Parkinsonism and Related Disorders 113 (2023) 105411

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Analysis of spontaneous speech in Parkinson's disease by natural language processing

Katsunori Yokoi [a,b], Yurie Iribe [c], Norihide Kitaoka [d], Takashi Tsuboi [a], Keita Hiraga [a], Yuki Satake [a], Makoto Hattori [a], Yasuhiro Tanaka [a,e], Maki Sato [a], Akihiro Hori [f], Masahisa Katsuno [a,g,*]

[a] Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan
[b] Department of Neurology, National Hospital for Geriatric Medicine, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan
[c] School of Information Science and Technology, Aichi Prefectural University, Nagakute, Japan
[d] Department of Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, Japan
[e] Department of Health Science, Aichi Gakuin University, 12 Araike, Iwasaki-cho, Nisshin-city, Aichi, Japan
[f] Kumiai Kosei Hospital, Takayama, Gifu, Japan
[g] Department of Clinical Research Education, Nagoya University Graduate School of Medicine, Nagoya, Japan

ARTICLE INFO

Keywords:
Parkinson's disease
Linguistic analysis
Natural language processing
Spontaneous speech
Part-of-speech

ABSTRACT

Introduction: Patients with Parkinson's disease (PD) encounter a variety of speech-related problems, including dysarthria and language disorders. To elucidate the pathophysiological mechanisms for linguistic alteration in PD, we compared the utterance of patients and that of healthy controls (HC) using automated morphological analysis tools.
Methods: We enrolled 53 PD patients with normal cognitive function and 53 HC, and assessed their spontaneous speech using natural language processing. Machine learning algorithms were used to identify the characteristics of spontaneous conversation in each group. Thirty-seven features focused on part-of-speech and syntactic complexity were used in this analysis. A support-vector machine (SVM) model was trained with ten-fold cross-validation.
Results: PD patients were found to speak less morphemes on one sentence than the HC group. Compared to HC, the speech of PD patients had a higher rate of verbs, case particles (dispersion), and verb utterances, and a lower rate of common noun utterances, proper noun utterances, and filler utterances. Using these conversational changes, the respective discrimination rates for PD or HC were more than 80%.
Conclusions: Our results demonstrate the potential of natural language processing for linguistic analysis and diagnosis of PD.

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease (AD). It is a relatively common disorder and its prevalence increases with age [1]. PD is characterized by motor signs of bradykinesia, rigidity, and resting tremor, as well as by nonmotor symptoms, including cognitive, neuropsychiatric, sleep, autonomic, and sensory disturbances [2].

Communicative changes are also common in PD; these are attributable to various factors including speech utterances, prosodic changes, pronunciation and articulation changes, language changes, discourse management and pragmatics, and psychosocial influences [3]. Studies have shown that more than 90% of PD patients experience some form of speech impairment [4].

Various studies have documented voice-speech deterioration in PD. For instance, repetitive speech disorder related to freezing of gait [5], and an overall down-scaling of speech production [6], have been reported in PD. Other reports suggested that PD affects all stages of

* Corresponding author. Department of Neurology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, 466-8550, Japan.
E-mail addresses: katsu-y@ncgg.go.jp (K. Yokoi), iribe@ist.aichi-pu.ac.jp (Y. Iribe), kitaoka@tut.jp (N. Kitaoka), takashi80@gmail.com (T. Tsuboi), gance2000@googlemail.com (K. Hiraga), y.satake0316@gmail.com (Y. Satake), hattori.makoto@b.mbox.nagoya-u.ac.jp (M. Hattori), acv27470@nyc.odn.ne.jp (Y. Tanaka), mi.suenyo.0013@gmail.com (M. Sato), a.hori@gfkosei.or.jp (A. Hori), ka2no@med.nagoya-u.ac.jp (M. Katsuno).

https://doi.org/10.1016/j.parkreldis.2023.105411
Received 20 February 2023; Received in revised form 14 April 2023; Accepted 21 April 2023
Available online 26 April 2023
1353-8020/© 2023 Elsevier Ltd. All rights reserved.

language production including conceptualization and functional and positional processing [7], and that PD patients exhibit linguistic and semantic deficits even in the absence of mild cognitive impairment (MCI) [8].

Recetly natural language processing has been applied to investigate the language production in patients with PD. Analysis using scores extracted from textual retelling data of content with movement showed 85% accuracy in discriminating between PD without mild cognitive impairment and HC [9]. By transforming spontaneous speech features into a multidimensional vector using a neural network, PD patients and healthy subjects can be discriminated with up to 72% accuracy [10]. However, there is a paucity of reports on the analysis of spontaneous speech texts of PD patients, particularly those with normal cognition.

The purpose of this study was to analyze speech texts using natural language processing to clarify the pathophysiological mechanisms for utterance alteration in cognitively unimpaired PD patients.

2. Methods

2.1. Ethical compliance

This study was conducted in accordance with the principles of Declaration of Helsinki, the Ethics Guidelines for Human Genome/Gene Analysis Research, and the Ethical Guidelines for Medical and Biological Research Involving Human Subjects endorsed by the Japanese government. The study protocol was approved by the Ethics Review Committee of Nagoya University Graduate School of Medicine (approval number: 2016-0238). Written informed consent was obtained from all subjects prior to their enrolment.

2.2. Participants

PD patients were recruited from Kumiai Kosei Hospital and Nagoya University Hospital. The inclusion criteria were: PD patients diagnosed according to the UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria [11]; age between 40 and 80 years; and Japanese version of the Montreal Cognitive Assessment (MoCA-J) score >23. The cut-off for MoCA-J score was set according to a previous report on the classification of PD by cognitive function [12]. The exclusion criteria were: dementia with Lewy body (DLB) or other neurodegenerative or psychiatric diseases; cerebrovascular diseases; any family history of parkinsonism; and insufficient clinical data. We also recruited age- and sex-matched healthy controls (HC) from Nagoya University Hospital or Kumiai Kosei Hospital. All HCs were voluntary participants, who incidentally visited the hospitals or had symptoms other than those of parkinsonism. None of the HCs had any history of neurodegenerative or cerebrovascular diseases.

2.3. Data acquisition

We acquired and assessed MoCA-J, phonemic and semantic fluency test, voice recording of all patients and controls, and Unified Parkinson's disease rating scale (UPDRS) part III of PD patients at ON condition. MoCA-J is a 30-point test for assessment of several cognitive domains such as verbal memory, visuospatial and executive abilities, attention, and language. The test is used to assess cognitive reserve and early cognitive decline [13]. Phonemic and semantic fluency discrepancy tests are commonly used as verbal tasks for assessment of MCI in neurodegenerative diseases including AD and PD [14]. In the phonemic fluency discrepancy test, the participants were asked to tell as many words starting with Ka as they could think of within 1 min. In the semantic fluency discrepancy test, participants were asked to name as many animals as they could think of within 1 min. Following these tests, the participants were provided the following instructions: "Tell me the flow of the day today from morning" and "Please tell me your favorite foods." We recorded the spontaneous responses of the participants to these questions and subjected them to analysis. K.Y, Y.T., or M.S. made the conversation with the participants, with standardization of the questioning approach in advance.

2.4. Linguistic analysis

The conversation was manually transcribed into text for natural language processing. As a preprocessing for language processing, we performed text normalization to remove notational distortions. Specifically, unification of character types, unification of full-width and half-width characters, and unification of upper and lower case characters were carried out. However, no text was deleted for the analysis. We utilized a machine learning approach for speech text analysis to discriminate PD and HC. A morphological analysis tool, MeCab, and a dependency structure analysis tool, Cabocha, was used to analyze and count the frequency of each part-of-speech and dependency distance (Fig. S1). MeCab was used to change texts to part-of-speech tag (Table S1). Cabocha was used to change texts to phase for analyzing syntax, such as the dependency distance and depth, which reflect the complexity of the text (Fig. S2, Table S2). Since the amount of speech differs from person to person, each item of the parts of speech was evaluated as a ratio to total counts. Among syntax valuables, type token ratio, different noun ratio, vocabulary level, maximum dependency distance, and dependency depth are known to be associated with cognitive function (Table S2) [15, 16]. Collectively, we employed 37 linguistic features of utterance for our analysis (Table S3). For discrimination between PD and HC, we selected the valid features for classification by the Wrapper method's positive best-first search method which was developed for extracting features from the data set and training them into the model. Using this method, we halved the number of features per trial. At a preprocessing stage of machine learning, we carried out hyper-parameter tuning by grid search to improve the accuracy of the model. We then performed analysis using support-vector machine (SVM) in a 10-fold cross-validation method. Cross-validation is an objective method for evaluating the performance of predictors; 10-fold cross-validation entails performing the fitting procedure for a total of 10 times, with each fit being performed on a training set consisting of 80% of the total training set selected at random, with the remaining 10% used as a held-out set for validation, and the other 10% used as a test set. There was no overlap of individuals among the sets. We presented the negative and positive predictive values, sensitivity, and specificity along with a 2 × 2 contingency table. Precision is the percentage of data predicted to be positive that are actually positive; Recall is the percentage of those predicted to be positive out of those that are actually positive; and F value is the harmonic mean of precision and recall.

2.5. Statistical analysis

Statistical analyses were conducted using EZR (Saitama Medical Center, Jichi Medical University, Saitama, Japan) [17], which is a graphical user interface for R (The R Foundation for Statistical Computing, Vienna, Austria). The demographic and clinical scores of patients with PD and HC groups were compared using Student's *t*-test. Continuous and categorical data are presented as mean ± SD and frequency, respectively, unless stated otherwise. *p*-values <0.05 were considered indicative of statistical significance. We performed multivariate analysis of the language characteristics of the participants' items that showed significant differences between PD and HC. Six variables (age, MoCA-J representing cognitive function, years of education, disease duration, UPDRS III, and levodopa equivalent daily dose [LEDD]) were included in the multivariate analysis. Similarly, we performed multivariate analysis of the items selected using the Wrapper method. For each item, we also performed univariate analyses using Pearson's product-rate correlation coefficients.

2.6. Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

3. Results

3.1. Participant characteristics

A total of 73 patients with PD and 54 HC were recruited between April 2012 and March 2020. Of these, 17 PD patients and one HC subject were excluded because of insufficient data; in addition, 3 patients who were eventually diagnosed as having DLB were also excluded. Finally, 53 PD patients (24 males and 39 females) and 53 HC (24 males and 39 females) were included in the analysis. There were no significant between-group differences with respect to age, sex, years of education, or MoCA-J score (Table 1).

3.2. Natural language analysis with conventional approach

Although there were no significant differences in phonemic or semantic fluency tests between PD and HC groups, the values of number of morphemes were significantly less in PD than in HC (Table 2). There were no significant between-group differences with respect to the number of sentences. Multivariate analysis revealed a significant correlation of number of morphemes with LEDD, but not with other background features (Table S4). In univariate analysis, the correlation was between the number of morphemes and LEDD was not found (Table S5).

3.3. Natural language analysis with machine learning approach

Using the positive best-first search methods of Wrapper method, we performed four trials to select the language items for discriminating PD from HC. The third trial showed the best points at F-measure in the results of 10-fold cross-validation (Table S6, Table S7). The sensitivity, specificity, positive predictive value, and negative predictive value were all above 0.83 at the third trial (Table 3). The precision, recall, and F-measure of HC or PD were all more than 0.83 at the third trial (Table S7). The area under the receiver operating characteristic (ROC) curves of the 10-fold cross-validation were more than 0.88 on average throughout the trials (Fig. S3). Thus, we selected six items used for the third trial: verb ratio, case particle ratio (dispersion), general noun utterance ratio, proper noun utterance ratio, verb utterance ratio, and filler utterance ratio.

When both sex groups were analyzed together, all six items showed a significant difference between the PD and HC groups (Fig. 1). In sex-specific analysis, verb ratio and case particle ratio showed significant differences only in males, while case particle and general noun utterance ratio showed significant differences only in females.

Although multivariate analysis of number of morphemes with the six selected items showed the correlation with verb ratio and verb utterance ratio (Table S8), univariate analysis detected no correlation (Table S9). Multivariate analysis of the six selected items with age, MoCA-J, years of education, disease duration of PD, UPDRS, and LEDD showed the correlation between the verb ratio and age (Table S10). Univariate analysis also suggested a weak correlation between the verb ratio and age (Table S11).

Table 1
Clinical features of the study population.

	Healthy control (n = 53)	PD patients with normal cognitive (n = 53)	*p*-value
Age, years	61.96 ± 6.06 (50–75)	62.75 ± 6.61 (48–75)	0.802
Sex (Male: Female)	24 : 29	24 : 29	1.000
MoCA-J	27.43 ± 1.42 (24–30)	26.94 ± 1.41 (24–30)	0.187
Years of education	13.00 ± 2.20 (9–18)	13.38 ± 2.77 (9–23)	0.860
Hoehn & Yahr stage	NA	2.34 ± 0.81 (1–4)	NA
Age of onset of PD	NA	49.87 ± 8.03 (32–67)	NA
Disease duration of PD	NA	12.51 ± 6.57 (1–43)	NA
UPDRSIII	NA	23.54 ± 14.94 (7–74)	NA
LEDD	NA	768.91 ± 345.03 (150–1558)	NA

Data presented as mean ± standard deviation (range). PD = Parkinson's disease, MoCA-J = Japanese version of the Montreal cognitive assessment, UPDRS = unified Parkinson's disease rating scale, LEDD = L-dopa equivalent daily dose, N.A. = not available.
p-values were determined by Student's *t*-test.

Table 2
Linguistic characteristics of the participants.

	Healthy control (n = 53)	PD patients with normal cognitive (n = 53)	*p*-value
Phonemic fluency task	12.40 ± 4.08 (3–22)	11.61 ± 3.60 (5–21)	0.550
Semantic fluency task	19.88 ± 5.39 (9–33)	19.02 ± 4.72 (8–30)	0.678
Number of morphemes	342.41 ± 290.38 (135–2056)	225.73 ± 112.54 (94–602)	0.007
Number of sentences	14.98 ± 9.13 (6–54)	13.28 ± 6.45 (2–34)	0.271

Data presented as mean ± standard deviation (range). PD = Parkinson's disease, NA = not available.
p-values determined by Student's *t*-test.

Table 3
Diagnostic accuracy of each group identification.

	Disease Positive (PD)	Disease Negative (HC)	Row Total	Predictive Values
1st				
NLP Positive (PD)	41	11	52	PPV = 0.79
NLP Negative (HC)	12	42	54	NPV = 0.78
Total	53 Sensitivity = 0.77	53 Specificity = 0.79		
2nd				
NLP Positive (PD)	44	10	54	PPV = 0.82
NLP Negative (HC)	9	43	52	NPV = 0.83
Total	53 Sensitivity = 0.83	53 Specificity = 0.81		
3rd				
NLP Positive (PD)	44	8	52	PPV = 0.85
NLP Negative (HC)	9	45	54	NPV = 0.83
Total	53 Sensitivity = 0.83	53 Specificity = 0.85		
4th				
NLP Positive (PD)	37	5	42	PPV = 0.88
NLP Negative (HC)	16	48	64	NPV = 0.75
Total	53 Sensitivity = 0.70	53 Specificity = 0.91		

HC = healthy control, PD = Parkinson's disease, NLP = natural language processing, PPV = positive predictive value, NPV = negative predictive value.

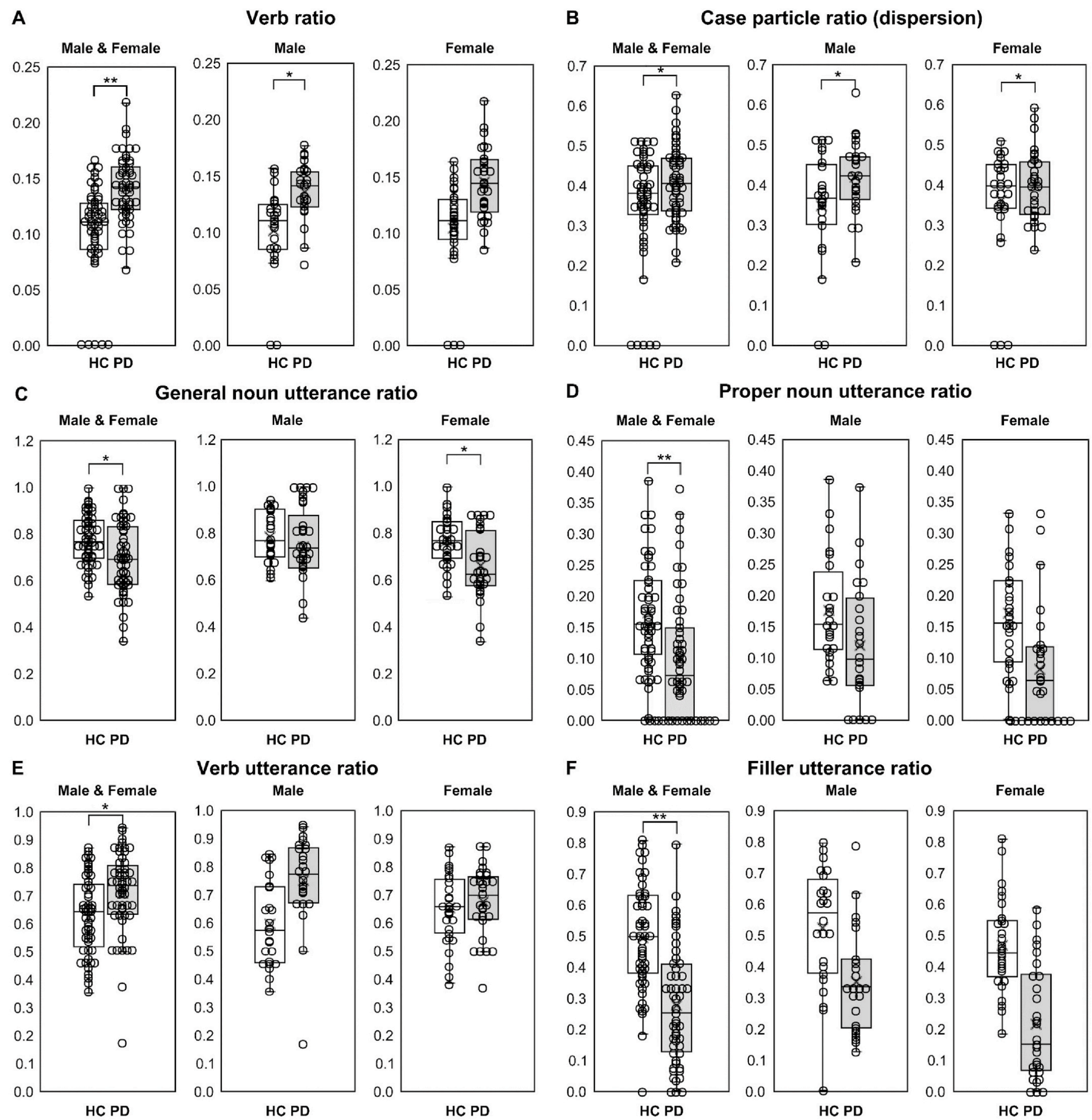

Fig. 1. Part-of-speech analysis
A–F. Box plots of values of verb ratio (A), case particle ratio (dispersion) (B), general noun utterance ratio (C), proper noun utterance ratio (D), verb utterance ratio (E), and filler utterance ratio (F).
*$p < 0.05$ and **$p < 0.001$ (Student's *t*-test). (A–E). Horizontal lines in the box plot indicate the median, and cross marks in the box plot indicate the average.
HC = Healthy control, PD = Parkinson's disease.

4. Discussion

In the present study, we clarified the characteristics of spontaneous speech of PD patients using machine learning with natural language processing. PD patients were found to spend less morphems on one sentence than the HC group. In addition, compared to HC, the utterance of PD patients had a higher rate of verbs, case particles (dispersion), and verb utterances, and a lower rate of common noun utterances, proper noun utterances, and filler utterances.

We observed no significant differences between the HC and PD groups with respect to semantic and phonemic fluency tasks, which are predictors of cognitive dysfunction in the early stages of AD [18]. Nevertheless, in the present study, even without overt cognitive decline, PD patients showed alteration in various topics of spontaneous conversation. Our results are consistent with those of previous studies which found reduced grammatical correctness and complexity in PD patients without dementia [19], and language and semantic impairment in PD patients without MCI [8]. Moreover, in a recent study, PD patients showed early differences in functional connectivity within the cortical systems that support language processing before the appearance of cognitive and language dysfunction [20].

In our conventional assessment, we found that PD patients spend less morphemes on one sentence than the HC group. This may be in line with a previous report that in the sentence generation task, the PD group was impaired in all language dimensions and overall performance, beyond the effects of cognitive ability differences [21]. The possibility that this morphological difference affected the results should be considered. The lack of correlation in the univariate analysis suggests that the difference in the number of morphemes is unrelated to the six items of part-of-speech we extracted.

Our morphological analysis using machine learning revealed the character of the speech of PD patients: verb ratio, case particle ratio (dispersion), and verb utterance ratio were higher than those in HC, while general noun utterance ratio, proper noun utterance ratio, and filler utterance ratio were lower than those in HC. Our analysis of 10-fold cross-validation using these six items showed high discrimination rates (>80%) in each trial, which is higher than the values in previous studies on AD or PD with a similar approach [10, 22]. These results suggest that speech analysis using natural language processing may potentially facilitate the diagnosis of PD, even if the language is different. We detected minor sex-related differences in part-of-speech, possibly due to data variability, but similar trends were observed in both sexes.

The increased ratio of verb indicates that the patients with PD connect many clauses in a single sentence, as verb is one of the most basic structures that makes up a sentence clause. A previous study showed that non-demented PD patients experience particular difficulty in generating verbs possibly due to dysfunction in the frontal lobes, which supports our results [23]. Multivariate analysis showed the correlation of the verb ratio with age, but not with other background factors including MoCA-J and UPDRS Part 3, suggesting that the linguistic change in speech is not closely related with general cognitive or motor function of PD. Case particles ratio (dispersion) was also increased in the PD group in the present study. Case particle in Japanese was shown to be associated with the left middle frontal gyrus and the inferior frontal gyrus [24]. As dysfunction of the frontal lobe in PD patients appears in the early stages of PD [25], the increase in case particles suggests early frontal dysfunction in PD.

In our study, the ratio of general nouns and proper nouns was lower in the PD group. In a previous study, PD patients were found to experience difficulty in the production of nouns [26]. Regional and functional neuroimaging studies suggested distinct neuroanatomical localization, with noun processing localizing to the temporoparietal lobes [27]. Others reported that in left-sided PD, noun production is significantly correlated with activity of the right uncinate fasciculus [28]. Given that the production of proper nouns is associated with the left hemisphere function [29], our results also suggest left temporal impairment in cognitively unimpaired PD.

Eexcept for the correlation between verb and age, as described above, multivariate and univariate analyses showed no correlation between each of the six selected items and age, MoCA-J, years of education, UPDRS3, LEDD, or disease duration, suggesting that the items selected in this study are not related to the severity or treatment of PD.

Some limitations of our study should be considered while interpreting the results. First, as we transcribed the speech to text, we could not analyze the voice and syllable indices in our study. Second, the sample size was small for machine learning analysis, and thus we did not separate the subjects into development and validation cohorts to investigate the diagnostic value of machine learning. Third, our study was limited by region and race, and the language was limited to Japanese. Japanese is a classifier language in which the noun is more likely to be assumed to refer to the substance, whereas English is a count/mass language in which the noun is more likely to be assumed to refer to the object kind [30]. Therefore, selection bias could not be completely eliminated. Multilinguistic studies with a larger sample size should be conducted in future. Fourth, the study used a 10-fold crossover method, but it was difficult to obtain a sufficient sample size, and validation based on unseen samples was not performed. In addition, because functional imaging was not included in the analysis, we did not evaluate the portion of the brain responsible for language deficit.

In conclusion, we found that PD patients without cognitive impairment have differences in natural language from healthy people. Specifically, PD patients speak less morphemes on one sentence in spontaneous conversation than healthy individuals. The conversations of PD patients include more verbs and case particles (dispersion) and less nouns and fillers than HC. By applying this conversational change, SVM was found to discriminate PD from HC with an accuracy of more than 80%. These results suggest the potential of natural language processing for linguistic analysis and possibly for diagnosis of PD.

Authors' roles

(1) Research project: A. Conception, B. Organization, C. Execution;
(2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique;
(3) Manuscript: A. Writing of the First Draft, B. Review and Critique.
K.Y.: 1A, 1C, 2A, 3A
Y,I.: 1C, 2A, 2B, 3B
N.K.: 1C, 2A, 2B, 3B
T.T.: 1C, 3B
K.H.: 1C.
Y.S.: 1C.
M.H.: 1C.
Y.T.: 1C.
M.S.: 1C.
A.H.: 1C.
M.K.: 1A, 1B, 2C, 3B

Funding

This work was supported by the National Centre for Geriatrics and Gerontology (NCGG) with the Research Funding for Longevity Sciences (Nos. 22–26) (M.K.); by AMED under Grant Numbers JP22ae0101077, JP22dk0207052, and JP22lk0201124 (M.K.).

Author disclosures

None of the authors report any financial interests or potential conflicts of interest.

Declaration of competing interest

None.

Acknowledgment

The authors are grateful to all the participants in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.parkreldis.2023.105411.

References

[1] L.M. de Lau, M.M. Breteler, Epidemiology of Parkinson's disease, Lancet Neurol. 5 (2006) 525–535, https://doi.org/10.1016/S1474-4422(06)70471-9.
[2] T. Simuni, K. Sethi, Nonmotor manifestations of Parkinson's disease, Ann. Neurol. 64 (2008) 65–80, https://doi.org/10.1002/ana.21505.

[3] N. Miller, Communication changes in Parkinson ' s disease, Date deposited 17 (2017) 266–274.
[4] A.B. New, D.A. Robin, A.L. Parkinson, C.R. Eickhoff, K. Reetz, F. Hoffstaedter, C. Mathys, M. Sudmeyer, C. Grefkes, C.R. Larson, L.O. Ramig, P.T. Fox, S. B. Eickhoff, The intrinsic resting state voice network in Parkinson's disease, Hum. Brain Mapp. 36 (2015) 1951–1962, https://doi.org/10.1002/hbm.22748.
[5] T. Tsuboi, H. Watanabe, Y. Tanaka, R. Ohdake, M. Sato, M. Hattori, K. Kawabata, K. Hara, D. Nakatsubo, S. Maesawa, Y. Kajita, M. Katsuno, G. Sobue, Clinical correlates of repetitive speech disorders in Parkinson's disease, J. Neurol. Sci. 401 (2019) 67–71, https://doi.org/10.1016/j.jns.2019.04.012.
[6] B. Walsh, A. Smith, Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease, Mov. Disord. 27 (2012) 843–850, https://doi.org/10.1002/mds.24888.
[7] L.J.P. Altmann, M.S. Troche, High-level language production in Parkinson's disease: a review, Parkinsons. Dis. 2011 (2011), https://doi.org/10.4061/2011/238956.
[8] Y. Bocanegra, A.M. García, D. Pineda, O. Buriticá, A. Villegas, F. Lopera, D. Gómez, C. Gómez-Arias, J.F. Cardona, N. Trujillo, A. Ibáñez, Syntax, action verbs, action semantics, and object semantics in Parkinson's disease: dissociability, progression, and executive influences, Cortex 69 (2015) 237–254, https://doi.org/10.1016/j.cortex.2015.05.022.
[9] A.M. García, D. Escobar-Grisales, J.C. Vásquez Correa, Y. Bocanegra, L. Moreno, J. Carmona, J.R. Orozco-Arroyave, Detecting Parkinson's disease and its cognitive phenotypes via automated semantic analyses of action stories, Npj Park. Dis. 8 (2022) 1–10, https://doi.org/10.1038/s41531-022-00422-8.
[10] M. Strauss, Natural Language Analysis to Detect Parkinson ' S Disease, (n.d.) 1–8..
[11] A.J. Hughes, S.E. Daniel, L. Kilford, A.J. Lees, Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases, J. Neurol. Neurosurg. Psychiatry 55 (1992) 181–184, https://doi.org/10.1136/jnnp.55.3.181.
[12] K. Ohta, K. Takahashi, J. Gotoh, K. Yamaguchi, M. Seki, Y. Nihei, S. Iwasawa, N. Suzuki, Screening for impaired cognitive domains in a large Parkinson's disease population and its application to the diagnostic procedure for Parkinson's disease dementia, Dement. Geriatr. Cogn. Dis. Extra 4 (2014) 147–159, https://doi.org/10.1159/000362124.
[13] J.M. Kang, Y.S. Cho, S. Park, B.H. Lee, B.K. Sohn, C.H. Choi, J.S. Choi, H.Y. Jeong, S.J. Cho, J.H. Lee, J.Y. Lee, Montreal cognitive assessment reflects cognitive reserve, BMC Geriatr. 18 (2018) 1–8, https://doi.org/10.1186/s12877-018-0951-8.
[14] W. Lee, D.R. Williams, E. Storey, Cognitive testing in the diagnosis of Parkinsonian disorders: a critical appraisal of the literature, Mov. Disord. 27 (2012) 1243–1254, https://doi.org/10.1002/mds.25113.
[15] J. Liang, Y. Fang, Q. Lv, H. Liu, Dependency distance differences across interpreting types: implications for cognitive demand, Front. Psychol. 8 (2017) 1–10, https://doi.org/10.3389/fpsyg.2017.02132.
[16] E. Aramaki, S. Shikata, M. Miyabe, A. Kinoshita, Vocabulary size in speech May Be an early indicator of cognitive impairment, PLoS One 11 (2016) 1–13, https://doi.org/10.1371/journal.pone.0155195.
[17] Y. Kanda, Investigation of the freely available easy-to-use software "EZR" for medical statistics, Bone Marrow Transplant. 48 (2013) 452–458, https://doi.org/10.1038/bmt.2012.244.
[18] R.M. Vaughan, R.F. Coen, R.A. Kenny, B.A. Lawlor, Semantic and phonemic verbal fluency discrepancy in mild cognitive impairment: potential predictor of progression to Alzheimer's disease, J. Am. Geriatr. Soc. 66 (2018) 755–759, https://doi.org/10.1111/jgs.15294.
[19] L.L. Murray, Spoken language production in Huntington's and Parkinson's diseases, J. Speech Lang. Hear. Res. 43 (2000) 1350–1366, https://doi.org/10.1044/jslhr.4306.1350.
[20] R. Hyder, M. Jensen, A. Højlund, L. Kimppa, C.J. Bailey, J.L. Schaldemose, M. B. Kinnerup, K. Østergaard, Y. Shtyrov, Functional connectivity of spoken language processing in early-stage Parkinson's disease: an MEG study, NeuroImage Clin 32 (2021), https://doi.org/10.1016/j.nicl.2021.102718.
[21] M.S. Troche, L.J.P. Altmann, Sentence production in Parkinson disease: effects of conceptual and task complexity, Appl. Psycholinguist. 33 (2012) 225–251, https://doi.org/10.1017/S0142716411000336.
[22] U. Petti, S. Baker, A. Korhonen, A systematic literature review of automatic Alzheimer's disease detection from speech and language, J. Am. Med. Inf. Assoc. 27 (2020) 1784–1797, https://doi.org/10.1093/jamia/ocaa174.
[23] P. Péran, O. Rascol, J.F. Démonet, P. Celsis, J.L. Nespoulous, B. Dubois, D. Cardebat, Deficit of verb generation in nondemented patients with Parkinson's disease, Mov. Disord. 18 (2003) 150–156, https://doi.org/10.1002/mds.10306.
[24] T. Inui, K. Ogawa, M. Ohba, Role of left inferior frontal gyrus in the processing of particles in Japanese, Neuroreport 18 (2007) 431–434, https://doi.org/10.1097/WNR.0b013e32805dfb7e.
[25] J.J. Tanner, T.H. Mareci, M.S. Okun, D. Bowers, D.J. Libon, C.C. Price, Temporal lobe and frontal-subcortical dissociations in non-demented Parkinson's disease with verbal memory impairment, PLoS One 10 (2015) 1–13, https://doi.org/10.1371/journal.pone.0133792.
[26] S. Di Tella, F. Baglio, M. Cabinio, R. Nemni, D. Traficante, M.C. Silveri, Selection processing in noun and verb production in left- and right-sided Parkinson's disease patients, Front. Psychol. 9 (2018) 1–13, https://doi.org/10.3389/fpsyg.2018.01241.
[27] K.A. Shapiro, L.R. Moo, A. Caramazza, Cortical signatures of noun and verb production, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 1644–1649, https://doi.org/10.1073/pnas.0504142103.
[28] S. Di Tella, F. Baglio, L. Pelizzari, M. Cabinio, R. Nemni, D. Traficante, M.C. Silveri, Uncinate fasciculus and word selection processing in Parkinson's disease, Neuropsychologia 146 (2020), 107504, https://doi.org/10.1016/j.neuropsychologia.2020.107504.
[29] S.Y. Yang, D. Van Lancker Sidtis, Cerebral processing of proper and common nouns: perception and production following left hemisphere damage, Clin. Linguist. Phon. 29 (2015) 319–332, https://doi.org/10.3109/02699206.2014.991450.
[30] M. Imai, D. Gentner, A cross-linguistic study of early word meaning: universal ontology and linguistic influence, Cognition 62 (1997) 169–200, https://doi.org/10.1016/S0010-0277(96)00784-6.