Validated Automatic Speech Biomarkers in Primary Progressive Aphasia

Journal:	*Annals of Neurology*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Nevler, Naomi; University of Pennsylvania Perelman School of Medicine, Neurology Ash, Sharon; University of Pennsylvania Perelman School of Medicine, Neurology Irwin, David J.; University of Pennsylvania Perelman School of Medicine, Neurology Liberman, Mark; University of Pennsylvania, Linguistics Grossman, Murray; University of Pennsylvania Perelman School of Medicine, Neurology
Keywords:	Frontotemporal dementia, Aphasia, Speech analysis
Domain:	Neuroscience

20 March 2018

Validated Automatic Speech Biomarkers in Primary Progressive Aphasia

Naomi Nevler, MD[1], Sharon Ash, PhD[1], David J Irwin, MD[1], Mark Liberman, PhD[2], Murray Grossman, MD, EdD[1]

1. Penn Frontotemporal Degeneration Center, Department of Neurology, University of Pennsylvania, Philadelphia, PA
2. Linguistic Data Consortium, Department of Linguistics, University of Pennsylvania, Philadelphia, PA

Please address correspondence to:

Dr. Naomi Nevler or Dr. Murray Grossman

Department of Neurology – 3 Gates

Hospital of the University of Pennsylvania

3400 Spruce St

Philadelphia, PA

19104-4283

email: mgrossma@pennmedicine.upenn.edu or naomine@pennmedicine.upenn.edu;

voice: 215-662-3361; fax: 215-349-8464

Character counts: Title – 68; Running head – 47

Word counts: Abstract - 249; Introduction – 248; Discussion – 1,437; Body of manuscript – 3,630

Figures – 5; Tables - 3

Abstract

Objective

To automatically extract and quantify specific disease biomarkers of prosody from the acoustic properties of speech in patients with primary progressive aphasia.

Methods

We analyzed speech samples from 59 progressive aphasic patients (non-fluent/agrammatic=15, semantic=21, logopenic=23; ages 50–85 years, 39% males) and 31 matched healthy controls (ages 54–89 years, 36% males). Using a novel, automated speech analysis protocol, we extracted acoustic measurements of prosody, including fundamental frequency and speech and silent pause durations, and compared these between groups. We then examined their relationships with clinical tests, gray matter atrophy, and cerebrospinal fluid analytes.

Results

We found a narrowed range of fundamental frequency in patients with non-fluent/agrammatic variant aphasia (mean 3.86±1.15 semitones) compared with healthy controls (6.06±1.95 semitones; $p<0.001$) and patients with semantic variant aphasia (6.12±1.77 semitones; $p=0.001$). Mean pause rate was significantly increased in the non-fluent/agrammatic group (mean 61.4±20.8 pauses per minute) and the logopenic group (58.7±16.4 pauses per minute) compared to controls. Narrowed fundamental frequency range was associated with atrophy in the left inferior frontal cortex. Cerebrospinal level of phosphorylated-tau was associated with an acoustic classifier combining fundamental frequency range and pause rate ($r=0.58$, $p=0.007$). Receiver Operating Characteristic analysis with this combined classifier distinguished non-fluent/agrammatic speakers from healthy controls (AUC=0.94) and from semantic variant patients (AUC=0.86).

Interpretation

Restricted fundamental frequency range and increased pause rate are characteristic markers of dysprosodic speech in non-fluent/agrammatic primary progressive aphasia. They can be extracted with automated speech analysis and are associated with left inferior frontal atrophy and cerebrospinal phosphorylated-tau level.

Introduction

Conversational speech is essential to our daily lives and allows us to vocalize thoughts and emotions in order to communicate a message to a listener. While language is often studied by analyses of segmental content such as words and sentences, speech involves the additional component of prosody. Prosody refers to suprasegmental aspects of speech, encompassing intonation, rhythm and stress properties that are crucial for conveying linguistic and emotional information.

Despite our natural sensitivity to prosodic features of speech, studies of its pathological form, dysprosody, are rare. This may stem from difficulties quantifying features of prosody in an objective manner. Most research on prosody has relied on subjective assessments, often focusing on the expression or comprehension of emotional speech[1-5]. We developed an automated technique for speech analysis, based on a Speech Activity Detector (SAD)[6], which we implemented to examine the prosodic characteristics of a semi-structured speech sample in patients with variants of primary progressive aphasia (PPA). We aimed to investigate the behavioral and neurobiologic basis for dysprosody in these patients, while testing the implementation of our automated speech analysis method. We hypothesized distinct acoustic dysprosodic markers in variants of PPA, in particular, restriction in the pitch range and increase in pause rate in the non-fluent/agrammatic variant of PPA (naPPA). Additionally, we expected to relate these changes to specific biologic markers of pathology frequently associated with naPPA, including inferior frontal atrophy and a cerebrospinal fluid (CSF) surrogate of Frontotemporal Lobar Degeneration (FTLD) pathology involving the accumulation of misfolded tau (FTLD-tau).

Methods

Subjects

We examined digitized speech samples from 67 native English speakers who met formal clinical criteria for a specific PPA syndrome[7], including naPPA (n=18), semantic variant PPA (svPPA, n=23), logopenic variant PPA (lvPPA, n=26) and 37 healthy controls (HC). All patients were assessed between April 1998 and September 2017 by experienced

neurologists (MG, DJI) in the Department of Neurology at the Hospital of the University of Pennsylvania, and were reviewed by a consensus conference according to published criteria[7], modified for lvPPA[8]. For this study, we excluded patients with a concurrent motor disorder such as progressive supranuclear palsy (PSP), corticobasal syndrome (CBS), or amyotrophic lateral sclerosis (ALS), to minimize potential motor confounds in our acoustic analyses. Fifteen of the svPPA cases had concomitant behavioral symptoms, but their speech acoustic pattern did not differ from their counterparts with isolated svPPA. We reviewed all speech samples with a pitch range above or below 1.5 SD of their group mean, and detected 7 patients (3 naPPA, 2 lvPPA, 2 svPPA) and 6 controls with extensive vocal-fry or “creaky voice". These vocal characteristics carry a high probability for pitch-tracking errors and so we excluded these 13 recordings from further analysis. Another lvPPA recording was excluded due to participation in an AD disease-modifying treatment trial. The final groups, totaling 59 PPAs and 31 HCs, were matched in all demographic characteristics except disease duration, which was shorter in naPPA compared to the other PPA groups (Table 1). Additional neuropsychological test data and manually coded linguistic data[9] are presented in Table 1 to confirm typical characteristics of each patient group.

Speech Samples

We used the Cookie Theft picture description task from the Boston Diagnostic Aphasia Examination[10] to elicit semi-structured narrative speech samples[9]. Total speech time of each recording averaged 41.1 seconds (range 4.4 – 76.4 seconds), excluding silent pauses and interviewer’s speech. Details of digital speech collection have been previously published[11]. Characteristics of speech reported previously[9, 12] in these phenotypes include speech rate measured as words per minute (WPM), and grammatical complexity reflected in dependent clauses per utterance (DC), mean length of utterance (MLU) and well-formed sentences per utterance (WFS).

Sound Processing

We used a SAD developed at the University of Pennsylvania Linguistic Data Consortium (LDC)[6] to time-segment the audio files and then pitch-tracked the segments of continuous

speech, using a protocol described previously[11]. We extracted the fundamental frequency (f0, defined as the inverse of the longest repeated waveform in a complex periodic signal, and closely related to perceived pitch intonation[13]) for each continuous speech segment, as well as the durations of speech and silent pause segements. From these, we calculated the following measures: f0 range, which is represented by the 90th percentile f0[11], mean speech segment and pause segment durations, and pause rate, which was calculated as the number of pauses per minute (ppm) over the total speech time for each subject. We validated our automatic measurements by comparing its results to a blinded assessment of restricted versus normal f0 range performed by experienced human raters (NN and SA). Inter-rater agreement was substantial (Cohen's kappa=0.81) and the cases of disagreement were reviewed and discussed until an agreement was reached. We compared these judgments to PPA subgroups formed by using a cutoff for normal f0 range at 4.8 semitones (ST), based on an ROC analysis for all PPA patients versus controls. A chi-square test showed no difference in the distributions of the normal and restricted f0 range categories when using the automated analysis compared to the subjective evaluation (X^2=1.48, df=1, p=0.22).

Analysis of Likely Pathology

Thirty-seven of our patients had a CSF sample collected within 5–39 months (mean 10.6) of cookie-theft speech recording. Following a pathologically validated algorithm, we screened for a non-Alzheimer's disease (AD) CSF profile (p-Tau/Aß<0.09, available in 32 samples)[14, 15]. This procedure identified 20 cases with a CSF-profile suggestive of non-AD FTLD underlying pathology[14]. These included 2 autopsy-confirmed cases (1 Tau, 1 TDP) and a third case with confirmed *MAPT* mutation.

To determine association of automated speech features with *in vivo* measures of pathology we examined the relationship between our acoustic variables and CSF biomarkers including beta-Amyloid (Abeta), total (tTau) and phosphorylated Tau (p-Tau) in this subset of high-probability FTLD pathology patients. We tested the effect of a combined acoustic parameter (see below) on each of these CSF analytes, applying multivariate regression analysis techniques (see below).

Statistical analysis

Demographic data were compared with analysis of variance (ANOVA) for continuous variables and the chi-square test for categorical variables. We used Kernel-density and Q-Q plots to examine speech and cognitive variable distributions and since these were normally distributed, we used ANOVA for between-group comparisons, co-varying for disease duration, and post-hoc tests with Tukey's Honest Significant Difference (HSD). Groups were compared for their f0 range, speech segment duration, and pause rate. Because of the effect of sex on f0, an additional f0 analysis was conducted within male and female subpopulations co-varying for disease duration. MMSE total scores differed between our male patient groups, and so we also introduced MMSE as a covariate in their analysis.

Within the naPPA group we compared patients with motor symptoms such as dysarthria or Apraxia of Speech (AoS, see below) to those without these speech features using a Student's t-test. Simple correlations were performed with Pearson's method. Regression analyses included generalized linear models (GLMs) with log transformation for p-Tau levels as the outcome measure and a polynomial logistic regression for clinical phenotype as the outcome variable. GLM validation was based on residuals plots. A stepwise backward elimination approach was implemented in the p-Tau GLM in order to examine the effects of potential confounders (see results section) and find the best fit model.

We performed receiver operating characteristic (ROC) curve analyses on f0 range and pause rate as acoustic classifiers for PPA phenotypes. These were tested individually and in combination (pause-rate/f0-range ratio, to control for opposite directionality) for patients versus controls and between patient groups. We used a bootstrap technique with 2000 permutations to compare ROC models of similar group-pairs. All calculations were conducted in RStudio[16] with additional packages[17-25].

Gray Matter (GM) Density Analysis

High-resolution structural brain MRIs were obtained on average within 1.5±2.5 months of recording in 16 controls and 9 naPPA patients. The reasons for unavailability of an MRI scan included various contra-indications for the test, absence of a T1 sequence or a difficulty obtaining a good-quality scan. Clinical and demographic characteristics of

these MRI subsets matched those of their original full sets. Details of data acquisition and pre-processing have been reported previously[11]. We calculated GM density. naPPA atrophy compared to HC was then mapped using voxel-wise comparisons in FSL[26, 27] with family-wise error correction and threshold-free cluster enhancement at a statistical threshold of $p \leq 0.01$ and cluster-size threshold of $k \geq 50$ voxels. We then performed a regression analysis within the naPPA areas of cortical atrophy, co-varying for age and disease duration. We applied 10,000 permutations equivalent to statistical protection controlling for type I error and set a statistical threshold of $p \leq 0.05$ and cluster size threshold of $k \geq 10$ voxels.

Ethical considerations

All participants were enrolled in study protocols and participated in an informed consent procedure approved by the Institutional Review Board of the University of Pennsylvania. All personnel exposed to personal patient data, including voice samples, have been specifically trained in ethical handling of patient data.

Results

Speech parameter results

We found a significantly reduced f0 range in naPPA (mean 3.86±1.15 ST) compared with HC (mean 6.06±1.95 ST; $p<0.001$) and svPPA (mean 6.12±1.77 ST; $p=0.001$, Fig 1A). Correlation analyses revealed that the f0 range is correlated with speech rate in all PPA patients ($r=0.29$, $p=0.02$). The narrow f0 range in naPPA did not correlate with any demographic or neuropsychological features (all p-values >0.1).

Pause rate differed significantly between groups (Fig 1B): each PPA group differed from HC (mean 32.24±9.75 ppm; $p \leq 0.002$ per contrast). naPPA (mean 61.36±20.8 ppm) differed from svPPA (mean 47.15±14.34 ppm; $p=0.02$), and lvPPA (58.74±16.41 ppm) also differed from svPPA ($p=0.04$). Pause rate did not correlate with disease duration, MMSE, or executive functioning. We found significant correlations between pause rate and manually coded measures of fluency and grammaticality in all PPA patients (Fig 2A to D), including: speech rate ($r=0.29$, $p=0.02$), WFS per utterance ($r=-0.47$, $p<0.001$), DC

per utterance (r=-0.48, p<0.001) and MLU (r=-0.52, p<0.001). We found a strong negative correlation between speech segment duration and pause rate across all patients (r=-0.87, p<0.001). Concordantly, speech segment duration correlated with speech rate, WFS per utterance, DC per utterance, and MLU (Fig 2E to H).

Mean speech segment duration was reduced significantly in each patient group compared to HC (p<0.001 for each contrast), but there were no significant differences between patient groups (Fig 1C). Speech segment duration did not correlate with demographic or cognitive measures.

Mean pause duration (overall mean 1.14±0.7 sec) was similar in all groups (Fig 1D).

Correlation analyses revealed that f0 range is not associated with pause rate, pause duration, or speech segment duration in any group (all p-values>0.1).

A regression model of f0 range and pause rate as main predictors of clinical phenotype indicated that a reduction of 1 semitone in f0 range with a constant pause rate would result in a 2.9-fold increase in the likelihood of a diagnosis of naPPA compared to HC (Table 2).

Within naPPA we compared patients with dysarthria or AoS (n=7) to those without these features at the time of recording (n=8). We found no differences in any acoustic marker between these subgroups. Likewise, an analysis by sex revealed a comparable restriction of f0 range in the naPPA group within each gender (males: 3.19±1.35 ST in naPPA vs. 5.96±1.49 ST in HC, p=0.002 and vs. 4.98±1.18 ST in svPPA, p=0.06; females: 4.31±0.79 ST in naPPA vs. 6.11±2.2 ST in HC, p=0.06 and vs. 7.15±1.58 ST in svPPA, p=0.004).

Finally, in ROC curve analyses, f0 range as a single predictor of naPPA versus HC had an area under the curve (AUC) of 0.84 (95% CI: 0.72-0.96) and best threshold at 4.8 ST, while pause rate showed an AUC=0.89 (95% CI: 0.75-1.00) and best threshold at 52.3

ppm, with no statistically significant difference between these two curves (p=0.6, Fig 3A). A combined acoustic parameter showed an AUC=0.94 (95% CI: 0.87-1.00 at best threshold; sensitivity 87%, specificity 90%) distinguishing naPPA from HC (Fig 3A-B). The same classifier distinguished naPPA from svPPA at an AUC=0.86 (95% CI: 0.73-0.98 at best threshold; sensitivity 71%, specificity 87%), and distinguished naPPA from lvPPA with an AUC=0.69 (95% CI: 0.50-0.87 at best threshold; sensitivity 87%, specificity 47%) (Fig 3B).

Neuroimaging

The naPPA group showed bilateral frontotemporal atrophy, most prominently left frontal atrophy. We associated f0 range with GM atrophy in the left inferior frontal gyrus (IFG). Refer to Table 3 and Fig 4 for details and visualization.

CSF results

We previously found CSF p-tau levels to correlate with the severity of postmortem tau pathology in FTLD (Irwin et al AON 2017). To determine if our automatically extracted speech variables relate to an *in vivo* marker of tau pathology we examined the relationship between these speech features and CSF biomarkers in the subset of PPA patients with a CSF profile suggestive of FTLD pathology (n=20).

We found the natural logarithm of p-Tau levels was linearly associated with the natural logarithm of the combined acoustic parameter (r=0.58, p=0.007, Fig 5). We tested the effect of potential confounding variables including age, disease duration, education, and the time interval between speech sample and CSF collection. These were found to have no significant effect (simple correlation was the best fit). We did not find an association of our prosodic marker with CSF biomarkers that do not directly associate with postmortem tau pathology (i.e. tTau and Abeta, data not shown).

Discussion

Our automated speech analysis protocol identified two basic acoustic markers that characterize patients with naPPA in a sensitive and specific way: f0 range, which correlates with perceived pitch; and pause rate, which is a measure of dysfluency. These

were associated with left inferior frontal atrophy and CSF level of p-Tau. Speech analyses in naPPA thus may be an informative marker to screen for FTLD-Tau pathology.

Prosody is a distinct but integral element of spoken language, associated with neural networks supporting language[28-30]. Although it plays a role in the phonological representation of some individual words in the auditory-aural system (e.g. stress provides the linguistic differentiation of "rècord" as a noun versus "recòrd" as a verb), prosody mainly contributes to suprasegmental aspects of sentence processing. For example, prosodic features mark the end of an utterance, distinguishing between a question and a statement (e.g. declining pitch for "you're tired! Versus ascending pitch for "you're tired?"). Several aspects of sentence-level speech are disturbed in naPPA, including grammatical expression[9, 12]. A restriction in f0 range may further limit naPPA patients' ability to express themselves at the sentence level. Indeed, our findings associate limited prosodic processing with difficulty expressing grammatically well-formed sentences. Although naPPA patients use shorter sentences, this does not apparently limit their ability to express pitch range in their speech since there was no correlation between f0 and mean length of utterance. Impaired prosodic comprehension has also been reported in naPPA[31]. Future work can determine whether dysprosody relates to impaired grammatical comprehension[32].

Our finding of highest pause rate in the naPPA group coincides with previous speech analyses[33, 34]. We correlated pause rate with a manually coded measure of reduced speech rate (words/minute) that is associated with the characteristic effortfulness heard in naPPA speech. This validates the use of our automated algorithm, which does not depend on time-consuming generation of transcripts and makes use of natural breath-group boundaries (see below). While pause rate is increased in naPPA, pause duration and speech segment duration do not differ between groups. Thus, these duration measures cannot easily explain the impression of effortful, non-fluent speech in naPPA.

In our cohort, 7 (~47%) patients with naPPA had either dysarthria or AoS as part of their clinical presentation. AoS has been reported to affect some acoustic measures of patients' speech, specifically prolonged duration of stressed syllables[35]. This was unlikely to be a confound in our study. Specifically, speech duration was measured over entire breath-groups (see below), not within words, and f0 originates at the level of the vocal folds and is mainly a function of sub-glottal air pressure[13]. Thus, impaired articulation due to difficulty coordinating the motor speech apparatus should minimally affect f0[13, 36]. f0 can also be affected by tension of the vocal folds[13]. Possible involvement of the vocal folds may be manifested as a coarse or hypophonic voice. Such voice quality is also highly susceptible to pitch-tracking errors, and, as mentioned above, we excluded samples that were detected as pitch outliers with these voice characteristics.

naPPA was characterized by reduced f0 range and increased pause rate. There was no co-linearity between these two acoustic variables, even though f0 range correlated with speech rate in the full set of PPA subjects. This suggests that the two acoustic parameters are relatively independent characteristics of the non-fluent variant of PPA. Thus, we sought to investigate whether these together can distinguish naPPA from healthy controls and from other PPA phenotypes. In addition to robustly discriminating naPPA from healthy controls, we found that the combination of these acoustic features can reasonably distinguish naPPA from svPPA, suggesting that these speech deficits are not non-specific impairments found in any aphasic but instead may be specific to a particular PPA phenotype. Others have used lengthy neuropsychological measures to distinguish between naPPA and svPPA[37, 38]. These observations are consistent with the clinical impression that svPPA is associated with a relatively intact suprasegmental speech pattern and emphasize that the disorder in svPPA is most prominently at the level of the representation of single word and object meaning[7, 39, 40].

We found previously that gender impacts f0 range in bvFTD[11]. However, we found that gender had no effect in PPA. This suggests that pitch range may interact with gender selectively in bvFTD as a component of their social disorder. While this observation emphasizes the importance of assessing f0 in both males and females with

neurodegenerative disorders, it does not appear that sex per se has a significant effect on prosody in PPA.

The acoustic characteristics of lvPPA were intermediate between those of naPPA and svPPA. The lvPPA phenotype has proven to be the least amenable to clinical identification[38, 41], although recent studies have begun to characterize it more reliably[8, 42]. It is possible that all lvPPA patients have an attenuated version of the speech disorder found in naPPA. Alternatively, a subset of patients with lvPPA may exhibit some of the speech characteristics of naPPA[38]. Additional work is needed to test these hypotheses in a larger cohort.

We associated f0 range impairment in naPPA with atrophy in left inferior frontal cortex. This observation coincides with our previous analysis in bvFTD[11], in which bilateral IFG involvement was established with a similar acoustic analysis. In the current study, we show the prosodic impairment in a group with non-fluent speech and grammatical deficits but no apparent social-behavioral impairment. Thus, association of f0 range restriction with the left IFG appears to be most consistent with the hypothesis that this acoustic marker may reflect derangement of a system of linguistic expression. This hypothesis is supported by other published reports, such as in Wildgruber et al.[30], where functional MRI studies associated linguistic prosody with the left IFG, while emotional prosodic processing was represented in orbitofrontal areas bilaterally.

We found a correlation of acoustic markers with CSF p-Tau levels in the non-AD subset of PPA patients. We recently found a linear association of antermortem CSF p-Tau, but not t-tau, to postmortem severity of tau pathology in the brain of FTLD patients[14]. Patients with confirmed FTLD-Tau pathology had higher CSF p-tau levels than their counterparts with confirmed FTLD-TDP pathology. Thus, the link we report here between acoustic speech markers and CSF p-Tau levels is consistent with the hypothesis connecting markers of dysprosody to the diagnosis of FTLD-Tau pathology[43, 44,45], which is the most prominent pathology underlying naPPA[46, 47]. This finding remains to be confirmed in a larger autopsy sample or in future studies with *in vivo* PET tau molecular

imaging that can detect FTLD-Tau pathology when this is developed. Because this speech analysis is highly repeatable with minimal learning effects, it can potentially serve as a surrogate end-point in treatment trials targeting tau pathology in patients with FTLD-Tau pathology.

Strengths & limitations

Strengths of our study include the objective and reliable measurement of speech intonation and rhythm without the use of subjective ratings. The use of the SAD enabled automatic analyses directly from digitized audio recording, freeing us from the time-consuming and laborious work of transcription. The SAD is also independent of a specific language and thus can theoretically be applied cross-linguistically without pre-programming. The use of a natural speech sample is an advantage when considering its possible implications in clinical research requiring repeated and frequent evaluations with minimal learning effects. Nevertheless, there are some shortcomings in our study. We were able to examine only a small group of PPA patients. Pitch-tracking involves complicated computational algorithms that estimate f0[48] and are subject to many errors, especially when confronting background noise, unfavorable voice quality, octave jumps in pitch and overlapping speech. We applied multiple quality control measures to minimize pitch-tracking inaccuracies and confounds both at the tracking level and in our statistical analyses. Due to the nature of the SAD we can only relate these acoustic data to the prosodic "breath-group". For a more detailed analysis at the sentence, word, syllable or phoneme level, a complete alignment of the sound to its transcript is needed. While we used autopsy-verified levels of CSF analytes to characterize participants in this study, we did not have autopsy evidence of specific pathology in all participants.

With these caveats in mind, this work reports implementation of the SAD as a novel automated speech analysis tool in the study of PPA. We identified characteristics of speech that distinguish PPA phenotypes, linking these to other language characteristics of naPPA, left frontal cortical atrophy and CSF levels of p-Tau. Speech analyses in naPPA thus may be an informative marker to screen for FTLD-Tau pathology. These findings support the potential use of the SAD in the study of the cognitive processes underlying

speech and for the measuring of a naturalistic, repeatable endpoint in clinical treatment trials.

Acknowledgments

This work was supported in part by the National Institutes of Health (AG017586, AG038490, NS053488, AG053940, AG052943, NS088341), the Institute on Aging and the Wyncote Foundation.

We thank our patients and their caregivers and family members for their continuous effort and contribution to our clinical research.

Author Contributions

Study concept and design: N.N, M.G; data acquisition, analysis and interpretation: all authors; drafting of the manuscript: N.N; critical review of manuscript: all authors.

Potential Conflicts of Interest

Dr. Nevler reports grants from Institute on Aging (IOA) and the National Institutes of Health (NIH); Dr. Ash reports grants from National Institutes of Health the Wyncote Foundation; Dr. Irwin reports grants from the National Institutes of Health (NIH); Dr. Liberman has nothing to disclose; Dr. Grossman reports grants from the National Institutes of Health, the Wyncote Foundation and Biogen, non-financial support from Avid Radiopharmaceuticals and Piramal, personal fees from UCB not related to the submitted work.

REFERENCES

1. Ethofer T, Anders S, Erb M, et al. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. Neuroimage. 2006 Apr 1;30(2):580-7.

2. Leitman DI, Wolf DH, Ragland JD, et al. "It's Not What You Say, But How You Say it": A Reciprocal Temporo-frontal Network for Affective Prosody. Front Hum Neurosci. 2010;4:1-13.

3. Pell MD. Fundamental Frequency Encoding of Linguistic and Emotional Prosody by Right Hemisphere-Damaged Speakers. Brain and Language. 1999 9//;69(2):161-92.

4. Pichon S, Kell CA. Affective and sensorimotor components of emotional prosody generation. J Neurosci. 2013 Jan 23;33(4):1640-50.

5. Ross ED, Monnot M. Neurology of affective prosody and its functional-anatomic organization in right hemisphere. Brain Lang. 2008 Jan;104(1):51-74.

6. Ryant N. LDC HMM Speech Activity Detector (v.1.0.4). University of Pennsylvania; 2013.

7. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011 Mar 15;76(11):1006-14.

8. Giannini LAA, Irwin DJ, McMillan CT, et al. Clinical marker for Alzheimer disease pathology in logopenic primary progressive aphasia. Neurology. 2017 Jun 13;88(24):2276-84.

9. Ash S, Evans E, O'Shea J, et al. Differentiating primary progressive aphasias in a brief sample of connected speech. Neurology. 2013 July 23, 2013;81(4):329-36.

10. Goodglass H, Kaplan E, Weintraub S. Boston Diagnostic Aphasia Examination. Philadelphia: Lea & Febiger; 1983.

11. Nevler N, Ash S, Jester C, Irwin DJ, Liberman M, Grossman M. Automatic measurement of prosody in behavioral variant FTD. Neurology. 2017 Aug 15 2017;89:650-6.

12. Gunawardena D, Ash S, McMillan C, Avants B, Gee J, Grossman M. Why are patients with progressive nonfluent aphasia nonfluent? Neurology. 2010 Aug 17;75(7):588-94.

13. Liberman P. Intonation, Perception, and Language. Cambridge, Massachusetts: M.I.T Press; 1968.

14. Irwin DJ, Lleo A, Xie SX, et al. Ante mortem cerebrospinal fluid tau levels correlate with postmortem tau pathology in frontotemporal lobar degeneration. Ann Neurol. 2017 Aug;82(2):247-58.

15. Lleó A, Irwin D, Illán-Gala I, et al. A CSF algorithm for the selection of FTLD subtypes in an autopsy cohort. JAMA Neurol.In press.

16. Team R. RStudio: Integrated Development for R. 0.99.879 ed. Boston, MA: RStudio, Inc.; 2015.

17. Marchetti GM, Drton M, Sadeghi K. ggm: Functions for graphical Markov models. 2.3 ed2015. p. R package.

18. Revelle W. psych: Procedures for Personality and Psychological Research. 1.7.5 ed. Northwestern University, Evanston, Illinois, USA.2017.

19. Wickham H. Reshaping data with the reshape package. Journal of Statistical Software. 2007;21(12).

20. Xavier R, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.

21. Yoshida K, Bohn. J. tableone: Create "Table 1" to Describe Baseline Characteristics. 2015.

22. Morales M. sciplot: Scientific Graphing Functions for Factorial Designs. 1.1-1 ed2017.

23. Fox J, Weisberg S. An {R} Companion to Applied Regression. Thousand Oaks CA: Sage; 2011.

24. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software. 2011;40(1):1-29.

25. Venables WN, Ripley BD. Modern applied statistics with S. 4 ed. New York: Springer; 2002.

26. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. NeuroImage. 2012;62:782-90.

27. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. NeuroImage. 2014;92:381-97.

28. Aziz-Zadeh L, Sheng T, Gheytanchi A. Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. PLoS One. 2010;5(1):e8759.

29. Pell MD. The Temporal Organization of Affective and Non-Affective Speech in Patients with Right-Hemisphere Infarcts. Cortex. 1999 //;35(4):455-77.

30. Wildgruber D, Ackermann H, Kreifelts B, Ethofer T. Cerebral processing of linguistic and emotional prosody: fMRI studies. Progress in brain research. 2006;156:249-68.

31. Rohrer JD, Sauter D, Scott S, Rossor MN, Warren JD. Receptive prosody in nonfluent primary progressive aphasias. Cortex. 2012 Mar;48(3):308-16.

32. Charles D, Olm C, Powers J, et al. Grammatical comprehension deficits in non-fluent/agrammatic primary progressive aphasia. J Neurol Neurosurg Psychiatry. 2014 Mar;85(3):249-56.

33. Mack JE, Chandler SD, Meltzer-Asscher A, et al. What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? Neuropsychologia. 2015 Oct;77:211-22.

34. Pakhomov SV, Smith GE, Chacon D, et al. Computerized analysis of speech and language to identify psycholinguistic correlates of frontotemporal lobar degeneration. Cogn Behav Neurol. 2010 Sep;23(3):165-77.

35. Duffy JR, Hanley H, Utianski R, et al. Temporal acoustic measures distinguish primary progressive apraxia of speech from primary progressive aphasia. Brain Lang. 2017 Feb 07;168:84-94.

36. Ladefoged P, Disner SF. Vowels and Consonants: Wiley; 2012.

37. Hodges JR, Patterson K. Nonfluent progressive aphasia and semantic dementia: a comparative neuropsychological study. Journal of the International Neuropsychological Society : JINS. 1996 Nov;2(6):511-24.

38. Mesulam MM, Wieneke C, Thompson C, Rogalski E, Weintraub S. Quantitative classification of primary progressive aphasia at early and mild impairment stages. Brain. 2012 May;135(Pt 5):1537-53.

39. Cousins KA, York C, Bauer L, Grossman M. Cognitive and anatomic double dissociation in the representation of concrete and abstract words in semantic variant and behavioral variant frontotemporal degeneration. Neuropsychologia. 2016 Apr;84:244-51.

40. Bonner MF, Price AR, Peelle JE, Grossman M. Semantics of the visual environment encoded in parahippocampal cortex. Journal of cognitive neuroscience. 2016 12/17;28(3):361-78.

41. Sajjadi SA, Patterson K, Arnold RJ, Watson PC, Nestor PJ. Primary progressive aphasia: a tale of two syndromes and the rest. Neurology. 2012 May 22;78(21):1670-7.

42. Mesulam MM, Weintraub S, Rogalski EJ, Wieneke C, Geula C, Bigio EH. Asymmetry and heterogeneity of Alzheimer's and frontotemporal pathology in primary progressive aphasia. Brain. 2014 Apr;137(Pt 4):1176-92.

43. Kuiperij HB, Versleijen AA, Beenes M, et al. Tau Rather than TDP-43 Proteins are Potential Cerebrospinal Fluid Biomarkers for Frontotemporal Lobar Degeneration Subtypes: A Pilot Study. J Alzheimers Dis. 2017;55(2):585-95.

44. Irwin DJ, Trojanowski JQ, Grossman M. Cerebrospinal fluid biomarkers for differentiation of frontotemporal lobar degeneration from Alzheimer's disease. Frontiers in aging neuroscience. 2013;5:1 - 11.

45. Irwin DJ, Lleó A, McMillan CT, et al. Ante mortem CSF tau levels correlate with post mortem tau pathology in FTLD. Annals of Neurology. 2017;82:247–58.

46. Grossman M. The non-fluent/agrammatic variant of primary progressive aphasia. The Lancet Neurology. 2012 Jun;11(6):545-55.

47. Josephs KA, Hodges JR, Snowden JS, et al. Neuropathological background of phenotypical variability in frontotemporal dementia. Acta neuropathologica. 2011 Aug;122(2):137-53.

48. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proceedings of the Institute of Phonetic Sciences. 1993;17(University of Amsterdam):97–110.

Figure 1: f0 and durations data

(A) f0 percentiles by clinical phenotype, expressed in semitones (ST). The 90th percentile represents the f0 range. (B) Pause rate, calculated as the number of pauses per minute of speech time. (C) Mean speech duration. (D) Silent pause mean duration.

f0 – fundamental frequency; ST – semitones; HC – healthy controls; lvPPA – logopenic variant Primary progressive aphasia; naPPA – nonfluent/agrammatic primary progressive aphasia; svPPA – semantic variant primary progressive aphasia; sec – seconds; ppm – pauses per minute.

Figure 2: Correlations of automated measures with manual coding

A – D Correlations of automatically extracted pause rate with manually coded measures of fluency and grammaticality.

E – H Correlations of automatically extracted mean speech segment duration with manual coding.

The mirror image between the upper and lower panels coincides with the strong negative correlation between speech duration and pause rate (see text).

lvPPA – logopenic variant Primary progressive aphasia; naPPA – nonfluent/agrammatic primary progressive aphasia; svPPA – semantic variant primary progressive aphasia; sec – seconds; ppm – pauses per minute; wpm – words per minute; WFS – well-formed sentences; DC – dependent clauses.

Figure 3: ROC analyses

A) f0 range and pause rate as single classifiers for Receiver operating characteristic (ROC) curve of naPPA vs. HC. A combined acoustic classifier (pause rate / f0 range) improves AUC (0.94).

B) Combined acoustic parameter (pause rate / f0 range) as classifier for naPPA vs. other phenotypes.

AUC – area under the curve; HC – healthy controls; lvPPA – logopenic variant Primary progressive aphasia; naPPA – nonfluent/agrammatic primary progressive aphasia; svPPA – semantic variant primary progressive aphasia; ROC – receiver operating curves.

Figure 4: Gray matter (GM) density analysis

Significant regression of impaired f0 range with cortical atrophy in naPPA (n=9) is shown in red in the left inferior frontal cortex (BA 44).

Figure 5: CSF p-Tau correlation with combined acoustic marker

Pearson correlation showing linear association between the natural logarithm of CSF p-Tau levels and the natural logarithm of the combined acoustic marker (r=0.58, p=0.007).

lvPPA – logopenic variant Primary progressive aphasia; naPPA – nonfluent/agrammatic primary progressive aphasia; svPPA – semantic variant primary progressive aphasia

Table 1: Mean (SD) demographic characteristics of patients and controls

	HC	naPPA	lvPPA	svPPA	p
n	31	15	23	21	
Age, y	69.29 (7.90)	69.67 (9.20)	65.91 (9.83)	64.48 (7.71)	0.14
Sex = Male (%)	11 (35.5)	6 (40.0)	7 (30.4)	10 (47.6)	0.68
Education, y	15.97 (2.58)	14.80 (3.12)	15.35 (3.19)	15.10 (2.81)	0.56
Disease duration, y	NA	2.60 (1.12)	4.00 (2.00)	4.05 (2.04)	0.04
MMSE total (0-30), n=85	29.00 (1.07)	24.73 (5.24)	23.05 (5.72)	23.05 (6.11)	<0.001
F letter fluency, n=41	17.75 (8.10)	6.33 (3.04)	6.36 (5.40)	8.21 (3.96)	0.001
Digit span forward, n=69	7.00 (1.37)	5.61 (1.30)	4.45 (1.54)	6.06 (1.89)	<0.001
Digit span backward, n=74	5.65 (1.31)	2.64 (1.11)	2.91 (1.08)	3.78 (1.70)	<0.001
Category fluency[a], n=54	19.67 (6.48)	10.11 (5.09)	9.85 (5.94)	5.36 (4.67)	<0.001
Speech rate, wpm	140.06 (36.74)	61.00 (24.85)	88.17 (36.09)	113.95 (40.76)	<0.001
MLU (words)	10.57 (1.98)	6.74 (2.38)	8.46 (2.45)	8.61 (2.75)	<0.001
DC / utterance[b]	0.37 (0.23)	0.05 (0.09)	0.21 (0.21)	0.32 (0.27)	<0.001
WFS / utterance	0.91 (0.11)	0.72 (0.32)	0.71 (0.25)	0.78 (0.19)	0.003

[*] MMSE total score did not differ between patient groups.

[a] Category = animals; [b] Data refers to the average number of clauses per utterance.

wpm - words per minute; MLU - mean length of utterance; DC - dependent clauses; WFS - well-formed sentences.

Table 2: Results of polynomial logistic regression

	f0 range			**Pause rate**		
	OR	95% CI	p	OR	95% CI	p
naPPA	0.35	0.18 - 0.69	**0.002**	1.18	1.10 - 1.26	**<0.001**
lvPPA	0.82	0.54 - 1.26	0.37	1.17	1.10 - 1.24	**<0.001**
svPPA	1.09	0.77 - 1.18	0.62	1.11	1.05 - 1.18	**<0.001**

OR – Odds Ratio
CI – Confidence Interval

Table 3: GM atrophy and regression results for f0 range in naPPA

Localization (BA)	MNI coordinates X	Y	Z	Maximal P	Cluster Size (Voxels)
GM Atrophy naPPA < Ctrl					
	Subpeak coordinates				
Lt. Orbitofrontal (11)	-10	14	-26	<0.001	4991
Lt. Prefrontal (9)	-30	36	46		
Lt. Premotor (6)	-34	6	64		
Lt. Anterior prefrontal (10)	-34	50	-16		
Lt. Premotor (6)	-56	8	38		
	Peak coordinates				
Lt. Prefrontal (8)	-4	18	40	0.003	86
Lt. Posterior cingulate (23)	-14	-28	36	<0.001	85
Lt. Primary somatosensory (1)	-40	-20	34	<0.001	73
Lt. Fusiform (37)	-58	-62	-18	<0.001	96
Lt. Visual association (18)	-2	-94	-14	<0.001	383
Rt. Pars opercularis (44)	40	8	22	<0.001	218
Rt. Supplementary motor (6)	16	-16	76	0.001	193
Rt. Anterior prefrontal (10)	30	52	-16	<0.001	129
Rt. Subgenual cingulate (25)	8	12	-26	<0.001	109
Rt. Anterior cingulate (24)	6	26	16	0.001	73
Rt. Anterior cingulate (24)	10	-18	38	<0.001	58
Rt. Prefrontal (9)	32	34	48	0.003	55
Rt. Pars orbitalis (47)	52	24	-10	0.001	52
Rt. Fusiform (37)	56	-66	-20	<0.001	181
Rt. Posterior cingulate (23)	8	-38	24	<0.001	160
Rt. Posterior cingulate (23)	10	-40	34	<0.001	146
Rt. Inferior temporal (20)	26	-6	-46	<0.001	483
Rt. Posterior medulla oblongata	6	-44	-52	<0.001	264
Regression of f0 Range with GM Atrophy in naPPA					
	Peak coordinates				
Lt. Pars opercularis (44)	-52	10	-2	<0.001	11

BA – Brodmann Area
MNI – Montreal Neurological Institute

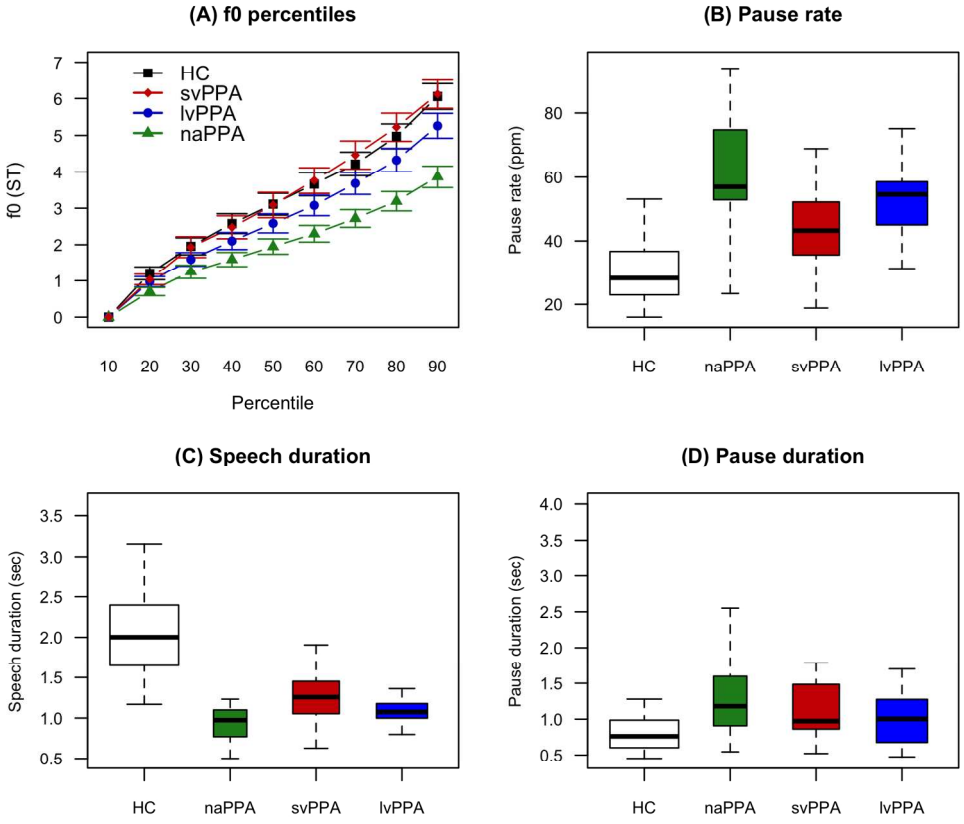

Figure 1: f0 and durations data
(A) f0 percentiles by clinical phenotype, expressed in semitones (ST). The 90th percentile represents the f0 range. (B) Pause rate, calculated as the number of pauses per minute of speech time. (C) Mean speech duration. (D) Silent pause mean duration.
f0 - fundamental frequency; ST - semitones; HC - healthy controls; lvPPA - logopenic variant Primary progressive aphasia; naPPA - nonfluent/agrammatic primary progressive aphasia; svPPA - semantic variant primary progressive aphasia; sec - seconds; ppm - pauses per minute.
f0 - fundamental frequency; ST - semitones; HC - healthy controls; lvPPA - logopenic variant Primary progressive aphasia; naPPA - nonfluent/agrammatic primary progressive aphasia; svPPA - semantic variant primary progressive aphasia; sec - seconds; ppm - pauses per minute.

708x624mm (72 x 72 DPI)

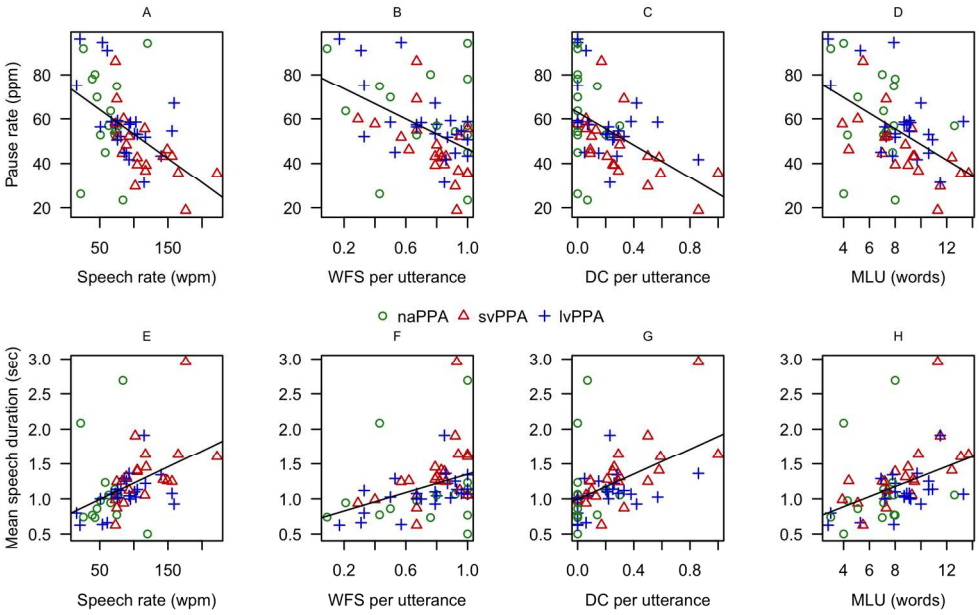

Figure 2: Correlations of automated measures with manual coding
A - D Correlations of automatically extracted pause rate with manually coded measures of fluency and grammaticality.
E - H Correlations of automatically extracted mean speech segment duration with manual coding.
The mirror image between the upper and lower panels coincides with the strong negative correlation between speech duration and pause rate (see text).
lvPPA - logopenic variant Primary progressive aphasia; naPPA - nonfluent/agrammatic primary progressive aphasia; svPPA - semantic variant primary progressive aphasia; sec - seconds; ppm - pauses per minute; wpm - words per minute; WFS - well-formed sentences; DC - dependent clauses.

708x458mm (72 x 72 DPI)

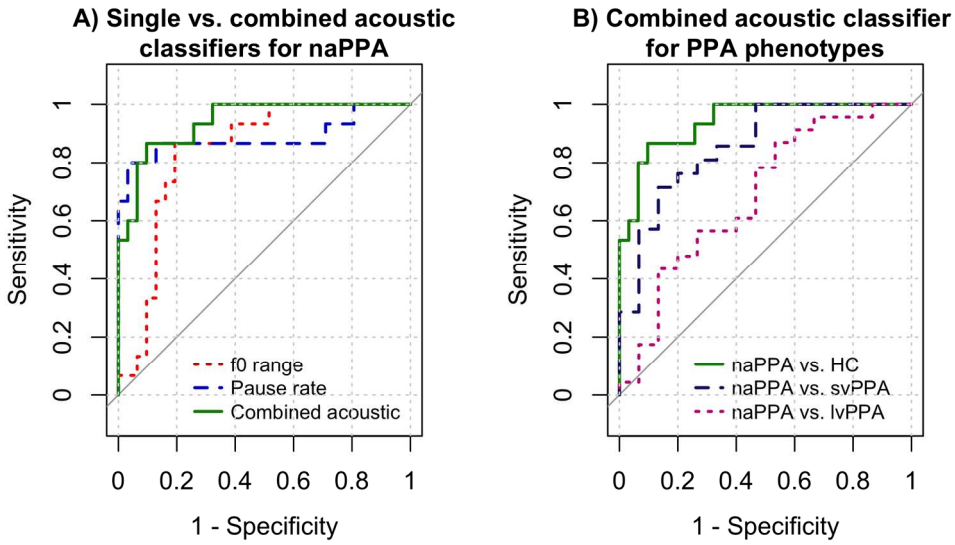

Figure 3: ROC analyses
A) f0 range and pause rate as single classifiers for Receiver operating characteristic (ROC) curve of naPPA vs. HC. A combined acoustic classifier (pause rate / f0 range) improves AUC (0.94).
B) Combined acoustic parameter (pause rate / f0 range) as classifier for naPPA vs. other phenotypes.
AUC - area under the curve; HC - healthy controls; lvPPA - logopenic variant Primary progressive aphasia; naPPA - nonfluent/agrammatic primary progressive aphasia; svPPA - semantic variant primary progressive aphasia; ROC - receiver operating curves.

708x416mm (72 x 72 DPI)

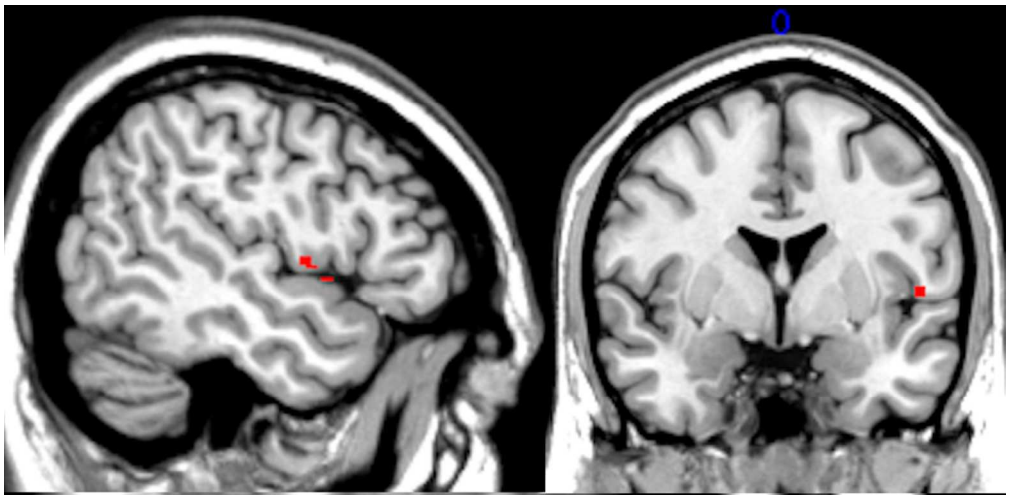

Figure 4: Gray matter (GM) density analysis
Significant regression of impaired f0 range with cortical atrophy in naPPA (n=9) is shown in red in the left inferior frontal cortex (BA 44).

80x39mm (300 x 300 DPI)

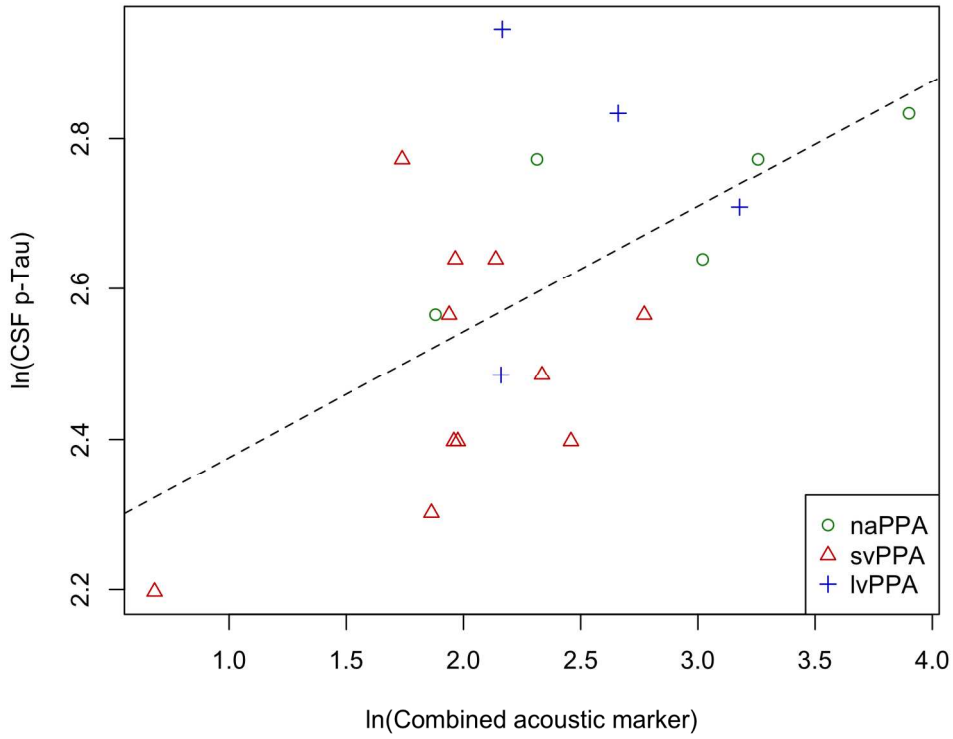

Figure 5: CSF p-Tau correlation with combined acoustic marker
Pearson correlation showing linear association between the natural logarithm of CSF p-Tau levels and the natural logarithm of the combined acoustic marker (r=0.58, p=0.007).
lvPPA - logopenic variant Primary progressive aphasia; naPPA - nonfluent/agrammatic primary progressive aphasia; svPPA - semantic variant primary progressive aphasia

708x624mm (72 x 72 DPI)