

## RESEARCH ARTICLE

# Speech and language characteristics differentiate Alzheimer's disease and dementia with Lewy bodies

Yasunori Yamada<sup>1</sup>  | Kaoru Shinkawa<sup>1</sup>  | Miyuki Nemoto<sup>2</sup>  | Miho Ota<sup>2</sup>  |  
Kiyotaka Nemoto<sup>2</sup>  | Tetsuaki Arai<sup>2</sup> 

<sup>1</sup>Digital Health, IBM Research, Chuo-ku, Tokyo, Japan

<sup>2</sup>Department of Psychiatry, Division of Clinical Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan

## Correspondence

Yasunori Yamada, Digital Health, IBM Research, 19-21 Nihonbashi, Hakozaiki-cho, Chuo-ku, Tokyo 103-8510, Japan.  
Email: [ysnr@jp.ibm.com](mailto:ysnr@jp.ibm.com)

## Funding information

Japan Society for the Promotion of Science, KAKENHI, Grant/Award Number: 19H01084

## Abstract

**Introduction:** Early differential diagnosis of Alzheimer's disease (AD) and dementia with Lewy bodies (DLB) is important, but it remains challenging. Different profiles of speech and language impairments between AD and DLB have been suggested, but direct comparisons have not been investigated.

**Methods:** We collected speech responses from 121 older adults comprising AD, DLB, and cognitively normal (CN) groups and investigated their acoustic, prosodic, and linguistic features.

**Results:** The AD group showed larger differences from the CN group than the DLB group in linguistic features, while the DLB group showed larger differences in prosodic and acoustic features. Machine-learning classifiers using these speech features achieved 87.0% accuracy for AD versus CN, 93.2% for DLB versus CN, and 87.4% for AD versus DLB.

**Discussion:** Our findings indicate the discriminative differences in speech features in AD and DLB and the feasibility of using these features in combination as a screening tool for identifying/differentiating AD and DLB.

## KEYWORDS

acoustic, digital health, language impairment, linguistic, machine learning, natural language processing, prosodic, spontaneous speech

## 1 | INTRODUCTION

Alzheimer's disease (AD) and dementia with Lewy bodies (DLB) are the two most common types of late-onset neurodegenerative dementias.<sup>1,2</sup> Early and accurate differentiation between AD and DLB is important to ensure appropriate management and treatment of the disease,<sup>3,4</sup> but similarities in clinical manifestations often result in difficulties in clinical diagnosis.<sup>2,3</sup> Although diagnostic biomarkers in cerebrospinal fluid and neuroimaging are the most well-validated biomarkers,<sup>3,5,6</sup> they can be invasive, time-consuming, and expensive.

Hence, novel approaches for screening candidates who should be examined with the biomarkers would help the differential diagnosis of AD and DLB.

Speech may be a promising data source for developing screening tools and obtaining multifaceted information encompassing prosodic, acoustic, and linguistic characteristics.<sup>7-18</sup> In fact, discernable differences in speech features characterizing these aspects have been reported in patients with different dementia types.<sup>8-11</sup> For instance, speech and language disturbances have been observed in the early stages of AD<sup>8,12,13</sup> and may enable the prediction of its

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, LLC on behalf of Alzheimer's Association.

onset.<sup>14,18</sup> In particular, numerous studies have investigated spontaneous, connected speech during picture description tasks<sup>14–16</sup> and reported linguistic differences in AD patients such as increased repetition and reduced informative content related to word-finding difficulties.<sup>7,8,15,16</sup> In addition, speech and language impairments do not constitute a central feature for DLB,<sup>19</sup> but they are also observed in the course of DLB.<sup>9,17,20,21</sup> DLB patients showing discernable differences in language characteristics encompassing all linguistic aspects have been suggested, with the exception of the lexico-semantic domain, although these deficits have been suggested to be less severe compared with AD patients by study with a communication test evaluated by human experts.<sup>22</sup> In the speech of DLB patients, prosodic differences, such as slower speech rate and longer pauses, have been consistently observed,<sup>9,20,17,21</sup> and these differences are thought to be at least partially related to motor impairment.<sup>17</sup> Furthermore, although there is no study for DLB, acoustic impairments, such as increased variabilities in voice frequency and amplitude (i.e., jitter and shimmer), have been known to be significant features of Parkinson's disease (another form of Lewy body spectrum disorders).<sup>23–26</sup> By contrast, prosodic and acoustic differences in AD patients such as pauses and variability in voice frequency have yielded mixed results with several studies showing no significant difference and others showing statistically significant differences.<sup>27–29</sup> Together, these previous studies suggest different profiles of speech and language impairments between AD and DLB,<sup>17,22,30</sup> but there is no study directly comparing them using speech analysis.

In this study, we aimed to identify differences in features characterizing speech and language impairments between AD and DLB and to examine the feasibility of using these features to identify and differentiate AD and DLB. On the basis of previous studies, we hypothesized that AD and DLB patients would have different profiles of speech and language impairments. Specifically, we hypothesized that patients with AD would show larger differences in linguistic features compared with DLB patients, and patients with DLB would show larger differences in prosodic and acoustic features than AD patients. A second hypothesis was that these features would be used for reliably classifying patients with AD, those with DLB, and cognitively normal (CN) individuals. We collected speech responses with a tablet-based application from participants consisting of three clinical diagnostic groups: AD, DLB, and CN. We then extracted speech features characterizing acoustic, prosodic, and linguistic aspects. Finally, we tested the first hypothesis by statistically comparing the speech features between the diagnostic groups, and we tested the second hypothesis by assessing the performance of machine-learning models using these features to identify and differentiate AD and DLB.

## 2 | METHODS

### 2.1 | Participants

We recruited outpatients from the Department of Psychiatry, University of Tsukuba Hospital, along with the spouses of the patients,

### RESEARCH IN CONTEXT

- **Systematic Review:** We reviewed the literature using medical and academic databases (e.g., PubMed and Google Scholar) and cited relevant review articles. Speech analysis has succeeded in quantifying speech and language impairments in various neurodegenerative diseases including Alzheimer's disease (AD) and dementia with Lewy bodies (DLB). However, whether these features can differentiate AD and DLB has not been directly investigated.
- **Interpretation:** Our results provide initial evidence of (1) discriminative differences in linguistic, prosodic, and acoustic features that would reflect cognitive and motor impairments in AD and DLB, and (2) the feasibility of machine-learning models by using these features to identify and differentiate AD and DLB as an easy-to-perform screening tool.
- **Future Directions:** Future studies may aim to confirm our findings with larger samples and neuropathological biomarkers and investigate the associations of these speech features with other types of dementia as well as Lewy body disorders.

### HIGHLIGHTS

- AD and DLB showed different profiles of speech and language impairments
- AD showed more severe impairments in linguistic aspects compared with DLB
- DLB showed more severe impairments in prosodic/acoustic aspects compared with AD
- Combining these speech features successfully identified/differentiated AD and DLB

and other participants either through local recruiting agencies or community advertisements in Ibaraki, Japan. The patients met the standard research diagnostic criteria for mild cognitive impairment (MCI)/dementia due to AD or Lewy body disease. Specifically, the patients in the AD group fulfilled the National Institute on Aging and Alzheimer's Association core clinical criteria for probable AD dementia<sup>5</sup> or MCI,<sup>31</sup> as well as the AD Neuroimaging Initiative criteria for AD or MCI.<sup>32</sup> The patients in the DLB group fulfilled McKeith et al.'s clinical diagnostic criteria for probable/possible DLB<sup>3</sup> or MCI with Lewy bodies.<sup>33</sup> Therefore, our samples were clinically diagnosed, though their diagnoses were not confirmed by biomarker or post-mortem examination. The CN participants were age-matched to the patients and did not meet any of the aforementioned criteria. Participants were excluded if they had diagnoses of other types of dementia

**TABLE 1** Participant demographics and cognitive/clinical measures

	CN (n = 49)		AD (n = 45)		DLB (n = 27)		BF	P value
Age, years	72.3	(3.8)	73.1	(6.7)	75.1	(5.0)	0.519	0.103
Sex, female, n (%)	31	(63.3)	20	(44.4)	12	(44.4)	0.517	0.126
Education, years	13.1	(2.0)	13.1	(2.7)	12.7	(2.8)	0.104	0.735
MCI, n (%)	N/A		25	(55.6)	19	(70.4)	0.608	0.318
Antipsychotic medication, n (%)	0	(0.0)	<sup>D</sup> 3	(6.8)	<sup>b</sup> 6	(23.1)	<sup>C, b</sup> 5.12	<b>0.002</b>
Mini-Mental State Examination <sup>a</sup>	28.0	(1.6)	<sup>A</sup> 23.5	(4.3)	<sup>D,C</sup> 26.5	(3.6)	<sup>A</sup> 1.82 × 10 <sup>6</sup>	<b>&lt;0.001</b>
Frontal Assessment Battery <sup>a</sup>	13.6	(2.5)	<sup>A,D</sup> 11.0	(3.7)	<sup>C</sup> 11.1	(4.1)	<sup>C</sup> 79.1	<b>&lt;0.001</b>
Logical Memory-immediate <sup>a</sup>	11.1	(3.3)	<sup>A,D</sup> 4.4	(3.5)	<sup>D,C</sup> 7.9	(3.9)	<sup>A,C</sup> 7.85 × 10 <sup>11</sup>	<b>&lt;0.001</b>
Logical Memory-delayed <sup>a</sup>	9.2	(3.0)	<sup>A,D</sup> 2.2	(2.6)	<sup>D,C</sup> 5.9	(4.1)	<sup>A,C</sup> 2.46 × 10 <sup>15</sup>	<b>&lt;0.001</b>
Trail Making Test part-A <sup>a</sup>	36.0	(11.7)	<sup>D</sup> 53.4	(46.6)	68.5	(54.0)	<sup>C</sup> 13.69	<b>0.002</b>
Trail Making Test part-B <sup>a</sup>	91.6	(39.9)	<sup>A,D</sup> 168.1	(86.4)	<sup>C</sup> 183.7	(86.8)	<sup>C</sup> 2.56 × 10 <sup>5</sup>	<b>&lt;0.001</b>
Clock Drawing Test <sup>a</sup>	6.7	(0.8)	6.1	(1.8)	6.6	(1.1)	0.872	0.063
Clinical Dementia Rating	0	(0.0)	<sup>A,D</sup> 0.7	(0.3)	<sup>C</sup> 0.6	(0.4)	<sup>C</sup> 9.21 × 10 <sup>19</sup>	<b>&lt;0.001</b>
Geriatric Depression Scale <sup>a</sup>	3.2	(3.0)	3.6	(3.1)	4.1	(3.9)	0.147	0.497
Activities of Daily Living <sup>a</sup>	99.7	(1.2)	<sup>A</sup> 98.7	(4.2)	<sup>D,C</sup> 97.2	(5.9)	<sup>A</sup> 1.47	<b>0.031</b>
Instrumental Activities of Daily Living <sup>a</sup>	7.8	(0.6)	<sup>D</sup> 6.7	(1.6)	6.1	(2.1)	<sup>C</sup> 6.65 × 10 <sup>3</sup>	<b>&lt;0.001</b>
Medial temporal lobe atrophy	0.8	(0.5)	<sup>A</sup> 1.6	(1.0)	<sup>D,C</sup> 1.1	(0.6)	<sup>A</sup> 2.52 × 10 <sup>3</sup>	<b>&lt;0.001</b>

Data are displayed as means (standard deviations), except for categorical data, which are displayed as numbers (percentages). Bold values highlight statistically significant differences (chi-square test,  $P < 0.05$ , for categorical data; one-way ANOVA,  $P < 0.05$ , for continuous data). Significant differences between individual diagnostic groups (chi-square test,  $P < 0.05$ , for categorical data; Tukey-Kramer test,  $P < 0.05$ , for continuous data) are marked with A, D, or C (A: different from AD; D: different from DLB; C: different from CN). Logical Memory-immediate and Logical Memory-delayed refer to immediate and delayed recall of Logical Memory Story A from the Wechsler Memory Scale-Revised, respectively.

<sup>a</sup>The total score ranges are as follows: Mini-Mental State Examination, 0 to 30; Frontal Assessment Battery, 0 to 18; Logical Memory (immediate and delayed), 0 to 25; Trail Making Test (parts A and B), 0 to 300; Clock Drawing Test, 0 to 7; Geriatric Depression Scale, 0 to 15; Activities of Daily Living, 0 to 100; Instrumental Activities of Daily Living, 0 to 8.

<sup>b</sup>Data missing for one participant.

Abbreviations: CN, cognitively normal; AD, Alzheimer's disease; DLB, dementia with Lewy bodies; ANOVA, analysis of variance; BF, Bayes factor; MCI, mild cognitive impairment.

(e.g., frontotemporal dementia or vascular dementia) or other serious diseases or disabilities that would interfere with the collection of speech data. Thus, the participants formed three clinical diagnostic groups of cognitive impairment due to AD (AD group) or Lewy body disease (DLB group), and cognitively normal controls (CN group). Patients in the AD and DLB groups ranged from MCI to moderate dementia.<sup>34</sup> The participants were administered cognitive and clinical examinations, which comprised 12 variables (see Table 1 for a full list and the Supplementary Methods in the Supporting Information for imaging details). Three psychiatrists (authors TA, KN, and MO), who are experts in dementia and were blind to the results of the speech data analysis, examined each case in terms of the clinical record, as well as the cognitive and clinical measures, and they confirmed the diagnoses.

The study was conducted with the approval of the Ethics Committee, University of Tsukuba Hospital (H29-065), and it followed the ethical code for research with humans as stated in the Declaration of Helsinki. All participants provided written informed consent to participate in the study. All examinations were conducted in Japanese. Experiment periods are described in the Supplementary Methods.

## 2.2 | Speech data collection and speech features

The participants sat in front of an iPad Air 2 tablet and answered questions presented by a voice-based application on the tablet in a quiet room with low reverberation in a lab setting (for more details, see the Supplementary Methods). The participants performed five speech tasks: counting backwards, subtraction, tasks for phonemic and semantic verbal fluency, and picture description with the Cookie Theft picture from the Boston Diagnostic Aphasia Examination.<sup>35</sup> The reasons for selecting these five tasks are described in the Supplementary Materials. Full descriptions of the five speech tasks are provided in Table S1.

From the participants' speech responses to the five tasks, we extracted a total of 42 speech features for each participant, which consisted of 11 prosodic features, 22 acoustic features, and 9 linguistic features on the basis of previous studies on AD, DLB, and Parkinson's disease.<sup>8,9,16,24,28,36</sup> They were extracted from each task response and investigated separately in subsequent analyses. A full list of speech features and tasks from which these features were extracted is given in Table S2. The acoustic and prosodic features were

extracted from the speech responses to all five tasks except where otherwise indicated. Specifically, prosodic features included a proportion of the pause duration and pitch variation (i.e., inflection) in addition to the phoneme rate in the picture description task. Acoustic features included jitter and shimmer, which measure cycle-to-cycle variations of the fundamental frequency and amplitude, respectively. Increased shimmer and jitter have been reported in patients of neurodegenerative diseases.<sup>23–26,37</sup> Furthermore, we used the variances of first-order derivatives of the first 12 Mel-frequency cepstral coefficients (MFCCs) during the picture description task as acoustic features for building classifiers. We excluded the MFCC features from statistical comparisons between the three diagnostic groups due to the difficulty in interpreting their differences. Linguistic features were extracted from manually transcribed text data. In addition to the number of correct answers during the verbal fluency tasks, the following linguistic features were extracted from response data during the picture description task: the number of filler words, type-token ratio of nouns for measuring vocabulary richness,<sup>16,36</sup> and five features for measuring informative content.<sup>7,8,15,16</sup> A higher value of the type-token ratio indicates greater vocabulary richness. As for the five features for measuring informative content, we counted the number of unique entities that a participant described in the picture, referred to as information units, for four predefined categories (people, places, objects, and actions), and used the number of information units for each category and the total number of information units as linguistic features after dividing by the speech duration. Details of the calculation methods including preprocessing are described in the Supplementary Methods. In sum, we used all 42 speech features for building classifiers and 30 features (all speech features except for the 12 MFCC-based features) for statistical comparisons.

### 2.3 | Statistical analysis

Group differences between CN, AD, and DLB were examined by using the chi-square test followed by a post hoc chi-square test for categorical data and one-way analysis of variance (ANOVA) followed by a post hoc Tukey-Kramer test for continuous data. Sex, disease stage (MCI, dementia), and the use of antipsychotic medication are categorical data, and the other data including all speech features are continuous data. For the multiple testing of the 30 speech features, the Benjamini-Hochberg correction was applied. We calculated a Bayes factor to assess the magnitude of evidence in favor of an alternative hypothesis versus the null hypothesis.<sup>38</sup> We also calculated the generalized eta-squared ( $\eta^2$ ) to assess the effect size of each feature, for which the values 0.01, 0.06, and 0.14 are considered to indicate small, medium, and large effects, respectively.<sup>39</sup> Between-group comparisons of speech features after controlling the use of antipsychotic medication were conducted with one-way analyses of covariance (ANCOVAs). A two-way ANOVA was used to examine the effects of the dementia type (AD, DLB) and disease stage (MCI, dementia) on the speech features. All statistical analyses were performed using R (version 4.0.5) with an alpha value of 0.05 ( $P < 0.05$ , two-sided).

### 2.4 | Machine-learning analysis

To evaluate the feasibility of using speech features to identify and differentiate AD and DLB, we used supervised machine-learning models to classify the clinical diagnostic groups of AD, DLB, and CN via the speech features. The input variables for the model were the 42 speech features. Binary-classification models were evaluated by using accuracy, sensitivity, specificity, F1 score, and the area under receiver operating characteristic curve (AUC) obtained from 10 iterations of a leave-two-subjects-out cross-validation procedure. To reduce overfitting through automatic feature selection, we also performed a sequential forward selection algorithm. For the classification algorithm, we used a support vector machine with a radial basis function kernel<sup>40</sup> implemented with the Python package scikit-learn (version 0.23.2). The hyperparameters in this study have been found in a prior study.<sup>41</sup> For missing values, we applied multivariate imputation by chained equations (version 3.13.3).<sup>42</sup> Spearman's rank correlation coefficient was used to assess the correlations between cognitive scores and output probability measures of the binary-classification models.

## 3 | RESULTS

A total of 123 participants met the inclusion criteria. Among the 123 participants, two were excluded from this analysis because they did not complete any of the five speech tasks. This yielded a total of 121 participants comprising three diagnostic groups of 45 AD, 27 DLB, and 49 CN participants (Table 1; S3 for additional clinical information on the DLB group; Supplementary Results for a power analysis of the sample size). The AD and DLB groups included 25 and 19 MCI patients, respectively, and their proportions were not statistically significantly different ( $P = 0.318$ ). Regarding the demographics, neither the age, proportion of female participants, nor years of education showed any statistically significant differences among the groups ( $P > 0.05$ ). The proportion of participants on antipsychotic medication was higher for the DLB group than for the CN group ( $P = 0.002$  among the three groups;  $P = 0.003$  for DLB vs. CN). All 12 cognitive and clinical measures except for the Clock Drawing Test and Geriatric Depression Scale were different among the diagnostic groups (all  $P < 0.05$ ; Table 1). Detailed information about missing values for the speech data are reported in the Supplementary Results.

We investigated whether the speech features had statistically discernable differences between the clinical diagnostic groups of AD, DLB, and CN. We found that nine speech features showed statistically significant differences between the groups (Benjamini-Hochberg adjusted  $P < 0.05$ ; Table 2). These nine features all showed a medium effect size ( $\eta^2 > 0.06$ ).<sup>39</sup>

Post hoc pairwise comparisons with the Tukey-Kramer test revealed the following patterns of statistically significant differences between the CN groups and the AD or DLB groups (Table 2 and Figure 1). Regarding the overall trends, the AD group demonstrated larger differences from the CN group in linguistic features, while the DLB group demonstrated larger differences in acoustic and prosodic features.

**TABLE 2** Speech features with statistically significant differences between the clinical diagnostic groups

Feature name	CN (n = 49)		AD (n = 45)			DLB (n = 27)			P value			
	Mean	SD	Significance	Mean	SD	Significance	Mean	SD	$\eta^2$	UNADJ	ADJ	
<b>Linguistic</b>												
Number of correct answers (words) (Semantic VFT)	17.7	(5.3)	A,D	13.8	(5.1)	C	14.6	(5.4)	C	0.108	<b>0.001</b>	<b>0.018</b>
Type-token ratio (Picture description)	0.81	(0.11)	A	0.74	(0.14)	C	0.78	(0.15)		0.054	<b>0.039</b>	0.099
Total number of information units (/s) (Picture description)	0.243	(0.110)	A	0.190	(0.079)	C	0.201	(0.098)		0.062	<b>0.024</b>	0.071
Number of information units in the action category (/s) (Picture description)	0.082	(0.049)	A	0.057	(0.026)	C	0.064	(0.038)		0.077	<b>0.009</b>	<b>0.047</b>
Number of information units in the place category (/s) (Picture description)	0.015	(0.014)	A	0.009	(0.009)	C	0.014	(0.011)		0.052	<b>0.043</b>	0.099
Number of information units in the object category (/s) (Picture description)	0.106	(0.049)	A	0.082	(0.044)	C	0.089	(0.044)		0.051	<b>0.046</b>	0.099
<b>Prosodic</b>												
Proportion of pause duration (Picture description)	0.386	(0.129)	D	0.446	(0.143)		0.507	(0.130)	C	0.110	<b>0.001</b>	<b>0.018</b>
Proportion of pause duration (Semantic VFT)	0.752	(0.074)	A,D	0.792	(0.079)	C	0.802	(0.069)	C	0.078	<b>0.008</b>	<b>0.047</b>
Proportion of pause duration (Subtraction)	0.339	(0.133)	D	0.396	(0.155)		0.433	(0.177)	C	0.059	<b>0.032</b>	0.088
Pitch variation (Subtraction)	25.5	(9.5)	D	22.5	(9.7)		18.8	(8.2)	C	0.072	<b>0.014</b>	<b>0.047</b>
Phoneme rate (/s) (Picture description)	2.37	(0.57)	D	2.11	(0.63)		1.92	(0.59)	C	0.085	<b>0.006</b>	<b>0.047</b>
<b>Acoustic</b>												
Jitter (Semantic VFT)	0.067	(0.010)	D	0.068	(0.009)	D	0.074	(0.011)	AC	0.072	<b>0.012</b>	<b>0.047</b>
Jitter (Picture description)	0.066	(0.010)	D	0.068	(0.011)		0.074	(0.011)	C	0.070	<b>0.014</b>	<b>0.047</b>
Shimmer (Picture description)	0.117	(0.017)	D	0.120	(0.019)	D	0.130	(0.019)	AC	0.075	<b>0.011</b>	<b>0.047</b>

Data are displayed as means (standard deviations). Bold values highlight statistically significant differences (one-way ANOVA,  $P < 0.05$ ). Significant differences between individual diagnostic groups (Tukey-Kramer test,  $P < 0.05$ ) are marked with A, D, or C (A: different from AD; D: different from DLB; C: different from CN).

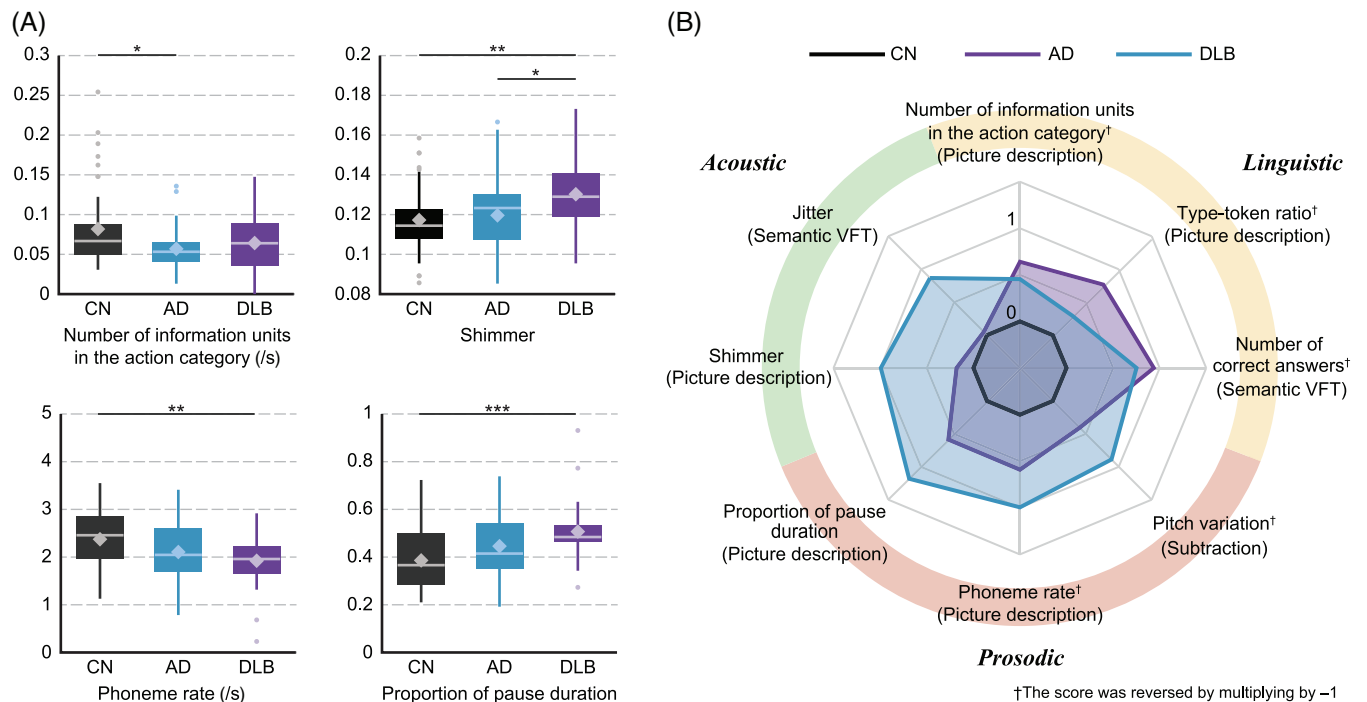
Abbreviations: CN, cognitively normal; AD, Alzheimer's disease; DLB, dementia with Lewy bodies; UNADJ, unadjusted; ADJ, Benjamini-Hochberg adjusted; VFT, verbal fluency task.

Specifically, the speech of the AD group showed significant differences in the linguistic features of reduced information units in the action category during the picture description task ( $P = 0.047$ ) and reduced correct answers of the semantic verbal fluency task ( $P = 0.018$ ). By contrast, the speech of the DLB group showed significant differences in the prosodic features of a slower phoneme rate ( $P = 0.047$ ), less pitch variation in the subtraction task ( $P = 0.047$ ), and increased proportion of pause duration in the semantic verbal fluency and picture description tasks ( $P = 0.047$  and  $P = 0.018$ , respectively); DLB speech was also significantly different in the acoustic features of increased jitter in the semantic verbal fluency and picture description tasks ( $P = 0.047$  and  $P = 0.047$ , respectively) and increased shimmer in the picture description task ( $P = 0.047$ ). After controlling for the use of antipsychotic medication, these speech features showed consistent trends: Seven

of the nine speech features remained significantly different between the groups (Benjamini-Hochberg adjusted  $P < 0.05$ ; Table S4), and the AD group showed significant differences from the CN group in linguistic features while the DLB group showed significant differences in prosodic and acoustic features (post hoc pairwise comparisons with the Tukey-Kramer test  $P < 0.05$ ; Table S4).

The ANOVA post hoc Tukey-Kramer tests also revealed that the AD and DLB groups had significant differences in two acoustic features: jitter in the semantic verbal fluency ( $P = 0.026$ ) and shimmer in the picture description tasks ( $P = 0.045$ ) (Table 2). We further conducted a two-way ANOVA, 2 dementia types (AD, DLB)  $\times$  2 disease stages (MCI, dementia), for these two acoustic features. The results showed significant effects of the dementia type, in which both jitter and shimmer were higher in the DLB group compared with the AD group (Figure 2):





**FIGURE 1** Differences in the linguistic, acoustic, and prosodic features between three clinical diagnostic groups: Alzheimer's disease (AD) patients, dementia with Lewy bodies (DLB) patients, and cognitively normal (CN) individuals. (A) Graphs of linguistic (upper left), acoustic (upper right), and prosodic (lower left and right) features. Boxes indicate the 25th (Q1) and 75th (Q3) percentiles. Whiskers indicate the upper and lower adjacent values that are most extreme within  $Q3+1.5(Q3-Q1)$  and  $Q1-1.5(Q3-Q1)$ , respectively. The line and diamond in each box represent the median and mean, respectively. Dots outside of the box represent outliers. Horizontal bars indicate significant differences (Tukey-Kramer test:  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ). (B) Radar chart comparing the linguistic, acoustic, and prosodic features of the AD, DLB, and CN groups, scaled according to Z-scores derived from the CN group's means and standard deviations.

**TABLE 3** Performance of classification models using speech features

	Mean [95% CI]				
	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)	F1 score (%)
AD vs. CN	87.0 [86.5, 87.5]	0.851 [0.849, 0.853]	86.4 [85.9, 86.9]	87.6 [86.7, 88.4]	86.4 [86.0, 86.9]
DLB vs. CN	93.2 [92.8, 93.6]	0.934 [0.931, 0.937]	88.9 [88.9, 88.9]	95.5 [94.9, 96.1]	90.2 [89.7, 90.7]
AD vs. DLB	87.4 [86.4, 88.3]	0.833 [0.830, 0.836]	N/A	N/A	N/A

Abbreviations: CI, confidence interval; AUC, the area under receiver operating characteristic curve; AD, Alzheimer's disease; DLB, dementia with Lewy bodies; CN, cognitively normal.

A two-way ANOVA of the jitter revealed a significant effect of the dementia type ( $P = 0.014$ ), no significant effect of the disease stage, and no interaction; that of the shimmer revealed significant effects of the dementia type ( $P = 0.014$ ) and disease stage ( $P = 0.030$ ), and no interaction.

We evaluated the model performance by using iterative cross-validation for classifying the three clinical diagnostic groups of AD, DLB, and CN on the basis of speech features (Table 3). The binary-classification models achieved 87.0% accuracy (86.4% sensitivity, 87.6% specificity, 86.4% F1 score, AUC of 0.851) for AD vs. CN; 93.2% accuracy (88.9% sensitivity, 95.5% specificity, 90.2% F1 score, AUC of 0.934) for DLB versus CN; and 87.4% accuracy (AUC of 0.833) for AD

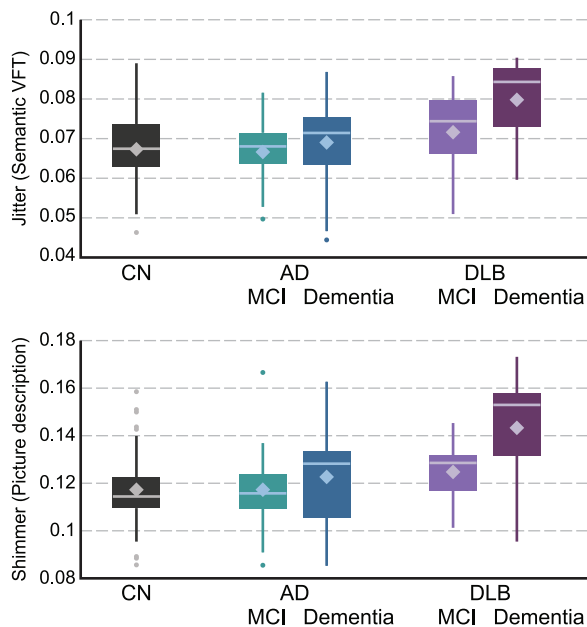
versus DLB (Table 3). The three-class classification model for AD, DLB, and CN groups achieved 79.9% accuracy (AUC of 0.873).

To better understand how the models could classify clinical diagnostic groups, we investigated the association between cognitive scores and output measures of each binary-classification model using the Spearman's rank correlation coefficient ( $\rho$ ). Consequently, the model output measures showed the highest correlation with the Logical Memory-delayed scores ( $|\rho| = 0.47$ ,  $P < 0.0001$ ; Table 4) in the model for AD versus CN, with the Trail Making Test part-B scores ( $|\rho| = 0.51$ ,  $P < 0.0001$ ; Table 4) in the model for DLB versus CN, and with the Trail Making Test part-A scores ( $|\rho| = 0.46$ ,  $P < 0.0001$ ; Table 4) in the model for AD versus DLB.

**TABLE 4** The Spearman's rank correlation coefficient ( $\rho$ ) of cognitive scores and output measures of binary-classification models of AD, DLB, and CN groups, ordered by the absolute correlation coefficients

AD vs. CN			DLB vs. CN			AD vs. DLB		
Cognitive measure	$ \rho $	P value	Cognitive measure	$ \rho $	P value	Cognitive measure	$ \rho $	P value
Logical Memory-delayed	0.47	<0.001	Trail Making Test part-B	0.51	<0.001	Trail Making Test part-A	0.46	<0.001
Mini-Mental State Examination	0.44	<0.001	Trail Making Test part-A	0.37	0.001	Trail Making Test part-B	0.27	0.026
Logical Memory-immediate	0.43	<0.001	Logical Memory-immediate	0.33	0.003	Logical Memory-delayed	0.19	0.105
Trail Making Test part-B	0.33	0.001	Logical Memory-delayed	0.32	0.005	Frontal Assessment Battery	0.12	0.296
Trail Making Test part-A	0.25	0.014	Frontal Assessment Battery	0.21	0.062	Logical Memory-immediate	0.09	0.461
Frontal Assessment Battery	0.24	0.019	Mini-Mental State Examination	0.13	0.275	Mini-Mental State Examination	0.03	0.797
Clock Drawing Test	0.19	0.075	Clock Drawing Test	0.11	0.350	Clock Drawing Test	0.01	0.945

Abbreviations: AD, Alzheimer's disease; DLB, dementia with Lewy bodies; CN, cognitively normal.

**FIGURE 2** Differences in acoustic features (upper: jitter; lower: shimmer) based on clinical diagnosis (cognitively normal [CN], Alzheimer's disease [AD], dementia with Lewy bodies [DLB]) and disease stages (mild cognitive impairment [MCI], dementia). Boxes indicate the 25th (Q1) and 75th (Q3) percentiles. Whiskers indicate the upper and lower adjacent values that are most extreme within  $Q3+1.5(Q3-Q1)$  and  $Q1-1.5(Q3-Q1)$ , respectively. The line and diamond in each box represent the median and mean, respectively. Dots outside of the box represent outliers.

## 4 | DISCUSSION

We investigated speech features characterizing linguistic, prosodic, and acoustic aspects by using data collected from 121 participants in the AD, DLB, and CN groups, and obtained two main findings. First, a statistical analysis showed that the AD group showed larger differences from the CN group than the DLB group in linguistic features, with reductions in informative contents and semantic verbal fluency, while the DLB group showed larger differences in prosodic and acoustic features, with reduced phoneme rate and increased pause proportion, less

inflections (i.e., monotony and dullness of speech in clinical descriptions), and increased variabilities in voice frequency and amplitude. Second, the combination of these speech features could identify and differentiate AD and DLB by capturing impairments in different cognitive measures: Logical Memory-delayed scores for AD versus CN, Trail Making Test part-B scores for DLB versus CN, and Trail Making Test part-A scores for AD versus DLB. To the best of our knowledge, this is the first study to identify discriminative patterns in speech and language impairments of AD and DLB and to demonstrate the feasibility of using them to identify and differentiate AD and DLB.

The profile of language and speech impairments were different in the AD and DLB groups, with larger differences from the CN group in the linguistic features of the AD group and in the prosodic and acoustic features of the DLB group. The trends in these differences in each dementia type were consistent with those reported in previous studies.<sup>8,16,24,28,36</sup> Hence, one of our contributions lies in providing empirical evidence of how these differences differ in AD and DLB via direct comparisons using speech analysis. Furthermore, to the best of our knowledge, this is the first study to suggest the usefulness of shimmer and jitter to differentiate DLB from CN and AD. Numerous studies on speech of Parkinson's disease patients have shown increases in shimmer and jitter,<sup>23-26</sup> which were also suggested to further increase as the disease progresses.<sup>26</sup> These differences are thought to be due in part to motor impairments including deteriorated control of respiratory and laryngeal muscles.<sup>43</sup> Comparing these acoustic features across Lewy body disorders may provide useful insights into neuropathological mechanisms underlying the speech impairments.<sup>23</sup>

The results of models using speech features could achieve a high performance for identifying and differentiating the AD and DLB groups. Our models differentiated the AD and DLB groups from the CN group primarily by capturing impairments in Logical Memory-delayed scores and Trail Making Test part-B scores, respectively. While AD is characterized clinically by prominent memory impairment,<sup>19</sup> DLB is characterized by more executive, attentional, and visuospatial impairment relative to memory impairment,<sup>19</sup> which may support the relevance of our results. Aligning with studies on gait and balance analysis for differentiating dementia types,<sup>44-46</sup> our results suggest the possibility of using speech and language characteristics as

behavioral markers, reflecting dementia-type-specific cognitive profiles and underlying pathologies. To confirm this suggestion, we will need further study with validated neuropathological biomarkers. From a clinical perspective, a classification model using speech analysis may assist clinicians in differential diagnosis as a screening tool because speech data can be acquired in routine clinical practice. In fact, a number of studies on speech analysis for data collected during neuropsychological assessment have succeeded in detecting MCI and dementia,<sup>14,16,47–49</sup> and several of them have shown that the addition of speech analysis to the neuropsychological assessment has been found to improve its accuracy.<sup>48,49</sup> Another clinical implication of this study is that our findings using the tablet-based application might help develop a self-administrated tool for the early detection of AD and DLB. According to the World Alzheimer Report published in 2021,<sup>50</sup> 83% of clinicians maintain that the COVID-19 pandemic has delayed access to diagnostic assessments. Thus, a self-administered screening tool may be especially important in the current COVID-19 pandemic given the difficulties of an in-person evaluation in a clinical setting. In our future research, we will endeavor to investigate the operability and acceptability of real-world data collection.

There were several limitations in this study. First, our analysis did not include neuropathological changes using validated biomarkers or postmortem follow-up. A significant proportion of dementia patients have mixed Lewy body and AD pathological changes.<sup>2</sup> Because we diagnosed DLB with clinical diagnostic criteria, we may have included DLB patients with mixed pathology. The association of speech and language characteristics with neuropathological changes warrants further research. Second, our dataset was a relatively small sample size and was imbalanced among the clinical diagnostic groups. The generality of our findings should be confirmed with larger samples collected in multiple sites. Third, residual confounding can still exist, and matching would have strengthened our findings. The DLB group had higher Mini-Mental State Examination scores than the AD group. Finally, we collected the speech data in a lab setting, and the controlled setting might have influenced how the participants responded to questions.

In summary, our results provide initial, empirical evidence of (1) the discriminative differences in speech features characterizing acoustic, prosodic, and linguistic aspects in AD and DLB, and (2) the feasibility of using these speech features as a screening tool for identifying and differentiating AD and DLB. Our findings require further validation with neuropathological biomarkers or postmortem follow-up.

## ACKNOWLEDGMENTS

This work was supported by the Japan Society for the Promotion of Science, KAKENHI (grant number 19H01084). The funders did not play any active role in the scientific investigation and reporting of the study.

## CONFLICT OF INTEREST

YY is employed by the IBM Corporation.

KS is employed by the IBM Corporation.

MN received funding from the Japan Society for the Promotion of Science, KAKENHI (grant number 19H01084).

MO has nothing to disclose.

KN has nothing to disclose.

TA received funding from Japan Society for the Promotion of Science, KAKENHI (grant number 19H01084). TA reports honoraria for lectures from Eisai, Daiichi-Sankyo, and Sumitomo Pharma.

Author disclosures are available in the [Supporting Information](#).

## ORCID

Yasunori Yamada  <https://orcid.org/0000-0003-3333-4368>

Kaoru Shinkawa  <https://orcid.org/0000-0002-3869-1922>

Miyuki Nemoto  <https://orcid.org/0000-0002-1969-5098>

Miho Ota  <https://orcid.org/0000-0001-8873-0822>

Kiyotaka Nemoto  <https://orcid.org/0000-0001-8623-9829>

Tetsuaki Arai  <https://orcid.org/0000-0003-4355-3900>

## REFERENCES

- Savica R, Boeve BF, Logroscino G. Epidemiology of alpha-synucleinopathies: from Parkinson disease to dementia with Lewy bodies. *Handb Clin Neurol*. 2016;138:153-158. <https://doi.org/10.1016/B978-0-12-802973-2.00009-4>
- Chatterjee A, Hirsch-Reinshagen V, Moussavi SA, et al. Clinico-pathological comparison of patients with autopsy-confirmed Alzheimer's disease, dementia with Lewy bodies, and mixed pathology. *Alzheimers Dement (Amst)*. 2021;13(1):e12189. <https://doi.org/10.1002/dad2.12189>
- McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with Lewy bodies: fourth consensus report of the DLB Consortium. *Neurology*. 2017;89(1):88-100. <https://doi.org/10.1212/WNL.0000000000004058>
- Rasmussen J, Langerman H. Alzheimer's disease—why we need early diagnosis. *Degener Neurol Neuromuscul Dis*. 2019;9:123. <https://doi.org/10.2147/DNND.S228939>
- McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*. 2014;13(6):614-629. [https://doi.org/10.1016/S1474-4422\(14\)70090-0](https://doi.org/10.1016/S1474-4422(14)70090-0)
- Croisile B, Ska B, Brabant MJ, et al. Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain Lang*. 1996;53(1):1-19. <https://doi.org/10.1006/brln.1996.0033>
- Ahmed S, de Jager CA, Haigh AM, et al. Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology*. 2013;27(1):79-85. <https://doi.org/10.1037/a0031288>
- Ash S, McMillan C, Gross RG, et al. The organization of narrative discourse in Lewy body spectrum disorder. *Brain Lang*. 2011;119(1):30-41. <https://doi.org/10.1016/j.bandl.2011.05.006>
- Nevler N, Ash S, Jester C, et al. Automatic measurement of prosody in behavioral variant FTD. *Neurology*. 2017;89(7):650-656. <https://doi.org/10.1212/WNL.0000000000004236>
- Yamada Y, Shinkawa K, Kobayashi M, et al. Tablet-based automatic assessment for early detection of Alzheimer's disease using speech responses to daily life questions. *Front Digit Health*. 2021;3:30. <https://doi.org/10.3389/fdgth.2021.653904>
- Beltrami D, Gagliardi G, Rossini Favretti R, et al. Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? *Front Aging Neurosci*. 2018;10:369. <https://doi.org/10.3389/fnagi.2018.00369>



13. O'Malley RPD, Mirheidari B, Harkness K, et al. Fully automated cognitive screening tool based on assessment of speech and language. *J Neurol Neurosurg Psychiatry*. 2021;92(1):12-15. <https://doi.org/10.1136/jnnp-2019-322517>
14. Eyigöz E, Mathur S, Santamaria M, et al. Linguistic markers predict onset of Alzheimer's disease. *EClinicalMedicine*. 2020;28:100583. <https://doi.org/10.1016/j.eclinm.2020.100583>
15. Kavé G, Levy Y. Morphology in picture descriptions provided by persons with Alzheimer's disease. *J Speech Lang Hear Res*. 2003;46:341-352.
16. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis*. 2016;49(2):407-422. doi:10.3233/JAD-150520
17. Boschi V, Catricala E, Consonni M, et al. Connected speech in neurodegenerative language disorders: a review. *Front Psychol*. 2017;8:269. <https://doi.org/10.3389/fpsyg.2017.00269>
18. Pakhomov SV, Hemmy LS. A computational linguistic measure of clustering behavior on semantic verbal fluency task predicts risk of future dementia in the nun study. *Cortex*. 2014;55:97-106. <https://doi.org/10.1016/j.cortex.2013.05.009>
19. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology*. 2005;65(12):1863-1872. <https://doi.org/10.1212/01.wnl.0000187889.17253.b1>
20. Ash S, McMillan C, Gross RG, et al. Impairments of speech fluency in Lewy body spectrum disorder. *Brain Lang*. 2012;120(3):290-302. <https://doi.org/10.1016/j.bandl.2011.09.004>
21. Ash S, Grossman M. Why study connected speech production. In *Cognitive neuroscience of natural language use*. 2015:29-58. <https://doi.org/10.1017/CBO9781107323667.003>
22. Rousseaux M, S'ève A, Vallet M, et al. An analysis of communication in conversation in patients with dementia. *Neuropsychologia*. 2010;48(13):3884-3890. <https://doi.org/10.1016/j.neuropsychologia.2010.09.026>
23. Ma A, Lau KK, Thyagarajan D. Voice changes in Parkinson's disease: what are they telling us? *J Clin Neurosci*. 2020;72:1-7. <https://doi.org/10.1016/j.jocn.2019.12.029>
24. Ruzs J, Cmejla R, Ruzickova H, Ruzicka E. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am*. 2011;129(1):350-367. <https://doi.org/10.1121/1.3514381>
25. Ruzs J, Čmejla R, Růžicková H, et al. Evaluation of speech impairment in early stages of Parkinson's disease: a prospective study with the role of pharmacotherapy. *J Neural Transm (Vienna)*. 2013;120(2):319-329. <https://doi.org/10.1007/s00702-012-0853-4>
26. Holmes RJ, Oates JM, Phyland DJ, et al. Voice characteristics in the progression of Parkinson's disease. *Int J Lang Commun Disord*. 2000;35(3):407-418. <https://doi.org/10.1080/136828200410654>
27. Gayraud F, Lee HR, Barkat-Defradas M. Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clin Linguist Phon*. 2011;25(3):198-209. <https://doi.org/10.3109/02699206.2010.521612>
28. Martínez-Sánchez F, García Meilán JJ, Pérez E, et al. Expressive prosodic patterns in individuals with Alzheimer's disease. *Psicothema*. 2012;24(1):16-21. pmid:22269358.
29. Themistocleous C, Eckerström M, Kokkinakis D. Voice quality and speech fluency distinguish individuals with mild cognitive impairment from healthy controls. *Plos one*. 2020;15(7):e0236009. <https://doi.org/10.1371/journal.pone.0236009>
30. Suárez-González A, Cassani A, Gopalan R, et al. When it is not primary progressive aphasia: a scoping review of spoken language impairment in other neurodegenerative dementias. *Alzheimers Dement (N Y)*. 2021;7(1):e12205. <https://doi.org/10.1002/trc2.12205>
31. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279. <https://doi.org/10.1016/j.jalz.2011.03.008>
32. Petersen RC, Aisen P, Beckett LA, et al. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209. <https://doi.org/10.1212/WNL.0b013e3181cb3e2>
33. McKeith IG, Ferman TJ, Thomas AJ, et al. Research criteria for the diagnosis of prodromal dementia with Lewy bodies. *Neurology*. 2020;94(17):743-755. <https://doi.org/10.1212/WNL.0000000000009323>
34. Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412-2414. <https://doi.org/10.1212/wnl.43.11.2412-a>
35. Goodglass H, Kaplan E, Weintraub S. BDAE: the boston diagnostic aphasia examination. *Lippincott Williams & Wilkins Philadelphia, PA*; 2001.
36. Bucks RS, Singh S, Cuerden JM, Wilcock GK. Analysis of spontaneous, conversational speech in dementia of Alzheimer type: evaluation of an objective technique for analysing lexical performance. *Aphasiology*. 2000;14(1):71-91. <https://doi.org/10.1080/026870300401603>
37. Ruzs J, Klempíř J, Baborová E, et al. Objective acoustic quantification of phonatory dysfunction in Huntington's disease. *PLoS one*. 2013;8(6):e65881. <https://doi.org/10.1371/journal.pone.0065881>
38. Jeffreys H. *The theory of probability*. OUP Oxford; 1998. ISBN: 9780191589676.
39. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;4:863. <https://doi.org/10.3389/fpsyg.2013.00863>
40. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory*. 1992, p. 144-152. <https://doi.org/10.1145/130385.130401>
41. Yamada Y, Shinkawa K, Kobayashi M, et al. Combining multimodal behavioral data of gait, speech, and drawing for classification of Alzheimer's disease and mild cognitive impairment. *J Alzheimers Dis*. 2021;84(1):315-327. <https://doi.org/10.3233/JAD-210684>
42. Buuren Sv, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Softw*. 2010;1:68. <https://doi.org/10.18637/jss.v045.i03>
43. Ruzs J, Hlavnička J, Tykalova T, et al. Smartphone allows capture of speech abnormalities associated with high risk of developing Parkinson's disease. *IEEE Trans Neural Syst Rehabil Eng*. 2018;26(8):1495-1507. <https://doi.org/10.1109/TNSRE.2018.2851787>
44. Mc Ardle R, Galna B, Donaghy P, et al. Do Alzheimer's and Lewy body disease have discrete pathological signatures of gait? *Alzheimers Dement*. 2019;15(10):1367-1377. <https://doi.org/10.1016/j.jalz.2019.06.4953>
45. Mc Ardle R, Pratt S, Buckley C, et al. Balance impairments as differential markers of dementia disease subtype. *Front Bioeng Biotechnol*. 2021;9:104. <https://doi.org/10.3389/fbioe.2021.639337>
46. Pieruccini-Faria F, Black SE, Masellis M, et al. Gait variability across neurodegenerative and cognitive disorders: results from the Canadian Consortium of Neurodegeneration in Aging (CCNA) and the gait and brain study. *Alzheimers Dement*. 2021;17(8):1317-1328. <https://doi.org/10.1002/alz.12298>
47. Orimaye SO, Wong JSM, Golden KJ. Learning predictive linguistic features for Alzheimer's disease and related dementias using verbal utterances. In: *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From linguistic signal to clinical reality*. 2014:78-87. <https://doi.org/10.3115/v1/W14-3210>
48. Roark B, Mitchell M, Hosom JP, et al. Spoken language derived measures for detecting mild cognitive impairment. *IEEE Trans Audio Speech Lang Process*. 2011;19(7):2081-2090. <https://doi.org/10.1109/TASL.2011.2112351>

49. Zhang L, Ngo A, Thomas JA, et al. Neuropsychological test validation of speech markers of cognitive impairment in the Framingham cognitive aging cohort. *Explor Med*. 2021;2:232. doi: [10.37349/emed.2021.00044](https://doi.org/10.37349/emed.2021.00044)
50. Gauthier S, Rosa-Neto P, Morais JA, et al. World Alzheimer report 2021: journey through the diagnosis of dementia. 2021. Accessed: March 9, 2022. <https://www.alzint.org/resource/world-alzheimer-report-2021/>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Yamada Y, Shinkawa K, Nemoto M, Ota M, Nemoto K, Arai T. Speech and language characteristics differentiate Alzheimer's disease and dementia with Lewy bodies. *Alzheimer's Dement*. 2022;14:e12364. <https://doi.org/10.1002/dad2.12364>