

Speech-Based Detection of Alzheimer's Dementia: Evaluating Acoustic, Temporal, and Complexity Biomarkers

Hamed Azami^{1,2,3}, Saturnino Luz⁴, and Sanjeev Kumar^{1,2,3}

¹ Adult Neurodevelopment and Geriatric Psychiatry Division, Centre for Addiction and Mental Health, Toronto, ON, Canada, ² Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada, ³ Toronto Dementia Research Alliance, Toronto, ON, Canada, ⁴ Usher Institute, Edinburgh Medical School, The University of Edinburgh, UK

Email: hamed.azami@camh.ca

Background

- ❖ Alzheimer's disease (AD) affects cognitive and motor control functions, altering speech characteristics.
- ❖ Speech-based biomarkers offer potential for non-invasive, scalable diagnosis and monitoring.
- ❖ Reliable acoustic features can enhance early detection and disease progression tracking.

Methods

• **Participants:** 237 individuals (115 Healthy Controls [HC], 110 AD), matched for age and gender via propensity scores.

- Mean age: HC = 66.4±6.6 years, AD = 69.4±6.9 years
- Cognitive assessment: Mean Mini Mental State Examination (MMSE) scores (HC = 29.0, AD = 17.9)

• **Task:** Participants described the standardized “Cookie Theft” picture

• **Feature Extraction:**

- ✓ **Rhythm and temporal dynamics:** articulation rate, PVI, pause statistics, tempo, and onset-rate measures
- ✓ **Spectral and timbre features:** MFCCs, spectral contrast, flux, entropy, spread, bandwidth, roll-off, and chroma
- ✓ **Voice quality and prosody**
- ✓ **Modulation and energy features**
- ✓ **Entropy-based complexity:** Dispersion Entropy (DispEn) and Multiscale DispEn (MDE; 20 scales)

“Cookie Theft” Picture Description Task



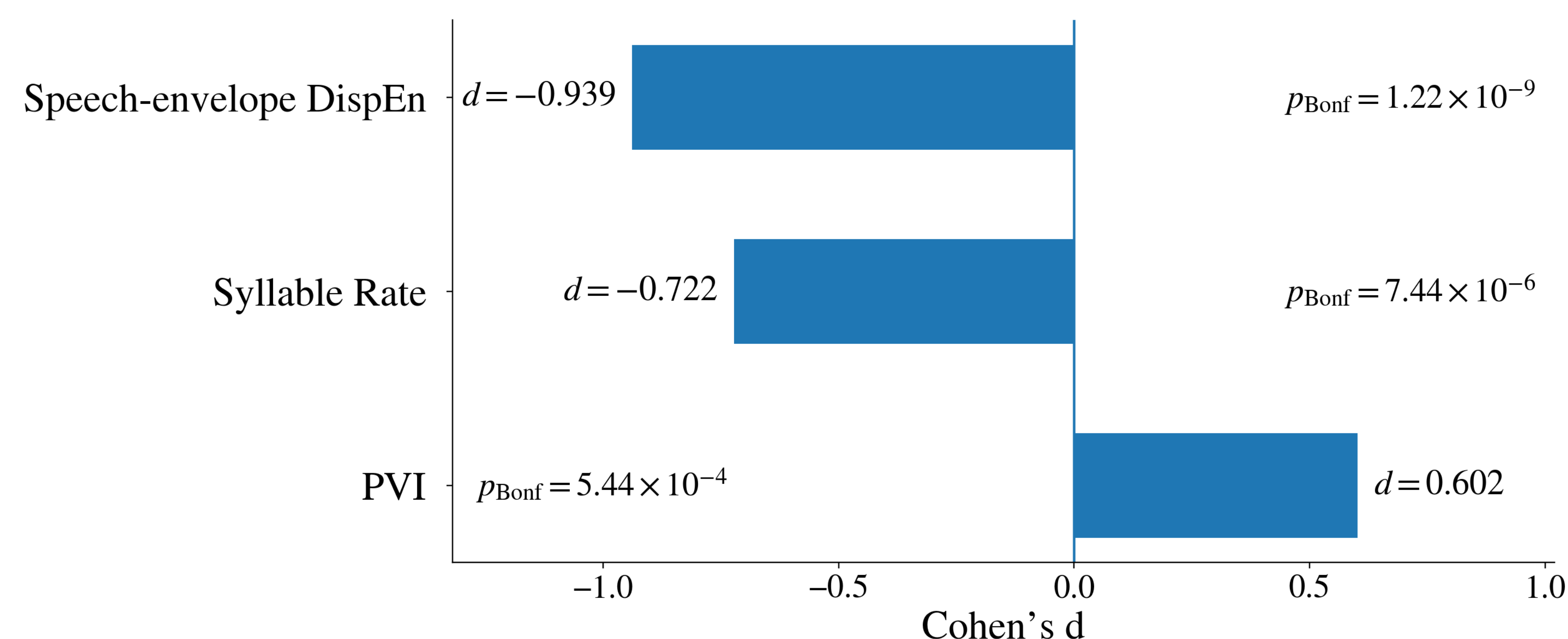
Methods

• **Statistical Analysis:**

- ✓ AD–HC group differences assessed using two-sample *t*-tests
- ✓ Bonferroni correction applied across 102 feature-wise tests
- ✓ Feature discriminability quantified using Cohen's *d*
- ✓ Features ranked using mutual information (*k* = 5–50)
- ✓ Classification performed using logistic regression with stratified 5-fold cross-validation
- ✓ Associations between selected acoustic features and MMSE assessed using Pearson correlation coefficients

Results

• **Top Three Discriminative Acoustic Features**



- Additional Bonferroni-significant markers included Chroma 1, Chroma 11–12, spectral contrast, and pitch mean.
- DispEn was associated with cognition (MMSE: $r = 0.37$, $p < 0.001$), with consistent associations across mid-range MDE scales ($r \approx 0.23$, $p < 0.001$).

Discussion

- Entropy and rhythm markers showed the strongest AD–HC differences, suggesting altered speech complexity and temporal organization in AD.
- DispEn was associated with MMSE, supporting its potential relevance as an interpretable marker of cognitive status.
- Combining complementary acoustic domains produced moderate classification performance (AUC = 0.80 ± 0.04), outperforming reliance on isolated speech characteristics.

Conclusions

- ❖ A multimodal acoustic ML pipeline integrating complexity, rhythm/pause, spectral, and voice-quality features provides interpretable discrimination between AD and HC.
- ❖ Future studies should validate and refine these markers in larger, independent cohorts and assess their added value alongside other AD biomarkers.

Reference

- F. Haider, S. de la Fuente and S. Luz, "An Assessment of Paralinguistic Acoustic Features for Detection of Alzheimer's Dementia in Spontaneous Speech," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 2, pp. 272-281, 2020.

Acknowledgments: We acknowledge the ADRess-M dataset providers and associated collaborators.