Most research on prosody has relied on subjective assessments, often focusing on the expression or comprehension of emotional speech [1].

A longitudinal study of automated analysis of acoustic speech markers in FTD & PPA

Naomi Nevler, MD[1], Sharon Ash, PhD[1], Sunghye Cho, PhD[2], Sanjana Shellikeri, PhD[1], Natalia Parjane[1], David J Irwin, MD[1], Mark Liberman, PhD[2], Murray Grossman, MDCM, EdD[1]

(1) Department of Neurology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; (2) Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA, USA

Introduction

- Speech is a complex activity requiring proper function and connectivity of multiple brain networks and as such is sensitive to focal neurodegeneration.
- We have previously reported acoustic markers of dysprosody in speech samples of speakers with frontotemporal dementia (FTD) phenotypes.
- In the current study, we explore the longitudinal changes in acoustic-prosodic markers in FTD.

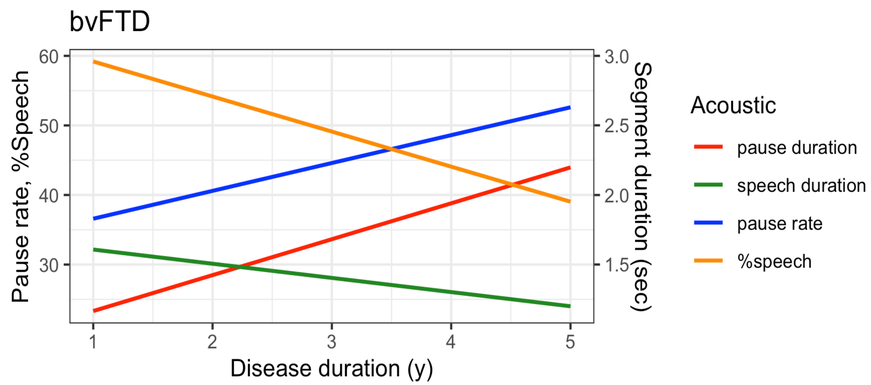
Participants & Methods

Table 1: Clinical & Demographic characteristics (mean (sd))

	bvFTD	lvPPA	naPPA	svPPA	p
n	14	10	8	14	
Sex = Male (%)	12 (85.7)	3 (30.0)	0 (0.0)	5 (35.7)	0.001
Age (y)	66.29 (6.40)	63.60 (8.85)	70.62 (7.09)	61.36 (9.33)	0.074
Age at onset (y)	64.00 (5.96)	60.70 (9.19)	67.88 (6.88)	58.71 (9.45)	0.068
Education (y)	15.64 (1.82)	14.60 (2.72)	13.88 (1.73)	15.00 (2.39)	0.331
Disease duration (y)	2.29 (1.33)	2.80 (1.32)	2.75 (0.89)	2.64 (0.74)	0.656

- We analyzed 102 speech samples of picture descriptions from 46 participants with FTD (Table 1): 8 with non-fluent/agrammatic primary progressive aphasia (naPPA), 14 with semantic variant PPA (svPPA), 10 with logopenic PPA (lvPPA) and 14 with behavioral variant FTD (bvFTD).
- We automatically segmented the acoustic signal into segments of continuous speech or silence, measured their durations, and derived other measures.
- We used linear mixed effects (lme) models to test changes over time for each acoustic measure, controlling for sex, education, and random intercepts. We also examined any interaction between phenotypes and disease duration.

Results



- bvFTD speakers increased their pause duration by 0.27 seconds per year and their pause rate by 3.9 pauses per minute (ppm) each year. Their speech segment duration shortened by 0.1 seconds per year (p=0.041), decreasing their total speech time by 6.6 seconds (p=0.003) per year. Thus, bvFTD patients reduced the proportion of speech in their samples by 5.16 percent per year (p=0.008).
- svPPA speakers increased their pause rate similarly, but in contrast, their pause duration decreased by 0.097 seconds per year and they increased their speech segment frequency by 8.32 per minute each year (p=0.054).
- naPPA and lvPPA speakers increased their pause rate over time and spent less total time (speech + pause) describing the picture (by 5.6 seconds per year; p=0.018). They did not differ from bvFTD and svPPA in these two acoustic measures.

Conclusions

- All FTD speakers became more dysfluent and produced shorter descriptions with time.
- Only bvFTD speakers exhibited reduced speech production.
- svPPA speakers produced more frequent pauses and speech segments over time, rendering their speech “fragmented” and inefficient.
- These findings support the role of automated acoustic analysis in characterizing speech longitudinally in FTD.

All authors have nothing to disclose. This work was supported by an Alzheimer’s Association grant AACSF-18-567131, NIH grants AG052943, AG017586, AG032953, AG038490, AG043503, NS053488, NS088341, and the Wyncote Foundation.

Correspondence: naomine@pennmedicine.upenn.edu