



Longitudinal vs Cross-Sectional Speech Analysis in Dementia

Peter S. Pressman(1), Francesca Dino(2), Chelsea Chandler(3), Kevin Bretonnel Cohen(4), Steven Bedrick(5), Elliott Ross(2), Peter Foltz(3), Lawrence E. Hunter(6), MD

1) Neurology, Oregon Health & Science University, Portland, OR; 2) Neurology, University of Colorado Anschutz Medical Campus, Aurora, CO; 3)Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO; 4) Third Battalion, Ukrainian Volunteer Army, Ukraine; 5) Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR; 6)Pediatrics, University of Chicago, Chicago, IL

Introduction

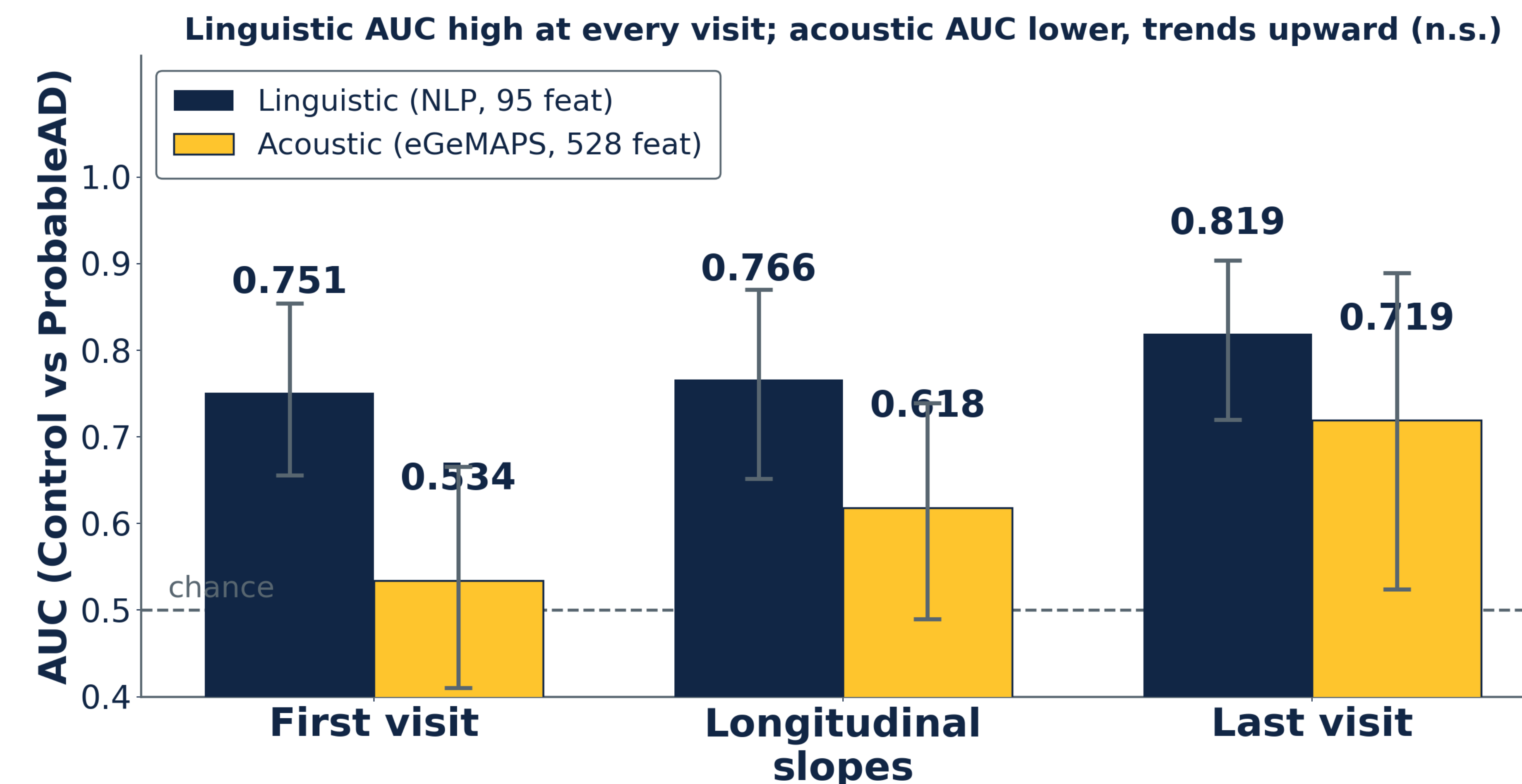
- Speech reflects cognition; dementia changes speech.
- Most speech-biomarker studies have been single-visit, patient versus control.
- But Alzheimer's disease and related disorders are diseases of trajectory.
- Q1: Does rate of change add diagnostic signal beyond one visit?
- Q2: Do linguistic and acoustic features differ in approximation of diagnosis versus progression?

Methods

- Corpus: DementiaBank Pitt Corpus, Cookie Theft; Control vs probable AD, two or more visits.
- Features: 95 linguistic (lexical diversity, psycholinguistic norms, syntax); 528 acoustic (eGeMAPS via OpenSMILE).
- Design: per-individual slopes; three analyses under nested 5-fold CV, first visit, last visit, trajectory.
- Selection: top features by importance within each training fold; five classifiers tested, highest-AUC reported.
- Cohort: 149 with two or more visits, 119 Control or AD; matched to 82 to 86 per analysis, 32 acoustic last visit.
- AUC with 1000-iteration bootstrap 95% CIs.

Results

Linguistic features classified well at every visit; acoustic separation was lower and trended upward, but not significantly.



Modality	Analysis	AUC [95% CI]	Best model	Matched N
Linguistic (95 NLP)	First visit	0.751 [0.656, 0.854]	SVM	86
	Longitudinal slopes	0.766 [0.652, 0.870]	SVM	84
	Last visit	0.819 [0.720, 0.904]	SVM	86
Acoustic (528 eGeMAPS)	First visit	0.534 [0.410, 0.666]	AdaBoost	86
	Longitudinal slopes	0.618 [0.490, 0.739]	SVM	82
	Last visit	0.719 [0.524, 0.889]	LogReg	32*
Multimodal (623)	Longitudinal slopes	0.626 [0.501, 0.754]	SVM	82

Differences are descriptive and confidence intervals overlap; only linguistic over acoustic at first visit reaches significance (approximate two-sided z-test, $p < 0.01$).

* Acoustic last-visit cohort reduced by matching; lower CI bound approaches chance. Five classifiers tested per analysis; highest-AUC reported.

Group	N	Age $\pm$ SD	Female	Visits	Span	Linguistic slope features	Acoustic slope features
Control	60	62.7 $\pm$ 7.0	63%	3.0	2.3	Type-token ratio variability $\uparrow$	Alpha-ratio variability $\downarrow$
ProbableAD	59	70.3 $\pm$ 8.5	71%	2.4	1.6	Sentiment (minimum) $\uparrow$	Spectral flux variability $\downarrow$
						Pronoun frequency $\downarrow$	Jitter (local) $\uparrow$
						Word meaningfulness $\downarrow$	F2 amplitude $\uparrow$
						Content density $\uparrow$	MFCC1 variability $\uparrow$

$\uparrow/\downarrow$ direction of change over time in probable AD

Language best distinguishes diagnoses. Acoustics may track change.
Different speech signals may serve detection and monitoring.

Discussion

- A good diagnostic marker is not always the best progression marker.
- Linguistic features index accumulated vocabulary and syntax damage: strong at baseline, largely saturated (0.751 to 0.766).
- Acoustic features weaker at baseline; point estimates higher at later stages (0.534 to 0.719) but not significant, may suit monitoring.
- Multimodal fusion did not beat the best single modality.
- Cookie Theft is language-loaded, favoring linguistic content.

Limitations

- One single-site, English corpus; clinical probable AD, no biomarker confirmation.
- Differences descriptive, CIs overlap; only linguistic over acoustic at baseline is significant.
- Best-of-five reporting optimistic; wide CIs, especially acoustic (528 features).
- Acoustic baseline and slopes straddle chance; last-visit cell $n = 32$.
- Short follow-up, shorter in AD; groups differ in age, so aging and disease are confounded.

Future directions

- Test the modality-by-timepoint interaction directly.
- Paired within-subject design; separate disease change from age drift; match on follow-up.
- Stronger selection: elastic net with stability selection, pre-reduce correlated acoustic features, fold model choice into CV.
- Replicate in biomarker-confirmed, multi-site cohorts with longer follow-up.

References

- Becker JT, Boller F, Lopez OL, Saxton J, McGonigle KL. The natural history of Alzheimer disease. Arch Neurol. 1994;51(6):585-594.
- Eyben F, Scherer KR, Schuller BW, et al. The Geneva Minimalistic Acoustic Parameter Set (eGeMAPS) for voice research and affective computing. IEEE Trans Affect Comput. 2016;7(2):190-202.

Acknowledgements

- Work was funded via NIH K23 AG 063900.
- Thanks to staff and faculty of OHSU, CU, and CSU.

