ABSTRACT

Background and Purpose: Language dysfunction is a symptom common to patients with Alzheimer’s disease (AD). Speech feature analysis may be a patient-friendly screening test for early-stage AD. We aimed to investigate the speech features of amnestic mild cognitive impairment (aMCI) compared to normal controls (NCs).

Methods: Spoken responses to test questions were recorded with a microphone placed 15 cm in front of each participant. Speech samples delivered in response to four spoken test prompts (free speech test, Mini-Mental State Examination [MMSE], picture description test, and sentence repetition test) were obtained from 98 patients with aMCI and 139 NCs. Each recording was transcribed, with speech features noted. The frequency of the ten speech features assessed was evaluated to compare speech abilities between the test groups.

Results: Among the ten speech features, the frequency of pauses ($p=0.001$) and mumbles ($p=0.001$) were significantly higher in patients with aMCI than in NCs. Moreover, MMSE score was found to negatively correlate with the frequency of pauses ($r=-0.441, p<0.001$) and mumbles ($r=-0.341, p<0.001$).

Conclusions: Frequent pauses and mumbles reflect cognitive decline in aMCI patients in episodic and semantic memory tests. Speech feature analysis may prove to be a speech-based biomarker for screening early-stage cognitive impairment.

Keywords: Alzheimer Disease; Biomarkers; Mild Cognitive Impairment; Speech Disorders

INTRODUCTION

The next generation of Alzheimer’s disease (AD) treatments would be most effective if administered in the early stages of the disease, before irreversible brain damage or significant cognitive decline has occurred. Early diagnosis of the disease is therefore particularly critical, and over the last decade numerous new techniques intended to achieve this have been developed. An AD diagnosis is currently made on the basis of laboratory test results, magnetic resonance imaging (MRI) images of the brain, and neuropsychological tests. While the standard neuropsychological tests used to detect AD are thorough and allow for the assessment of diverse domains of cognition, the examination is time-consuming and
somewhat burdensome for the patients. Developing AD screening tests that are patient-friendly and able to reflect real life is a priority.

Speech feature analysis may be a less uncomfortable means of screening for AD. Recording the interviews between doctors/neuropsychologists and patients and subsequently analyzing the patient’s speech features based on that recording would obviate the need for the patient to undergo a cognitive test. Language dysfunction is a common symptom in patients with AD which are known to have various speech features such as word finding difficulties, verbal fluency, and temporal changes in spontaneous speech. Prior research has confirmed that these speech features are affected in cases of amnestic mild cognitive impairment (aMCI).

The language and speech dysfunctions associated with AD are receiving increasing scrutiny by researchers. In most of prior studies, patients were asked to perform free speech tests by recalling a day or describing a picture. The number and duration of pauses, speech duration and length, and speech rate were the most informative features in the comparison of AD/aMCI and control group in free speech tests. In other studies, verbal fluency tests were performed by asking patients to generate as many words as possible falling within specific phonemic or semantic categories. The duration of silent and voiced segment were the most informative features in the comparison of AD/aMCI and control group in verbal fluency tests. In some studies, patients were asked to repeat a sentence, read a paragraph out loud, and count down numbers to investigate different memory, semantic processing, and acoustic or context-related processing. Most studies to date have been conducted in European languages.

In this study, the speech features of Korean native speakers with aMCI were investigated and compared to normal controls (NCs). We assessed the speech features examined in conversations in the same way that we assessed interviews during dementia screenings.

METHODS

Recruitment of participants

The study was conducted at the Jeonbuk National University Hospital Dementia Clinic between May 2019 and December 2020. Ninety-eight aMCI patients and 139 NC participants were recruited. Patients with aMCI were diagnosed based on the criteria proposed by Petersen et al., after a rigorous examination that included neuropsychological tests and brain MRI, as well as assessment by neurologists, neuropsychologists, and radiologists. None of the aMCI patients fulfilled the standard criteria for primary progressive aphasia (PPA). The NC participants were age-matched healthy people considered intact because their scores on the Korean Mini-Mental State Examination (K-MMSE) were within one standard deviation of the age- and education-matched mean. Individuals who satisfied the following criteria were recruited for participation: between 60 and 85 years old and no hearing impairment. We excluded participants with diseases that could affect speech function. This study was performed in compliance with the protocols established by Jeonbuk National University Hospital Institutional Review Board, pursuant to which written informed consent was obtained from all participants.

Recording protocol

The spoken tests were recorded with a microphone (C414 B-ULS; AKG Acoustics, Vienna, Austria) and a smartphone (Galaxy S7; Samsung, Seoul, Korea) placed 15 cm in front of each participant. Before the spoken tests were performed, each participant was asked to speak for
30 seconds, during which a recording was taken. As each participant spoke, the experimenter verified that the participant’s volume scale was over 3,000 (sample value) to make sure that the sound volume was sufficient for recording. The experimenter then delivered the following instructions in Korean: “You will be answering various questions. Please try not to interrupt while the experimenter is speaking. After each question, you can start answering at intervals of 1 second. Take your time, please tell me enough to answer and let me know when you are done.” The 20- to 30-minute-long spoken tests were then administered.

The following four spoken tests were administered and recorded, consistent with prior studies⁷⁻¹⁵:

(i) Free speech test: The aim of the free speech test was to provoke spontaneous speech. Participants were asked which food they ate today and how they arrived at the hospital. If the answers were insufficient, the participants were encouraged to provide additional answers.

(ii) MMSE: The aim of the MMSE test was to examine general cognition. Participants were asked to answer all questions asked as part of the K-MMSE.

(iii) Picture description test: The aim of the picture description test was to provoke spontaneous speech with semantic processing. Participants were asked to describe the “Picnic Scene” from the Western Aphasia Battery, Revised (WAB-R; NCS Pearson, Inc., Bloomington, MN, USA).

(iv) Sentence repetition test: The aim of the sentence repetition test was to examine working memory as well as measure acoustic and phonetic measures. Participants were asked to repeat 6 sentences articulated by the experimenter; Standard Korean sentences were used.¹⁹

All responses were recorded, and a transcript of each speech sample was generated by two experimenters using Transcriber software (http://trans.sourceforge.net/).

**Speech features**

We annotated several speech features that could not be transcribed into text based on previous research. This mainly consisted of extending the transcripts by annotating pause, breath, and laughter in a manner consistent with previous research.²⁰ In addition, we defined the final 10 speech features by adding annotations that were relevant to our needs: adding, breath, continuous, interrupt, laugh, mumble, pause, self-talk, stutter, and interjection (Table 1). These speech features were labeled manually by our two experimenters. Then, we determined the number of speech features that appeared during each spoken test.

| Feature     | Description                                                      |
|-------------|------------------------------------------------------------------|
| Adding      | Participants interrupting the examiner                           |
| Breathing   | Participants making breathing sounds, coughing, etc.             |
| Continuous  | Continuous vocalization without one-second interval after questions |
| Interrupt   | No voice for more than 5 seconds or examiner intervention         |
| Laugh       | Situations involving laughter during utterances                  |
| Mumble      | Unable to understand the content of words                        |
| Pause       | No voice for more than 3 seconds                                  |
| Self-talk   | Other than a reply to the conversation                           |
| Stutter     | Stuttering during speech                                         |
| Interjection| Meaningless words like “ah” and “mmm”                            |
**Statistical analysis**

To assess any demographic or cognitive profile differences between the 2 groups (aMCI and NCs), we used a $\chi^2$ test or Fisher’s exact test for categorical variables, and a Mann-Whitney $U$ test for continuous variables that did not follow a normal distribution. The Mann-Whitney $U$ test was also used to investigate differences in speech features among the two groups, as the relevant variables did not follow a normal distribution. Pearson’s correlation was used to investigate the association between the number of speech features and MMSE scores. Correlations were calculated for both groups and separately for each group. The statistical software package SPSS version 23 was used for data analyses. For all tests, the level of statistical significance was set to $p<0.05$.

**RESULTS**

**Demographics**

Our study group was comprised of 98 patients with aMCI and 139 NCs. There was no significant difference in median age and education between the aMCI patients and NCs. Apolipoprotein was only analyzed in 87 of the aMCI patients. The prevalence of apolipoprotein $\varepsilon_4$ (APOE4) carriers among aMCI patients was 35.6% (31/87) (participants with one or more copies of the $\varepsilon_4$ allele [i.e., $\varepsilon_2/4$, $\varepsilon_3/4$, $\varepsilon_4/4$] were considered $\varepsilon_4$ carriers). In the MMSE, aMCI patients achieved significantly lower scores than the NCs. The demographic features of the participants are summarized in Table 2.

**Speech features across all tests**

The frequency of mumbles ($p=0.001$), pauses ($p=0.001$), and stutters ($p=0.004$) were significantly higher in aMCI patients than in NCs (Fig. 1A and B). Moreover, the frequency of continuous ($p=0.004$) was significantly lower in aMCI patients than in NCs. Table 3 presents the median and interquartile range of speech features in aMCI patients and NCs.

**Speech features in each test**

The frequency of mumbles ($p=0.012$), pauses ($p<0.001$), and stutters ($p=0.01$) was significantly higher in aMCI patients administered the free speech test than in NCs. The frequency of breathing ($p=0.016$) and continuous ($p<0.001$) were significantly lower in aMCI patients than in NCs (Supplementary Table 1).

### Table 2. Demographics

| Characteristics | aMCI     | NC       | $p$-value |
|-----------------|----------|----------|-----------|
| No. of patients | 98       | 139      |           |
| Age (yrs)       | 71 (65–77) | 70 (65–74) | 0.181    |
| Sex F:M         | 53:45    | 104:35   |           |
| APOE $\varepsilon_4$ carrier* | 35.6% (31/87) |           | 0.051    |
| Education       | 9 (6–12) | 9 (7–12) |           |
| MMSE            | 26 (25–29) | 29 (27–30) | $<0.001$ |
| CDR             | 0.5 (0.5–0.5) | 0 (0–0) | $<0.001$ |
| CDR-SB          | 1.5 (1–2) | 0 (0–0) | $<0.001$ |

Values are presented as median (range). Bold type indicates statistical significance.

APOE: apolipoprotein, aMCI: amnestic mild cognitive impairment, NC: normal control, CDR: Clinical Dementia Rating, CDR-SB: Clinical Dementia Rating-Sum of Boxes, MMSE: Mini-Mental State Examination.

*APOE was analyzed in 87 aMCI patients. Participants with one or more copies of the $\varepsilon_4$ allele (i.e., $\varepsilon_2/4$, $\varepsilon_3/4$, $\varepsilon_4/4$) were considered $\varepsilon_4$ carriers.
In the MMSE test, the frequency of breathing ($p=0.025$), interrupts ($p=0.044$), laughs ($p=0.014$), mumbles ($p<0.001$), pauses ($p<0.001$), and stutters ($p=0.012$) was significantly higher in aMCI patients than in NCs (Supplementary Table 2).

The frequency of interrupts ($p=0.033$), mumbles ($p<0.001$), pause ($p<0.001$), and stutters ($p=0.027$) was significantly higher in aMCI patients than in NCs in the picture description test. The

**Table 3.** The median number of speech features in aMCI and NC participants

| Feature     | aMCI       | NC        | $p$-value |
|-------------|------------|-----------|-----------|
| Adding      | 0 (0–0)    | 0 (0–0)   | 0.223     |
| Breathing   | 11 (6–20)  | 13 (6–28) | 0.188     |
| Continuous  | 3 (1–5)    | 4 (2–9)   | 0.004     |
| Interrupt   | 0 (0–1)    | 0 (0–1)   | 0.053     |
| Laugh       | 2 (0–4)    | 1 (0–4)   | 0.955     |
| Mumble      | 6 (4–13)   | 4 (2–8)   | 0.001     |
| Pause       | 12 (7–19)  | 9 (5–14)  | 0.001     |
| Self-talk   | 0 (0–0)    | 0 (0–0)   | 0.828     |
| Stutter     | 0 (0–2)    | 0 (0–1)   | 0.004     |
| Interjection| 1 (0–3)    | 1 (0–5)   | 0.817     |

Values are presented as median (interquartile range). Bold type indicates statistical significance.

aMCI: amnestic mild cognitive impairment, NC: normal control.
frequency of continuous ($p=0.001$) was significantly lower in aMCI patients than in NCs (Supplementary Table 3).

In the sentence repetition test, the frequency of continuous ($p=0.017$), mumbles ($p=0.002$), and stutters ($p<0.001$) was significantly higher in aMCI patients than in NCs (Supplementary Table 4).

**Correlations between MMSE score and speech features**

Across both groups, the MMSE score negatively correlated with the frequency of interrupts ($r=-0.146$, $p=0.025$), laughs ($r=-0.197$, $p=0.002$), mumbles ($r=-0.341$, $p<0.001$), pauses ($r=-0.441$, $p<0.001$), self-talks ($r=-0.235$, $p<0.001$), and stutters ($r=-0.215$, $p=0.001$), but positively correlated with the frequency of continuous ($r=0.130$, $p=0.045$) (Table 4, Fig. 1C).

In aMCI patients the MMSE score negatively correlated with the frequency of laughs ($r=-0.225$, $p=0.026$), mumbles ($r=-0.321$, $p=0.001$), pauses ($r=-0.439$, $p<0.001$), and self-talks ($r=-0.378$, $p<0.001$) (Table 4, Fig. 1D), while in The MMSE score negatively correlated with the frequency of adding ($r=-0.194$, $p=0.022$), interrupts ($r=-0.257$, $p=0.002$), laughs ($r=-0.255$, $p=0.003$), and pauses ($r=-0.173$, $p=0.043$; Table 4) in NCs.

**DISCUSSION**

In this study, we investigated the speech features of aMCI and NC participants. Four spoken tests were administered—i) the free speech test, ii) the MMSE, iii) the picture description test, and iv) the sentence repetition test—and each participant’s speech was recorded. Then, we identified ten speech features and the number of these speech features in each spoken test. The frequency of these speech features was used to compare the speaking abilities of each group. In summary, pauses and mumbles were consistently more present in aMCI patients than in NCs. Moreover, the frequency of pauses and mumbles was negatively correlated with MMSE score. Pauses in the speech of AD patients are typically regarded as dysfluency.21,22 Mumbles similarly reflect hesitation in speech. According to prior studies, the frequency of pauses is significantly higher in AD patients than in NCs.23-24 Specifically, one prior study revealed that the number of pauses is significantly different between aMCI patients and NCs performing the question-answer test.25 This is consistent with our results, in which significant differences in the numbers of pauses and mumbles were found between the two tested groups across three (free speech, MMSE, picture description) of the four (sentence repetition) tests. Stutter had...
relatively low absolute frequency though analysis of median numbers of speech features did reveal some differences between aMCI patients and NCs.

Only in the sentence repetition test did the numbers of pauses and mumbles not differ across the aMCI and NCs test groups. In the free speech test participants were asked which food they ate that day and how they arrived at the hospital. Consequently, free speech test is associated with recall. The MMSE test is also designed to spot delayed recall. The picture description test requires that participants rely less on episodic memory and more on semantic knowledge. In the context of these tests, more frequent pauses and mumbles might reflect greater effort to recall and higher mental loading. In contrast, the sentence repetition test relies on working memory rather than episodic or semantic memory. Pauses and mumbles might reflect cognitive decline in episodic and semantic memory impairments.

Interestingly, the number of continuous was significantly lower in aMCI participants than their NC counterparts. While we instructed all participants they could begin answering 1 second after each question and to take care that their voice did not overlap with the experimenter’s, NCs were more likely to respond before the 1 second interval ended, possibly reflecting faster cognitive processing on their part.

The frequency of pauses and mumbles were negatively correlated with MMSE scores while continuous was positively correlated with the MMSE scores of the aMCI and NC groups. This suggests that participants with lower MMSE scores were less fluent and more hesitant in their language expression. Prior studies suggested that pauses and mumbles might be associated with compensatory mechanisms in the earliest stages of AD, as speech rate and hesitation have been found to be negatively correlated with MMSE scores, and speech features (including the frequency of pauses) explain 47.8% of the variation in MMSE scores. This study confirms that less fluent speech is associated with cognitive decline.

The present study had several limitations. First, speech dysfunction can appear as a result of normal aging or the presence of an array of neurodegenerative diseases as well as other diseases affecting the muscles of the tongue, lips, jaw, and vocal tract. In this study, we excluded participants with diseases that could affect speech function, but patients could have mixed diseases in the real world. Also, aMCI patients were compared against NCs, though comparison between patients with different neurodegenerative diseases will be needed to validate our results. Second, we could not assess inter-rater reliability between two experimenters. Future studies should perform linguistic as well as acoustic feature analyses to improve accuracy and reliability. In summary, speech feature analysis may be a speech-based biomarker for the screening of early-stage cognitive impairment. Speech can be recorded with a smartphone, which could be useful remote and/or frequent monitoring. Despite these limitations, our findings are interesting, in as much as they validate findings from previous studies conducted in European languages.

**SUPPLEMENTARY MATERIALS**

**Supplementary Table 1**
The median and interquartile range of the frequency of 10 speech features in the free speech test

Click here to view
Supplementary Table 2
The median and interquartile range of the frequency of 10 speech features in the MMSE test
Click here to view

Supplementary Table 3
The median and interquartile range of the frequency of 10 speech features in the picture description test
Click here to view

Supplementary Table 4
The median and interquartile range of the frequency of 10 speech features in the sentence repetition test
Click here to view

REFERENCES

1. Cummings J, Lee G, Ritter A, Sabbagh M, Zhong K. Alzheimer’s disease drug development pipeline: 2020. Alzheimers Dement (N Y) 2020;6:e12050.
   PUBMED | CROSSREF

2. Kim KW, Choi JD, Lee H, Lee NK, Park S, Chin J, et al. Social Event Memory Test (SEMT): a video-based memory test for predicting amyloid positivity for Alzheimer’s disease. Sci Rep 2018;8:10421.
   PUBMED | CROSSREF

3. Taler V, Phillips NA. Language performance in Alzheimer’s disease and mild cognitive impairment: a comparative review. J Clin Exp Neuropsychol 2008;30:501-556.
   PUBMED | CROSSREF

4. Khodabakhsh A, Yesil F, Guner E, Demiroglu C. Evaluation of linguistic and prosodic features for detection of Alzheimer’s disease in Turkish conversational speech. EURASIP J Audio Speech Music Process 2015;2015:9.
   CROSSREF

5. Pereiro AX, Semenza C, Juncos-Rabadán O. Editorial: language and mild cognitive impairment. Front Psychol 2020;11:2264.
   PUBMED | CROSSREF

6. Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J; Consortium for the Early Identification of Alzheimer’s disease-Quebec. Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer’s type dementia in older adults: a systematic review and meta-analysis. Neuropsychol Rev 2017;27:328-353.
   PUBMED | CROSSREF

7. Petti U, Baker S, Korhonen A. A systematic literature review of automatic Alzheimer’s disease detection from speech and language. J Am Med Inform Assoc 2020;27:1784-1707.
   PUBMED | CROSSREF

8. Au R, Christianson K, McManus C, Knight A, Karjadi C. Global challenges in cognitive assessment: how technology will drive neuropsychologists to the forefront. Alzheimers Dement 2020;16:e038142.
   CROSSREF

9. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJ. Ten years of research on automatic voice and speech analysis of people with Alzheimer’s disease and mild cognitive impairment: a systematic review article. Front Psychol 2021;12:620251.
   PUBMED | CROSSREF

10. Beltrami D, Gagliardi G, Rossini Favretti R, Ghidoni E, Tamburini F, Calzà L. Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? Front Aging Neurosci 2018;10:369.
    PUBMED | CROSSREF
11. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer’s disease in narrative speech. J Alzheimers Dis 2016;49:407-422.

12. Hernández-Dominguez L, Ratré S, Sierra-Martínez G, Roche-Bergua A. Computer-based evaluation of Alzheimer’s disease and mild cognitive impairment patients during a picture description task. Alzheimers Dement (Amst) 2018;10:260-268.

13. König A, Satt A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer’s disease. Alzheimers Dement (Amst) 2015;1:112-124.

14. Clark DG, Wadley VG, Kapur P, DeRamus TP, Singletary B, Nicholas AP, et al. Lexical factors and cerebral regions influencing verbal fluency performance in MCI. Neuropsychologia 2014;54:98-111.

15. Clark DG, McLaughlin PM, Woo E, Hwang K, Hurtz S, Ramirez L, et al. Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. Alzheimers Dement (Amst) 2016;2:113-122.

16. Mirzaei S, El Yacoubi M, Garcia-Salicetti S, Boudy J, Kahindo C, Cristancho-Lacroix V, et al. Two-stage feature selection of voice parameters for early Alzheimer’s disease prediction. IRBM 2018;39:430-435.

17. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018;90:126-135.

18. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology 2011;76:1006-1014.

19. Choi SH. Development of Korean standardized sentences on voice quality evaluation for dysphonia. Audiol Speech Res 2018;14:128-142.

20. Toth L, Hoffmann I, Gosszolay G, Vincze V, Szatloczki G, Banreti Z, et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. Curr Alzheimer Res 2018;15:130-138.

21. Davis BH, Maclagan M. Examining pauses in Alzheimer’s discourse. Am J Alzheimers Dis Other Demen 2009;24:141-154.

22. Sluis RA, Angus D, Wiles J, Back A, Gibson TA, Liddle J, et al. An automated approach to examining pausing in the speech of people with dementia. Am J Alzheimers Dis Other Demen 2020;35:1533317520939773.

23. 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;25:198-209.

24. Pistono A, Jucla M, Barbeau EJ, Saint-Aubert L, Lemesle B, Calvet B, et al. Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer’s disease. J Alzheimers Dis 2016;50:687-698.

25. Mueller KD, Hermann B, Mecollari J, Turkstra LS. Connected speech and language in mild cognitive impairment and Alzheimer’s disease: a review of picture description tasks. J Clin Exp Neuropsychol 2018;40:917-939.

26. Singh S, Bucks RS, Cuerden JM. Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. Aphasiology 2001;15:571-583.

27. Small IA, Kemper S, Lyons K. Sentence repetition and processing resources in Alzheimer’s disease. Brain Lang 2000;75:232-258.
28. Pistono A, Pariente J, Bézy C, Lemesle B, Le Men I, Jucla M. What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer’s disease. Neuropsychologia 2019;124:133-143.  
   [PUBMED] [CROSSREF]

29. Hoffmann I, Nemeth D, Dye CD, Pákáski M, Irinyi T, Kálmán J. Temporal parameters of spontaneous speech in Alzheimer’s disease. Int J Speech-Language Pathol 2010;12:29-34.  
   [PUBMED] [CROSSREF]

30. Qiao Y, Xie XY, Lin GZ, Zou Y, Chen SD, Ren RJ, et al. Computer-assisted speech analysis in mild cognitive impairment and Alzheimer’s disease: a pilot study from Shanghai, China. J Alzheimers Dis 2020;75:211-221.  
   [PUBMED] [CROSSREF]

31. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of speech-based digital biomarkers: review and recommendations. Digit Biomark 2020;4:99-108.  
   [PUBMED] [CROSSREF]