

Creating and validating a normative dataset of speech duration measures to support cognitive decline detection

Junhyuk “Louis” Kwak¹, Mark Y. Liberman¹, Laili Soleimani², Nadia Ramsingh², Carolyn W. Zhu³, Mary Sano², Sunghye Cho^{1,4}

¹Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA, USA

²Department of Psychiatry and ³Brookdale Department of Geriatrics and Palliative Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

⁴Department of Communication Sciences and Disorders, University of Cincinnati, Cincinnati, OH, USA

Introduction

- Picture description tasks have been used to gauge communication capabilities of neurodegenerative patients.
- Previous studies have demonstrated that speech duration measures can help distinguish between healthy individuals and those with dementia.
- It is crucial to establish the baseline of those measures in healthy individuals in a wide age range, as deviations could indicate neurodegeneration.
- We present a normative dataset of speech duration measures from picture description tasks performed by healthy volunteers to expedite cognitive decline detection.
- We validated the dataset of normalized scores with a different picture description dataset.

Method

- Participants: $N=282$, aged 20–90 years ($M=50.8$, $SD=17.6$)
- Materials: The Cookie Theft scene, the Picnic scene, and two similar scenes of our own design (Figure 2).
- Procedure
 - Data was collected online, where healthy individuals voluntarily participated in the study.
 - Participants completed 1–4 scenes ($M=2.6$, $SD=1.3$).
- Analysis
 - In-house speech activity detector for segmentation
 - Semi-parametric norming (cNORM package)
 - Age bins: [20, 30), [30, 40), ..., [60, 70), [70, 90]
 - Validation of norm scores
 - Internal: Cronbach's α , intraclass correlation coefficient
 - External: Picture description dataset from Mount Sinai
 - Healthy: $N=60$, age= 73.7 ± 6.2 yrs, edu= 16.5 ± 2.6 yrs
 - MCI: $N=30$, age= 75.9 ± 6.4 yrs, edu= 16.0 ± 2.7 yrs
 - AD: $N=16$, age= 74.4 ± 7.8 yrs, edu= 16.9 ± 2.3 yrs
 - A Welch's two-sample t-test to distinguish between cognitively normal and impaired groups

Results

Semi-parametric norming

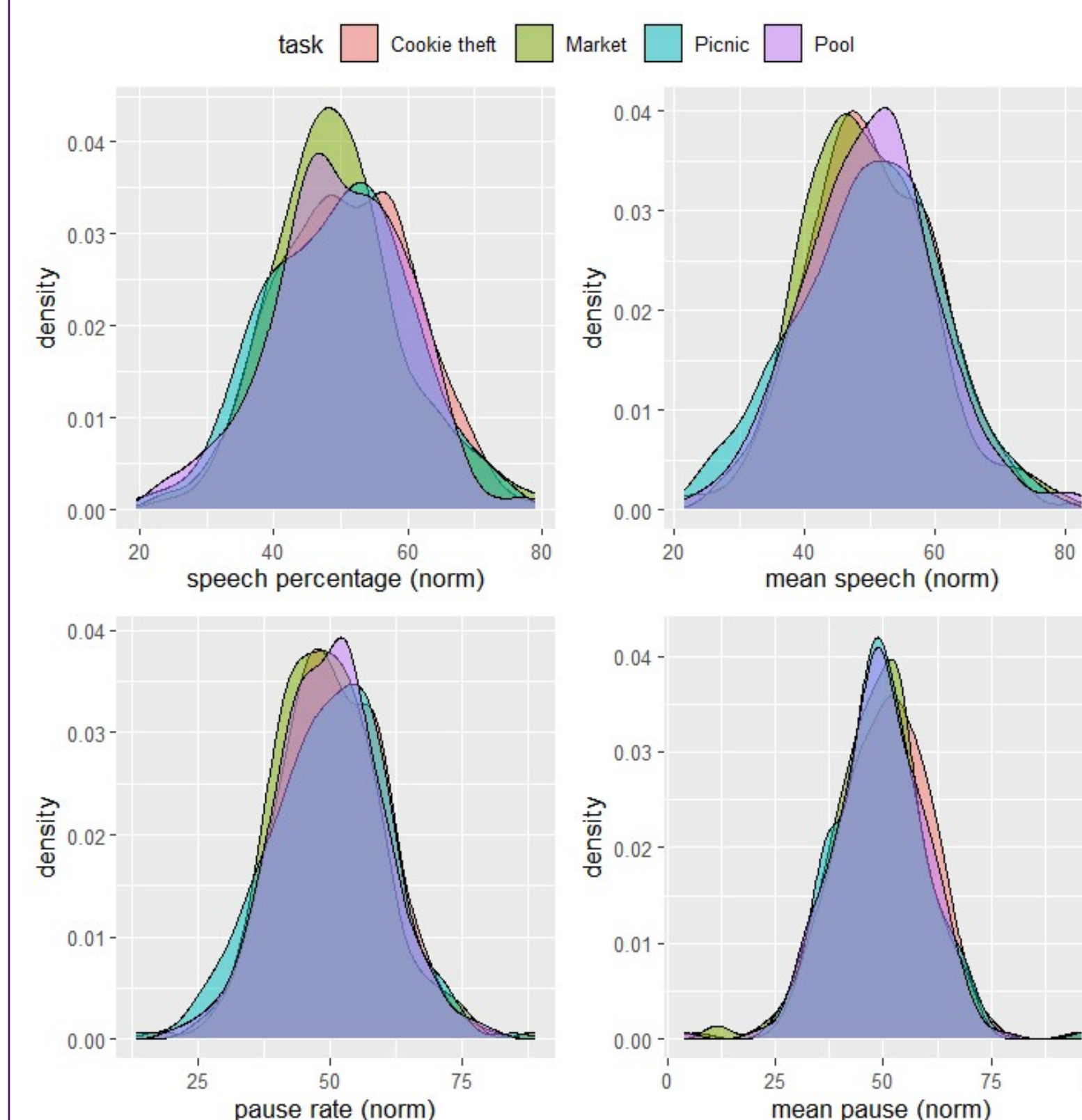


Figure 1: Density plot by task and measure. Each panel includes the density plots for one measure, where the plots for different tasks are overlaid on one another.

- For each measure, data from different tasks show similar density trends across their ranges.
- We combined data from different tasks as one group when proceeding with norming.



Figure 2: Two pictures customized for the study. The “Market” scene (left) and the “Pool” scene (right).

- For each measure, its norming model was selected such that:
 - $R_{adj}^2 \geq 0.99$;
 - It does not violate monotonicity;
 - Its number of predictors was minimal.
- Norming results provide norm scores of each measure for each age (Table 1).

Norm Score	Raw Score	Percentile
30	66.40 %	2.3
40	72.71 %	15.9
50	79.42 %	50.0
60	85.90 %	84.1
70	90.32 %	97.7

Table 1: Subset of speech percentage norm scores at age 65.

Normalized score validation

Internal (Consistency & Reliability Tests)

- Pause measure norm scores were reversed by $(100 - x)$ to be aligned with speech measure norm scores.
- Excluded measures due to negative correlation
 - Total speech length
 - $100 - \text{total pause length}$
- Cronbach's α
 - 4 items
 - Speech percentage, mean speech length
 - $100 - \text{pause rate}$, $100 - \text{mean pause length}$
 - $\alpha = .89$, $N = 732$
 - Good consistency
- Intraclass correlation coefficient
 - ICC(3,k) = 0.85
 - 95% CI = [0.87, 0.90]
 - $F(731, 2196) = 8.9$, $p < .001$
 - Good reliability

External

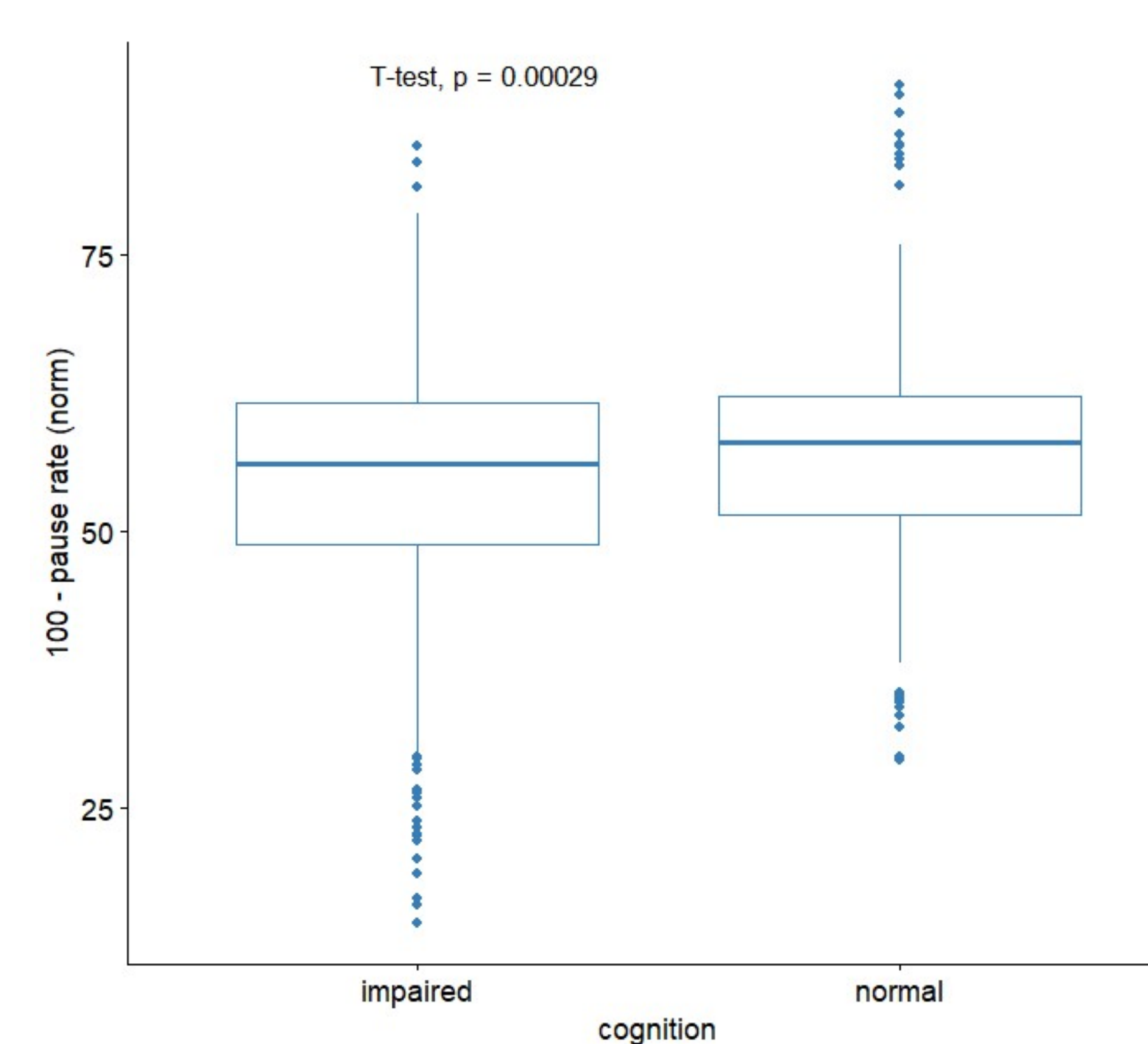


Figure 3: Boxplot for pause rate norm scores. Cognitively normal ($M=57.0$, $SD=9.8$) and impaired ($M=54.1$, $SD=11.8$) groups in the external validation dataset differed significantly ($t(656.33)=-3.6436$, $p<.001$) in their pause rate scores. With the score direction reversed ($100 - \text{pause rate}$), greater scores indicate less pauses.

Discussion

- We demonstrated the utility of age-normalized speech duration measures in cognitive decline detection.
- Combining the results from semi-parametric norming and validation, pause rate seems to be the most promising measure to detect abnormal speech patterns.
- These outcomes contribute to the applicative advancement of speech biomarker research for cognitive decline, calling for further studies on the potential of speech as a screening measure of dementia.

Limitations

- The study incorporated English data only, requiring cross-linguistic validation.
- The education levels of participants in the dataset for norming were high overall (17.8 ± 2.1 yrs), which should be further addressed in the norming and validation processes.
- Since we only used duration measures, future studies may expand this study with lexical and acoustic measures as potential speech biomarkers.

Acknowledgement

This study was funded by the Alzheimer's Association (AARF-21-851126).