneration, Bluefield Project to Cure Frontotemporal Dementia, Corticobasal Degeneration Solutions, the Alzheimers Drug Discovery Foundation and the Alzheimer’s Association; consulting for Applied Genetic Therapies, Alector, Arkuda, Arvinas, AZTherapeutics, Boehringer Ingelheim, GlaxoSmithKline, Humana, Lundbeck, Oligomerix, Ono, Roche, Samumed, Stealth, Third Rock, Transposon, TrueBinding, and Wave; receiving stock and options from Alector Stock/options; receiving options from Arvinas, Arkuda, AZTherapeutics, and True Binding; receiving grants from Biogen, Eisai, and Regeneron; and receiving personal fees from Denali, GlaxoSmithKline, Humana, Boehringer Ingelheim, Oscotec, Oligomerix, Roche, Transposon, and Wave outside the submitted work. No other disclosures were reported. B.L.M. reported receiving grants from NIH and receiving royalties from Cambridge University Press, Guilford Publications, Johns Hopkins Press, Oxford University Press, Taylor & Francis Group, Elsevier, and UpToDate outside the submitted work. W.W.S. reported

receiving personal fees from BridgeBio, GLG Council, Guidepoint Global, and Corcept Therapeutics outside the submitted work.

The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## Supplementary material

Supplementary material is available at *Brain* online.

## References

1. Tee BL, Gorno-Tempini ML. Primary progressive aphasia: A model for neurodegenerative disease. *Curr Opin Neurol*. 2019; 32:255-265.
2. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;76:1006-1014.
3. Murley AG, Coyle-Gilchrist I, Rouse MA, et al. Redefining the multidimensional clinical phenotypes of frontotemporal lobar degeneration syndromes. *Brain*. 2020;143:1555-1571.
4. Josephs KA, Duffy JR, Strand EA, et al. The evolution of primary progressive apraxia of speech. *Brain*. 2014;137:2783-2795.
5. Tetzloff KA, Duffy JR, Clark HM, et al. Progressive agrammatic aphasia without apraxia of speech as a distinct syndrome. *Brain*. 2019;142:2466-2482.
6. Josephs KA, Duffy JR, Strand EA, et al. Characterizing a neurodegenerative syndrome: Primary progressive apraxia of speech. *Brain*. 2012;135:1522-1536.
7. Boeve BF, Boxer AL, Kumfor F, Pijnenburg Y, Rohrer JD. Advances and controversies in frontotemporal dementia: Diagnosis, biomarkers, and therapeutic considerations. *Lancet Neurol*. 2022;21:258-272.
8. Thompson CK, Cho S, Price C, et al. Semantic interference during object naming in agrammatic and logopenic primary progressive aphasia (PPA). *Brain Lang*. 2012;120:237-250.
9. Josephs KA, Duffy JR, Clark HM, et al. A molecular pathology, neurobiology, biochemical, genetic and neuroimaging study of progressive apraxia of speech. *Nat Commun*. 2021;12:3452.
10. Henry ML, Hubbard HI, Grasso SM, et al. Retraining speech production and fluency in non-fluent/agrammatic primary progressive aphasia. *Brain*. 2018;141:1799-1814.
11. Spinelli EG, Mandelli ML, Miller ZA, et al. Typical and atypical pathology in primary progressive aphasia variants. *Ann Neurol*. 2017;81:430-443.
12. Whitwell JL, Tosakulwong N, Schwarz CC, et al. Longitudinal anatomic, functional, and molecular characterization of pick disease phenotypes. *Neurology*. 2020;95:e3190-e3202.
13. Caso F, Mandelli ML, Henry M, et al. In vivo signatures of nonfluent/agrammatic primary progressive aphasia caused by FTLD pathology. *Neurology*. 2014;82:239-247.
14. Santos-Santos MA, Mandelli ML, Binney RJ, et al. Features of patients with nonfluent/agrammatic primary progressive aphasia with underlying progressive supranuclear palsy pathology or corticobasal degeneration. *JAMA Neurol*. 2016;73:733-710.
15. Rogalski E, Sridhar J, Rader B, et al. Aphasic variant of Alzheimer disease: Clinical, anatomic, and genetic features. *Neurology*. 2016;87:1337-1343.
16. Josephs KA. Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*. 2006;129:1385-1398.
17. Boxer AL, Gold M, Feldman H, et al. New directions in clinical trials for frontotemporal lobar degeneration: Methods and outcome measures. *Alzheimers Dement*. 2020;16:131-143.
18. Strafella C, Caputo V, Galota MR, et al. Application of precision medicine in neurodegenerative diseases. *Front Neurol*. 2018;9: 701.
19. Matias-Guiu JA, Díaz-Álvarez J, Cuetos F, et al. Machine learning in the clinical and language characterisation of primary progressive aphasia variants. *Cortex*. 2019;119:312-323.
20. Hoffman P, Sajjadi SA, Patterson K, Nestor PJ. Data-driven classification of patients with primary progressive aphasia. *Brain Lang*. 2017;174:86-93.
21. Mesulam MM. Primary progressive aphasia. *Ann Neurol*. 2001;49: 425-432.
22. Lorca-Puls DL, Gajardo-Vidal A, Mandelli ML, et al. Neural basis of speech and grammar symptoms in non-fluent variant primary progressive aphasia spectrum. *Brain*. 2024;147(2): 607-626.
23. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. *J Psychiatr Res*. 1975;12:189-198.
24. Morris JC. The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*. 1993;43:2412-2414.
25. Litvan I, Agid Y, Calne D, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): Report of the NINDS-SPSP international workshop. *Neurology*. 1996;47:1-9.
26. Boxer AL, Geschwind MD, Belfor N, et al. Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Arch Neurol*. 2006;63:81.
27. Perry DC, Brown JA, Possin KL, et al. Clinicopathological correlations in behavioural variant frontotemporal dementia. *Brain*. 2017;140:3329-3345.
28. Seo SW, Thibodeau MP, Perry DC, et al. Early vs late age at onset frontotemporal dementia and frontotemporal lobar degeneration. *Neurology*. 2018;90:e1047-e1056.
29. Mackenzie IRA, Neumann M, Bigio EH, et al. Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: An update. *Acta Neuropathol (Berl)*. 2010;119:1-4.
30. Gorno-Tempini ML, Brambati SM, Ginex V, et al. The logopenic/phonological variant of primary progressive aphasia. *Neurology*. 2008;71:1227-1234.
31. Wilson SM, DeMarco AT, Henry ML, et al. Variable disruption of a syntactic processing network in primary progressive aphasia. *Brain*. 2016;139:2994-3006.
32. Kramer JH, Jurik J, Sha SJ, et al. Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cogn Behav Neurol*. 2003;16:211-218.
33. Miyagawa T, Brushaber D, Syrjanen J, et al. Utility of the global CDR® plus NACC FTLT rating and development of scoring rules: Data from the ARTFL/LEFFTDS consortium. *Alzheimers Dement*. 2020;16:106-117.
34. Wertz RT, LaPointe LL, Rosenbek JC. *Apraxia of speech in adults: The disorder and its management*. Grune and Stratton; 1984.
35. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6:284-290.
36. Kertesz A. *The western aphasia battery*. Grune and Stratton; 1982.
37. Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. *Cognition*. 2004;92(1-2):145-177.
38. Wilson SM, Dronkers NF, Ogar JM, et al. Neural correlates of syntactic processing in the nonfluent variant of primary progressive aphasia. *J Neurosci*. 2010;30:16845-16854.
39. Yi A. California Verbal learning test (California verbal learning test-II). In: Kreutzer JS, DeLuca J, Caplan B, eds. *Encyclopedia of clinical neuropsychology*. Springer; 2011:475-476.

40. Possin KL, Laluz VR, Alcantar OZ, Miller BL, Kramer JH. Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia*. 2011;49:43-48.
41. Golden CJ, Freshwater SM, Zarabeth G, University NS, eds. *Stroop color and word test children's version for ages 5–14: A manual for clinical and experimental uses*. Stoelting; 2003.
42. Delis DC, Kaplan E, Kramer JH, eds. *Delis kaplan executive function system (D-KEFS)*. Psychological Corporation; 2001.
43. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept Mot Skills*. 1958;8:271-276.
44. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston naming test: Shortened versions for use in Alzheimer's disease. *J Gerontol*. 1992;47:P154-P158.
45. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*. 1994;44:2308-2308.
46. Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. *Brain*. 2007;130(Pt 6):1552-1565.
47. Goetz CG, Tilley BC, Shaftman SR, et al. Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord*. 2008;23:2129-2170.
48. Crawford JR, Garthwaite PH, Ryan K. Comparing a single case to a control sample: Testing for neuropsychological deficits and dissociations in the presence of covariates. *Cortex*. 2011;47:1166-1178.
49. Ranasinghe KG, Rankin KP, Lobach IV, et al. Cognition and neuropsychiatry in behavioral variant frontotemporal dementia by disease stage. *Neurology*. 2016;86:600-610.
50. Gómez-Andrés D, Oulhissane A, Quijano-Roy S. Two decades of advances in muscle imaging in children: From pattern recognition of muscle diseases to quantification and machine learning approaches. *Neuromuscul Disord*. 2021;31:1038-1050.
51. Lawson RG, Jurs PC. New index for clustering tendency and its application to chemical problems. *J Chem Inf Comput Sci*. 1990;30:36-41.
52. fpc package - RDocumentation. Accessed 18 September 2023. <https://www.rdocumentation.org/packages/fpc/versions/2.2-10>
53. Bartlett MS. THE EFFECT OF STANDARDIZATION ON A  $\chi^2$  APPROXIMATION IN FACTOR ANALYSIS. *Biometrika*. 1951;38(3-4):337-344.
54. Kaiser HF, Rice J. Little Jiffy, Mark IV. *Educ Psychol Meas*. 1974;34:111-117.
55. Field AP, Miles J, Field Z. *Discovering statistics using R*. Sage; 2012.
56. Staffaroni AM, Ljubenkova PA, Kornak J, et al. Longitudinal multimodal imaging and clinical endpoints for frontotemporal dementia clinical trials. *Brain*. 2019;142:443-459.
57. Illán-Gala I, Falgàs N, Friedberg A, et al. Diagnostic utility of measuring cerebral atrophy in the behavioral variant of frontotemporal dementia and association with clinical deterioration. *JAMA Netw Open*. 2021;4:e211290.
58. Staffaroni AM, Weintraub S, Rascovsky K, et al. Uniform data set language measures for bvFTD and PPA diagnosis and monitoring. *Alzheimers Dement (Amst)*. 2021;13:e12148.
59. Staffaroni AM, Goh SYM, Cobigo Y, et al. Rates of brain atrophy across disease stages in familial frontotemporal dementia associated with MAPT, GRN, and C9orf72 pathogenic variants. *JAMA Netw Open*. 2020;3:e2022847.
60. Robinaugh G, Henry ML. Behavioral interventions for primary progressive aphasia. *Handb Clin Neurol*. 2022;185:221-240.
61. Illán-Gala I, Nigro S, VandeVrede L, et al. Diagnostic accuracy of magnetic resonance imaging measures of brain atrophy across the Spectrum of progressive supranuclear palsy and corticobasal degeneration. *JAMA Netw Open*. 2022;5:e229588.
62. Horie K, Barthélemy NR, Spina S, et al. CSF tau microtubule-binding region identifies pathological changes in primary tauopathies. *Nat Med*. 2022;28:2547-2554.
63. Höglinger GU, Respondek G, Stamelou M, et al. Clinical diagnosis of progressive supranuclear palsy: The movement disorder society criteria. *Mov Disord*. 2017;32:853-864.
64. Armstrong MJ, Litvan I, Lang AE, et al. Criteria for the diagnosis of corticobasal degeneration. *Neurology*. 2013;80:496-503.
65. VandeVrede L, Ljubenkova PA, Rojas JC, Welch AE, Boxer AL. Four-Repeat tauopathies: Current management and future treatments. *Neurotherapeutics*. 2020;17:1563-1581.