



Measurement of speech parameters in casual speech of dementia patients

SOA 2017 12 Sept 13:30

Roelant Ossewaarde^{a,b}, Roel Jonkers^a, Fedor Jalvingh^{a,c}, Roelien Bastiaanse^a

^a Center for Language and Cognition, University of Groningen (NL)

^b Institute for ICT, HU University of Applied Science, Utrecht (NL)

^c St. Marienhospital Vechta, Geriatric Clinic Vechta (DE)



Motivation

- Pilot for a larger study with goal:
 - Follow participants with Primary Progressive Aphasia (PPA) to map decline of language
 - Study how to use software to measure decline
- Often cited characteristics of PPA speech:
 - Non-fluent:
 - agrammatism, hesitant or labored speech, word finding problems
 - Fluent:
 - fluent but empty speech, comprehension of word meaning and object recognition are affected
- Speech type: connected speech.



Motivation

- › Available data:
 - Larger study with non-brain-damaged participants and participants with PPA, Parkinson's (PD), Alzheimer's (AD), Minor Cognitive Impairment (MCI).
 - Spontaneous speech.
 - German
- › Easiest accessible information:
 - Acoustic properties
- › Can we expect to build software that can tell apart the groups based on their voice in connected speech?

Background

- › Other approaches: Fraser et al. (2014), Jarrold et al. (2014), Orimaye et al. (2014) report success in a similar task for English.
 - Acoustic properties as well as other measures.
 - Machine learning to combine variables for prediction
- › Reported symptoms:
 - Alzheimer's: reduced pitch variation and syllable timing
 - Parkinson's: impaired voice quality, monopitch, late VOT
 - Non-fluent PPA: increased number of pauses

Expected findings

		PD	AD	PPA
Fluency	Pause length	+	=	+
	Pause frequency	+	=	+
Prosody	Pitch level	-	-	-
	Pitch range	-	-	-
Voice quality	Jitter	+	=	=
	Shimmer	+	=	=

(+: increase, -: decrease, =: no deviation expected)

Data

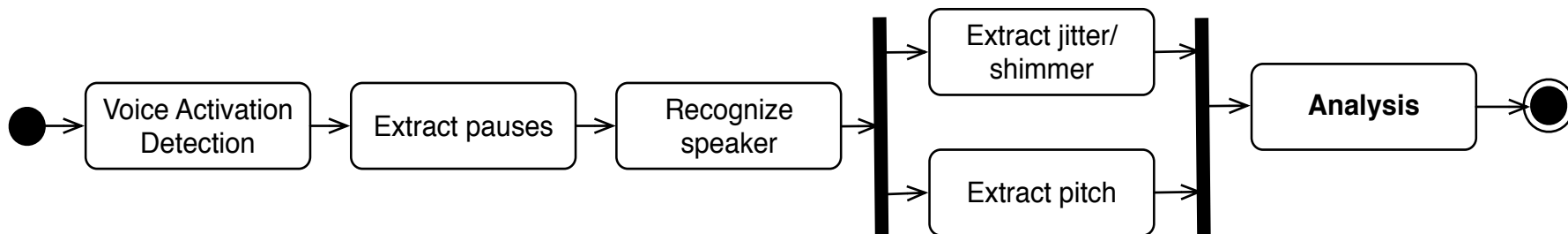
Participants

NBD	2m, 6f
AD	5m, 4f
PD	4m, 2f
PPA non-fluent	2m
PPA fluent	1f

Sample characteristics

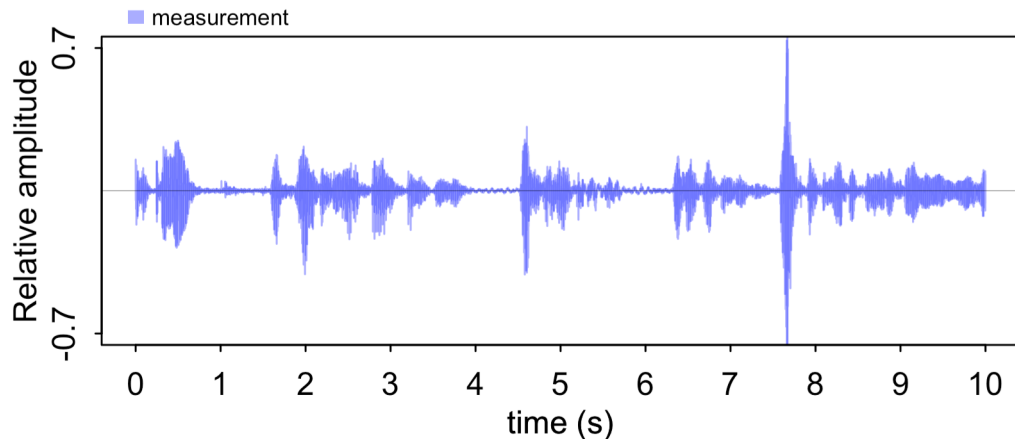
- > Three conversations, about a typical day in past, present, near future.
- > Repeated 3 times, at least 6 months apart.
- > Relatively long fragments:
 - μ 5:47 ($\pm$ 2:30)
 - 26 x 3 x 3 x 5:47: over 22 hrs speech
 - Manual broad transcription: 700 hours!
- > Interviewer and participant in same signal.

Extracting acoustics-only variables

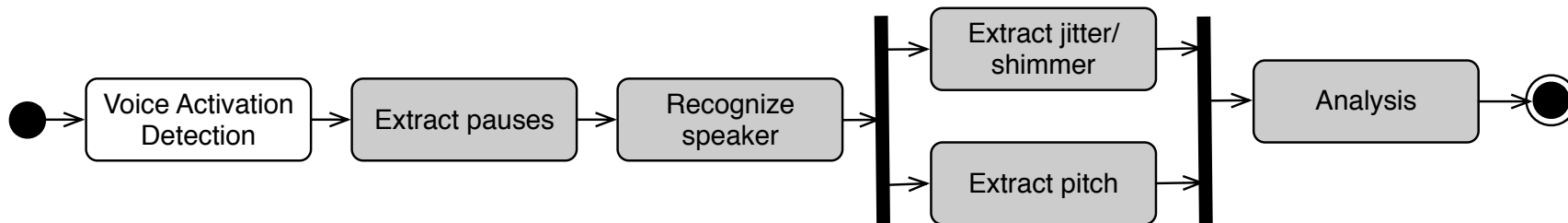


Relevant variables:

- Fluency and phonation
 - Pause length, frequency
- Voice quality
 - Jitter, shimmer
- Prosody
 - Pitch level, range

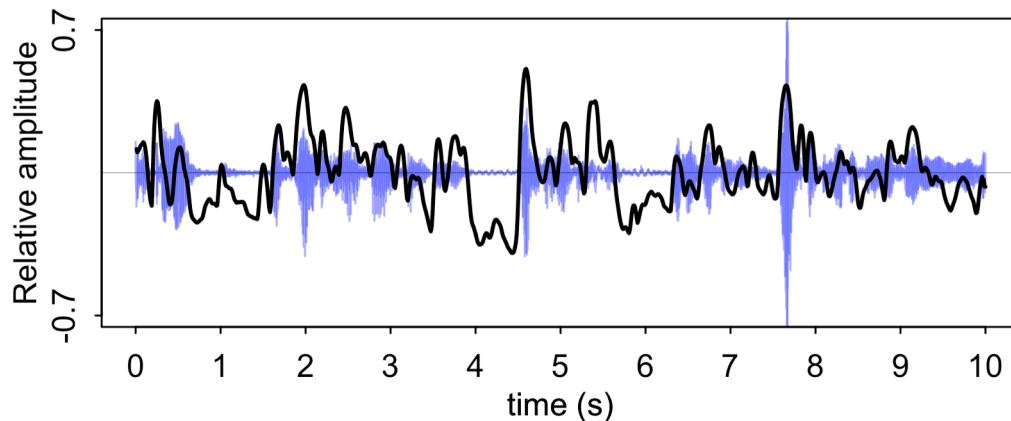


Extracting acoustics-only variables

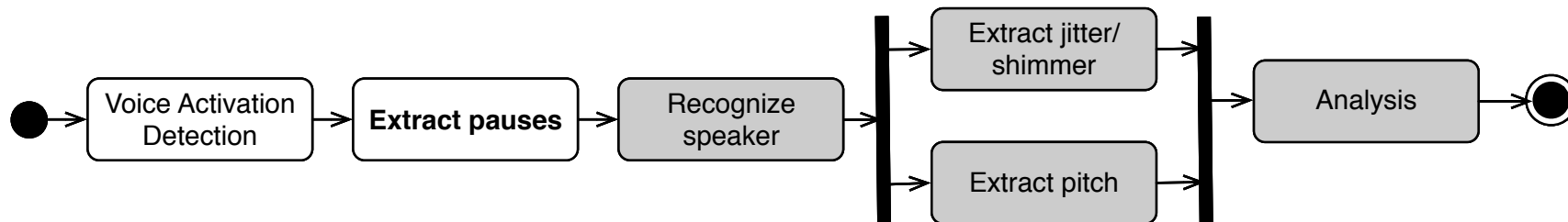


Relevant variables:

- Fluency and phonation
 - Pause length, frequency
- Voice quality
 - Jitter, shimmer
- Prosody
 - Pitch level, range

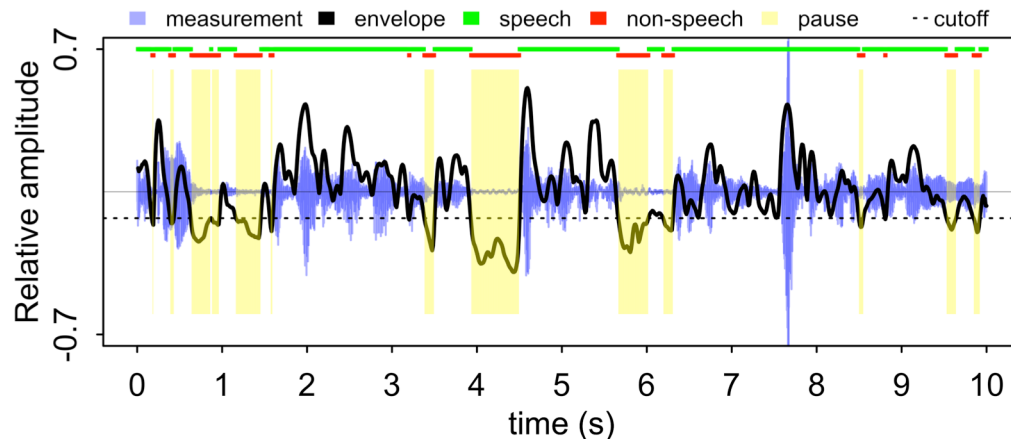


Extracting acoustics-only variables

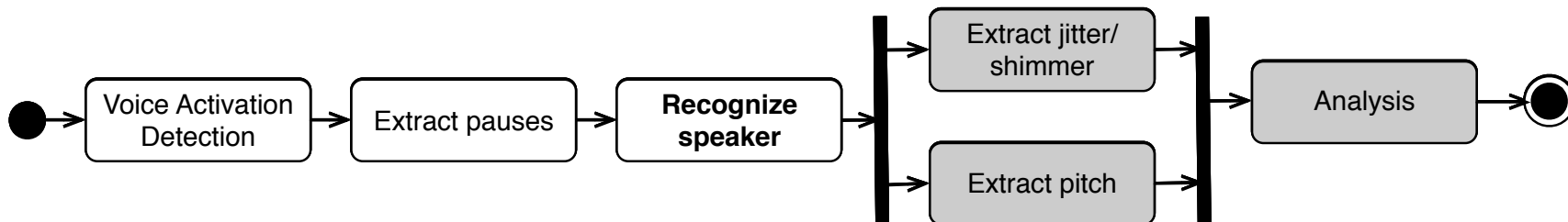


Relevant variables:

- Fluency and phonation
 - **Pause length, frequency**
- Voice quality
 - Jitter, shimmer
- Prosody
 - Pitch level, range

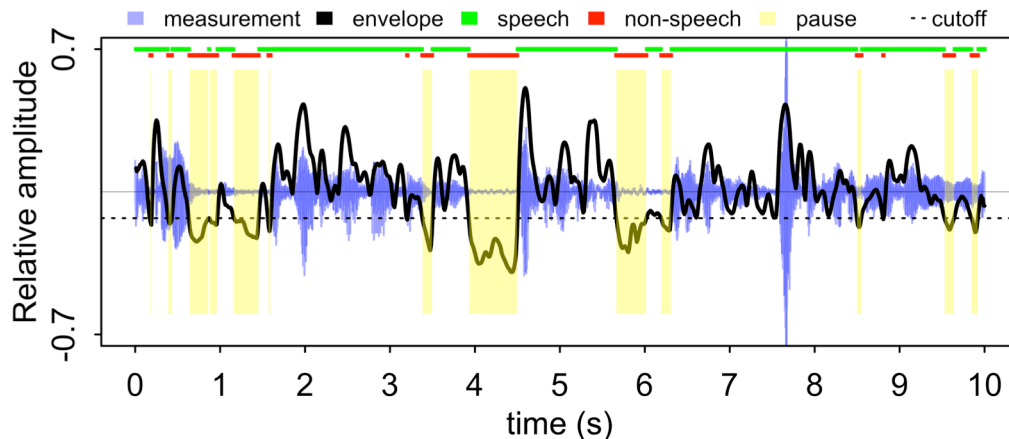


Extracting acoustics-only variables

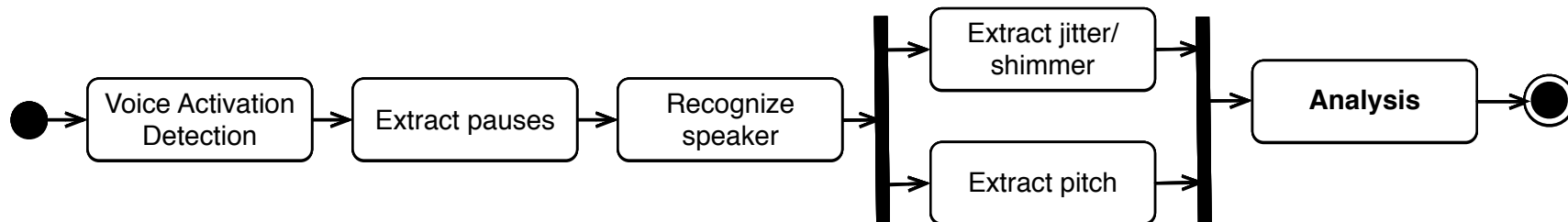


Relevant variables:

- Fluency and phonation
 - Pause length, frequency
- Voice quality
 - Jitter, shimmer
- Prosody
 - Pitch level, range

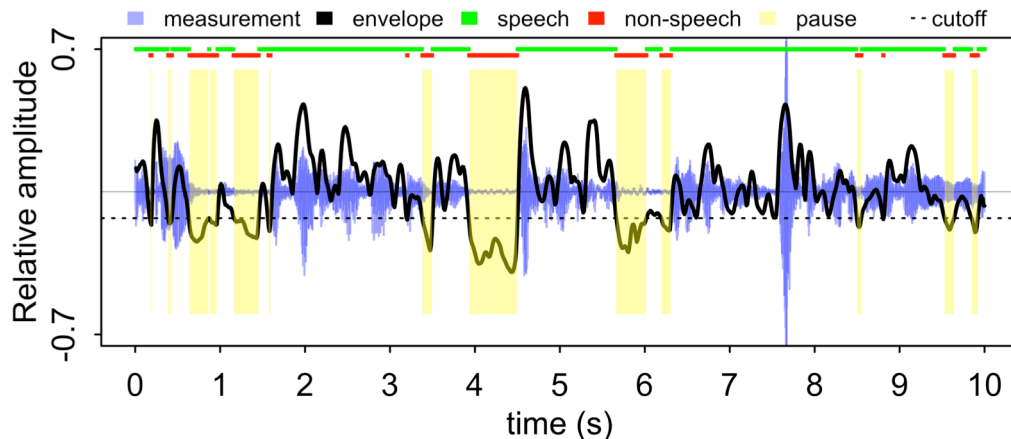


Extracting acoustics-only variables



Relevant variables:

- Fluency and phonation
 - Pause length, frequency
- Voice quality
 - Jitter, shimmer
- Prosody
 - Pitch level, range

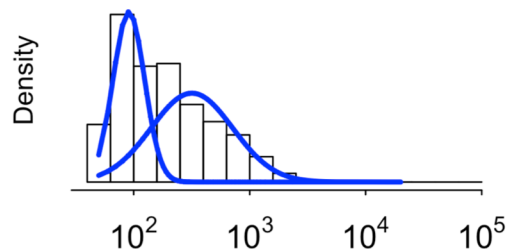


Outline of results

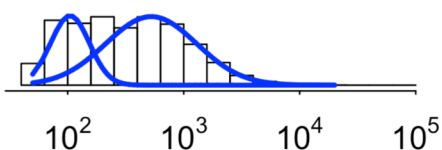
- › Fluency - pauses
 - No evidence that *frequency* or *length* of pauses is different between groups.
 - The *ratio short:long pause* is different for PPA-NF fragments.
- › Prosody
 - No evidence that *pitch level* or *pitch range* are different between groups.
- › Voice Quality
 - Measurements of jitter/shimmer in Parkinson's fragments are slightly different (if you measure them the right way).

Results: fluency - pauses

Non-brain-damaged, N=14661

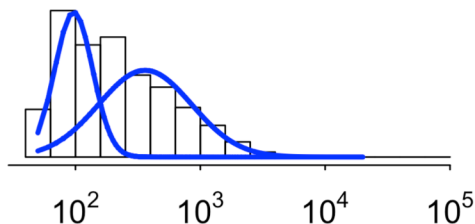


PPA PNFA #1, N=4525

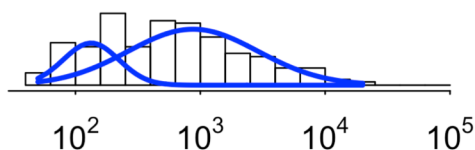


Pause length (milliseconds)

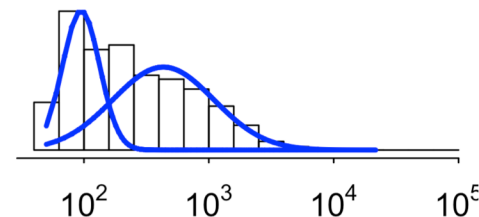
Alzheimer, N=38380



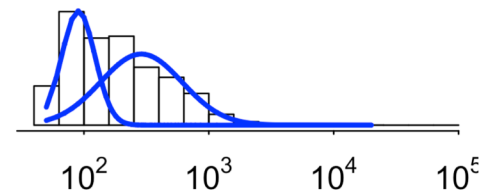
PPA PNFA #2, N=588



Parkinson, N=43051



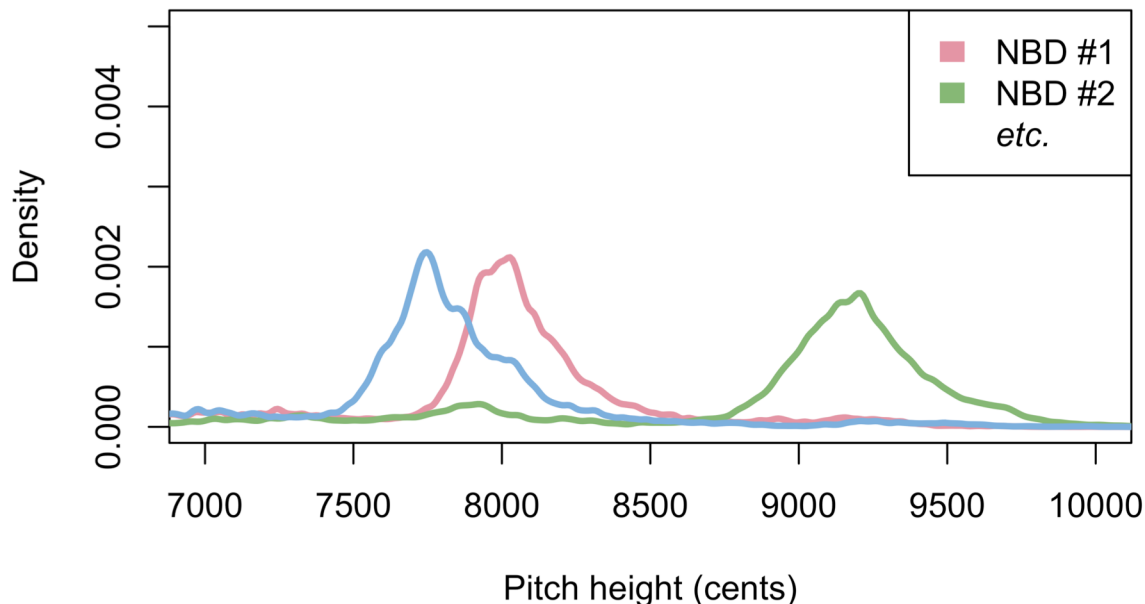
PPA SD, N=7251



Results: pitch

- Measured as pitch interval (Olness et al., 2010).
- More intra-group variation than inter-group variation.
- No evidence that pitch range or pitch height differ between groups.
- Relatively stable over time.

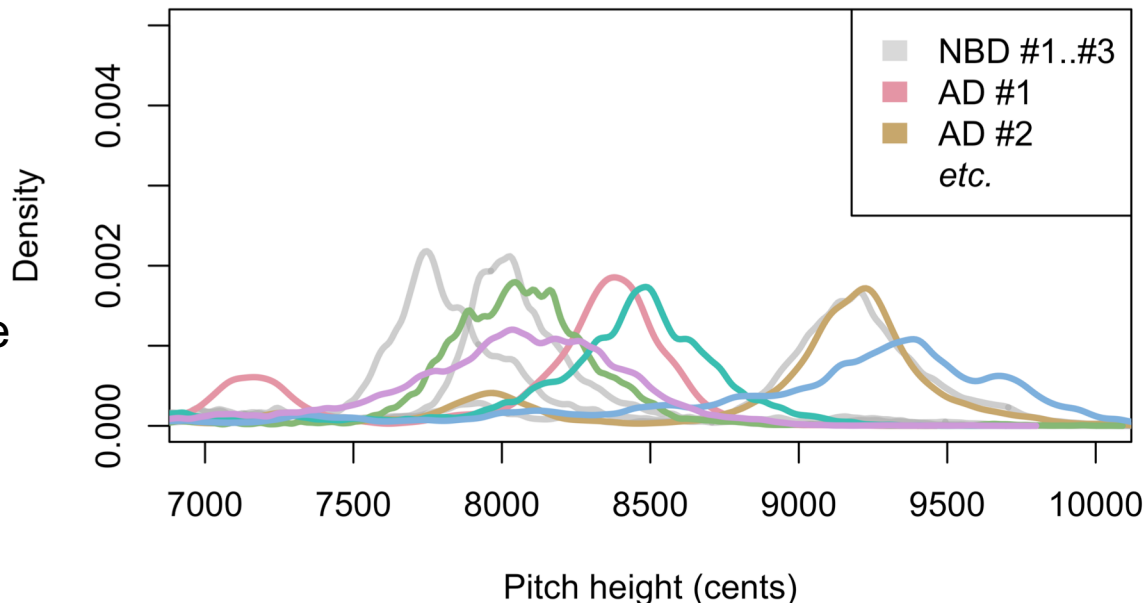
Non-brain-damaged



Results: pitch

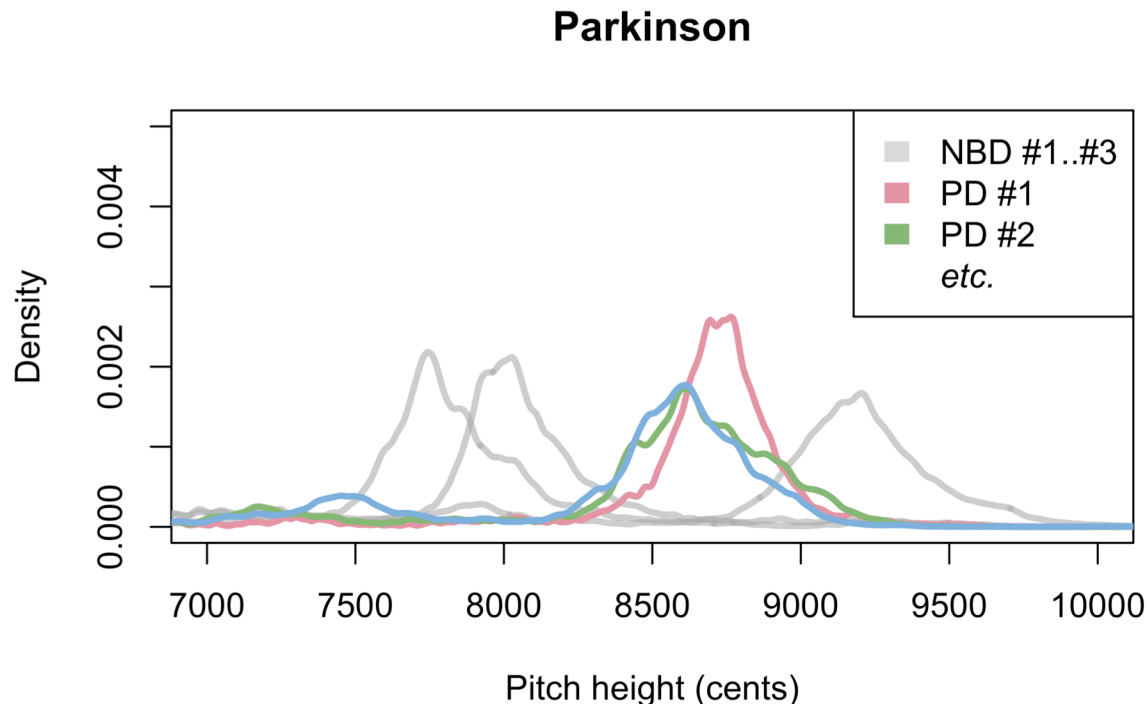
- Measured as pitch interval (Olness et al., 2010).
- More intra-group variation than inter-group variation.
- No evidence that pitch range or pitch height differs between groups.
- Pitch doesn't become 'more' monotonous over time

Alzheimer



Results: pitch

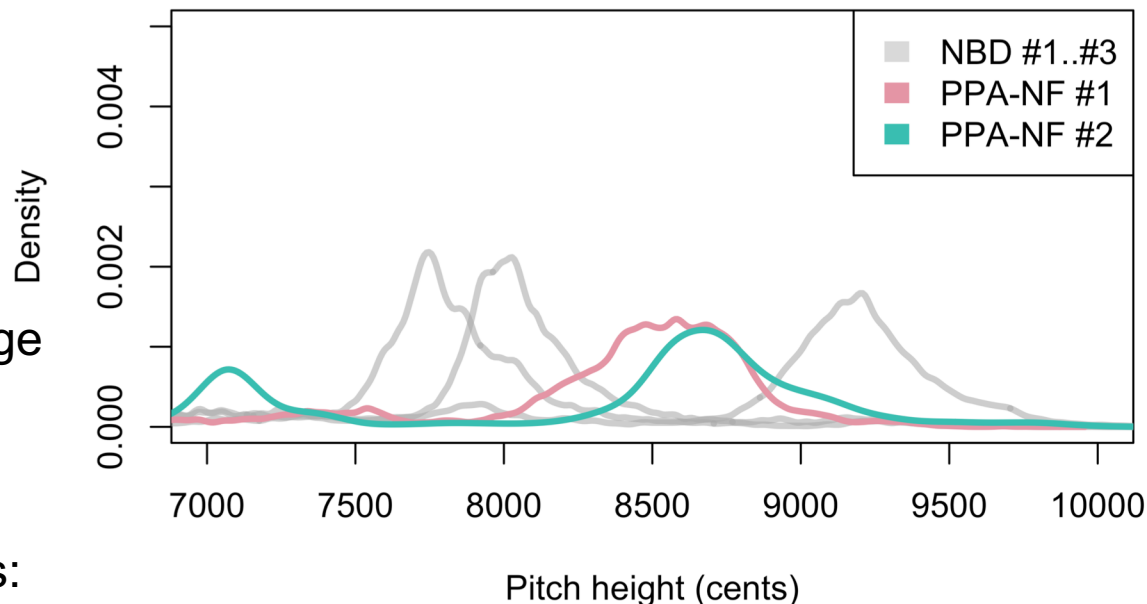
- Measured as pitch interval (Olness et al., 2010).
- More intra-group variation than inter-group variation.
- No evidence that pitch range or pitch height differ between groups.



Results: pitch

- Measured as pitch interval (Olness et al., 2010).
- More intra-group variation than inter-group variation.
- No evidence that pitch range or pitch height differ between groups.
- HDI comparisons of curves: no difference from NBD.

Primary Progressive Aphasia



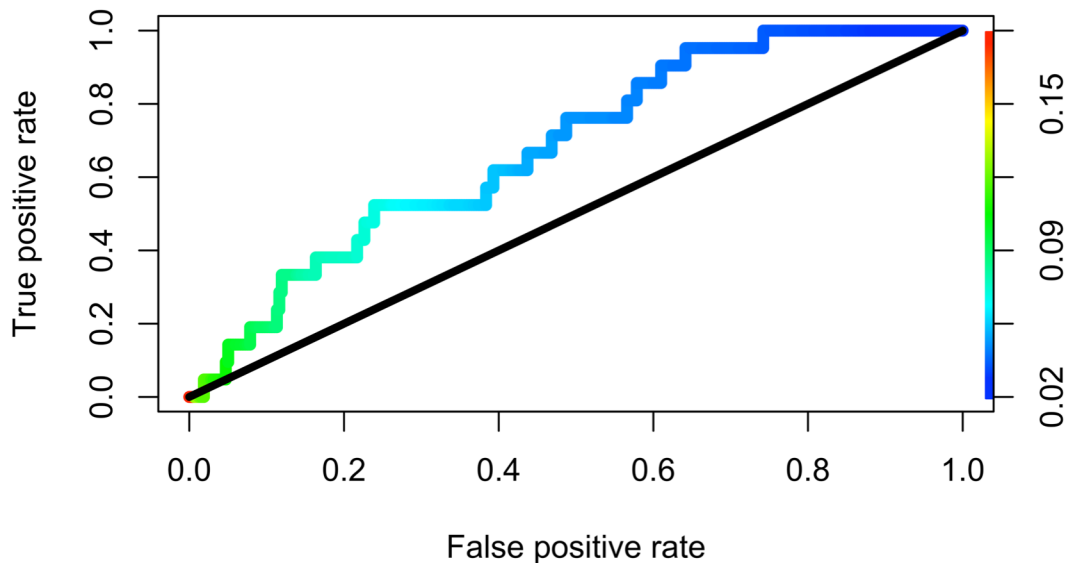


Results: voice quality - jitter

- Randomly sampled 10% of highest energy fragments and extracted vowels.
 - Different than typical clinical measure, but correlates ($R^2 = 0.7$) according to Rovirosa et al. (2008), Brockmann et al. (2011).
- Jitter/Shimmer as a measure ill-understood:
 - Perceptory correlates still unclear.
 - Each software package computes the same measure in a different way.
 - We used Baken & Orlikoff (2000) formulas – and our measurements differed too!
- Instability as computed over 5 and 3 second periods significant for PD fragments. (PPQ5/APQ3)
- All other measures: roughly the same.

Machine learning results

ROC curve (AUROC = 0.68), NBD versus dementia



Classifier details

- › Trained a Support Vector Machine (Vapnik 1995).
- › Uses variables:
 - Ratio long:short pauses
 - Voice quality (jitter/shimmer)
- › AUROC value of 0.68, borderline between “poor” and “fair”.

Expected findings

		PD	AD	PPA	Confirmed?
Fluency	Pause length	+	=	+	no
	Pause frequency	+	=	+	no
	Pause ratio	=	=	+	yes
Prosody	Pitch level	-	-	-	no
	Pitch variation	-	-	-	no
Voice quality	Jitter	+	=	=	yes
	Shimmer	+	=	=	yes

Discussion

Findings

- › Monopitch is not detectable in this sample.
- › Duration can be helpful after transformation.
- › Jitter/shimmer is promising, but as measure ill understood.
- › A model with pause length and jitter/shimmer outperforms chance, but borderline “poor”.
 - But fully automated!

Future directions

- › Perhaps this task isn't challenging enough.
- › Should compare to human ratings.
- › Add linguistic variables.
- › Thank you.

Results: monopitch in AD

- MMSE scores quantify progression of disease.
- No monopitch in this sample!

