Verb Motility Dynamics Reveals Cognitive Impairment in Parkinson's Disease: A Speech-Language Fusion Approach

Jhon Fredy Mercado-Agudelo[1⋆[0009−0000−4867−4352]], Daniel Escobar-Grisales[1*[0000−0002−3257−0134]], Cristian David Ríos-Urrego[1[0000−0003−0174−1452]], Adolfo M. García[3[0000−0002−6936−0114]], Yamile Bocanegra[3[0000−0003−3064−2509]], Leonardo Moreno[4], Elmar Nöth[5[0000−0002−3396−555X]], and Juan Rafael Orozco-Arroyave[1,5[0000−1111−2222−3333]]

[1] GITA Lab, Faculty of Engineering, Universidad de Antioquia, Medellín, Colombia
[2] CNC, Universidad de San Andrés, Argentina
[3] GNA Lab, Universidad de Antioquia, Colombia
[4] Hospital Pablo Tobón Uribe, Colombia
[5] LME Lab, University of Erlangen-Nuremberg, Erlangen, Germany

Abstract. Parkinson's disease (PD) and mild cognitive impairment (MCI) affect both motor and cognitive linguistic abilities. This work proposes a multimodal framework for detecting PD and MCI from naturalistic retelling tasks by combining acoustic and linguistic features. Speech features capture prosodic, articulatory, and phonemic properties linked to hypokinetic dysarthria, while language features model lexical,syntactic, and semantic complexity, including a novel motility-based representation that quantifies the use of action-related verbs.
Each modality is independently evaluated and combined following early and late fusion strategies based on support vector machines. The results confirm speech features as good biomarkers to model PD and MCI, and show the motility-based language features improve specificity, particularly in distinguishing cognitive decline in Parkinson's patients, i.e., PD patients with MCI vs. patients without MCI. Fusion strategies further improve classification performance, confirming the complementarity of speech and language. These findings support the use of retelling-based multimodal analyses as promising tools for early and non-invasive screening of neurodegenerative conditions.

Keywords: Parkinson's disease · Mild cognitive impairment · Multimodal analysis · Speech and language modeling · Support vector machine · Feature fusion.

1 Introduction

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder worldwide after Alzheimer's disease [15]. PD is characterized by different

⋆ These authors contributed equally.

motor and non-motor symptoms. Motor alterations include tremors, bradykinesia, and rigidity. Non-motor signs include cognitive decline, sleep disorders, and others. The manifestation and progression of these symptoms vary among individuals. One of the main challenges in PD is its diagnosis, which is expensive and requires neurologists experts to screen the patient. Speech and language biosignals emerged as promising biomarkers thanks to their non-invasive and cost-effective nature, and also because they easily enable remote screenings.

Speech impairments suffered by PD patients are typically grouped and called hypokinetic dysarthria. It affects different dimensions of speech production, including respiration, phonation, articulation, and prosody [26, 19]. Classical approaches include measures like pitch, jitter, shimmer, formant frequencies, and Mel Frequency Cepstral Coefficients (MFCCs) [16]. Other studies have focused on representations that aim to model specific speech dimensions [26]. In [17], the authors used different feature sets to model phonation, prosody, and articulation. These features were used to estimate the PD severity based on diadochokinetic (DDK) tasks performed by PD patients of the PC-GITA database [18]. Patients were categorized into four severity levels (Normal, Slight, Mild, Moderate) based on the MDS-UPDRS-III score. The authors reported an overall accuracy of 72% in discriminating the four severity levels. Other works have considered Deep Learning (DL) approaches to analyze PD using speech signals. In [4], the authors considered 246 Italian speakers (160 with PD) who were asked to produce the sustained vowel /e/. In the study, traditional methods were compared with methods based on DL. The traditional approach considered hand-crafted features like MFCCs, jitter, shimmer, and formant frequencies, and achieved accuracies of up to 82%. The DL approach was based on Mel-spectrograms as input to a CNN and achieved accuracies of up to 70%. Contrary to the typically reported in the literature, the hand-crafted features outperformed the DL approach. This is likely because the authors did not incorporate data augmentation or transfer learning methods to improve data size and therefore better optimize their models. In [27], the authors used classical feature sets and autoencoders. These features were used to classify between PD patients and HCs using sustained vowels in a Chinese dataset. The authors reported an accuracy of 85% and indicated the autoencoders-based features as the most discriminant ones. In [9], the authors considered a corpus composed of 80 Spanish-speaking subjects (40 PD patients) and extracted prosody, articulation, and phonemic identifiability features to classify between PD patients and HCs. The fusion of all features yielded an accuracy of up to 84%, where articulation features contributed the most to discriminate between PD patients and HCs. In addition, the authors classified between PD patients with and without MCI and found that phonemic identifiability features effectively discriminate between these two subgroups. Finally, a recent study proposed the use of foundation models, such as Wav2vec and WavLM, to model PD. In [6], the authors explored four characterization strategies: three pathology-oriented representations (articulation, prosody, and phonemic identifiability) and one non-pathology-oriented representation based on the Wav2Vec

2.0 model. Accuracies of up to 81% were reported when pathology-oriented representations were incorporated, outperforming the Wav2vec 2.0 results.

On the other hand, although language production has been less studied in the context of PD, some works have reported different language patterns associated with this disease. PD is characterized by various lexico-semantic deficits, including reduced verbal fluency, alterations in the use and production of motor verbs (i.e., verbs denoting bodily movements), and deficits related to learning new verbs [8, 14]. In [8], three corpora in different languages were considered: Spanish [18], German [24], and Czech [22]. The authors considered morphological tagging and graphs to model the three corpora and reported accuracies of up to 80% when classifying bertween PD and HC subjects. In [10], the authors used the same database as in [9] and focused on manual transcriptions of the retelling tasks. They proposed the Proximity-to-Reference Semantic Field (P-RSF) characterization to capture the weight of action and non-action concepts across retold texts. Accuracies of up to 85% were obtained in the classification of PD patients and HC when analyzing the retelling of an action-based story. The P-RSF score showed a significant difference between PD patients with MCI vs. HC. However, when comparing P-RSF scores between PD patients with and without MCI, no significant difference were observed. DL approaches have also been explored in language analysis. Representations based on word-embeddings, such as Word2vec or Bidirectional Encoder Representations from Transformers (BERT), have been used to analyze language abnormalities in PD. Although these general representations were not explicitly developed for mapping PD traits, some of them have demonstrated correlations with certain cognitive tests and achieved accuracies of up to 72% in classifying PD patients and HCs [12, 21]. In [6], the authors computed language representations based on BETO (a Spanish version of BERT [3]) and reported the best results with a Verbs+CNN representation. This finding aligns with the literature, where different non-automatic approaches have reported that PD patients show difficulties processing verbs compared to other grammatical units, such as nouns [1, 11]. Previous studies have shown impaired verb production in PD patients compared to HCs, particularly in the production of verbs that denote action or movement. However, the studies have focused on the frequency of such verbs in discourse or on their semantic proximity to reference action-related fields. This paper aims to analyze the temporal evolution of the motor content in verbs produced during narrative retelling to assess whether the dynamic patterns in the motor content of verbs serve as a sensitive linguistic biomarker of alterations associated with PD. This paper considers the story with high-action content included in [9]. Our analysis includes both speech and language modalities: we extract three well-established speech representations as well as a new language-based feature, which aims to capture the dynamics of motor content in the verbs produced during the retelling. We evaluated all representations in four classification scenarios: (i) PD vs. HC, (ii) PD with MCI (PD-MCI) vs. HC, (iii) PD without MCI (PD-nMCI) vs. HC, and (iv) PD-MCI vs. PD-nMCI. While speech-based features yielded good results, in line with previous literature, the proposed motility-based characterization improved performance

in scenarios involving cognitive status. These findings suggest the potential of motor-content dynamics to capture subtle linguistic alterations associated with cognitive impairment in PD.

2 Material and methods

2.1 Data

We used the same dataset referenced in [9] and [10], which includes speech recordings from 40 HC subjects and 40 early-stage Spanish-speaking PD patients, 16 PD-MCI and 24 PD-nMCI. The dataset has two subsets of HCs matched matched by gender and age with respect to their corresponding subsets of PD patients. Although the corpus includes two stories, we focused exclusively on the retelling of the high motor content story. This enables us to jointly model speech and language [10]. Classification of PD-MCI and PD-nMCI was developed based on their MoCA scores and their level of functional independence. Original recordings were sampled at 44.1 kHz. For this study, we applied the channel normalization procedure described in [2] and the recordings were down sampled to 16 kHz. The demographic and clinical characteristics of the participants are summarized in Table 1.

Table 1. Clinical and demographic information. (F/M): Female/Male. UPDRS-III: Unified Parkinson's Disease Rating Scale, section III, H&Y: Hoehn & Yahr scale, MoCA: Montreal Cognitive Assessment. MCI screening followed level-1 criteria of the Movement Disorder Society Task Force [13].

	PD patients (n = 40)	**HC subjetcs** (n = 40)	**PD vs. HCs**
Sociodemographic variables			
Sex (F/M)	15/25	15/25	–
Age (years)	62.3± 9.3	61.9±7.3	0.84*
Education (years)	12.2± 5.0	12.8±4.6	0.63*
Clinical variables			
Years since diagnosis	5.7± 3.7	–	–
UPDRS-III	31.0±12.5	–	–
Hoehn & Yahr stage	2.1± 0.3	–	–
MoCA	24.8± 3.0	26.7±1.6	< 0.01*
wMCI/nMCI	16/24	–	–

* $p-$values computed using Mann–Whitney U tests.

2.2 Methodology

The general methodology addressed in this work is shown in Figure 1. We process speech recordings from a retelling task to extract acoustic and linguistic informa-

tion. These modalities are first modeled independently and then combined using fusion strategies. The methodology involves three main stages: speech modeling using prosodic, articulatory, and phonemic features; language modeling based on dynamic analysis of the motor content in the verbs produced during the retelling; and fusion of speech and language representations. The proposed approach was evaluated in four binary classification tasks: (a) PD patients and HCs, (b) PD-MCI and HCs, (c) PD-nMCI and HCs, and (d) PD-MCI and PD-nMCI.

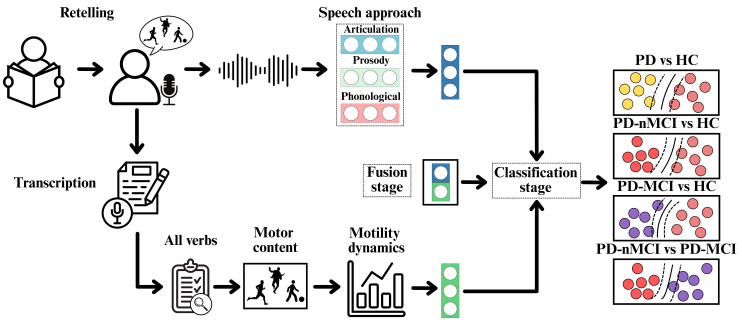

Fig. 1. Proposed methodology.

2.3 Speech Modeling

We extracted three sets of acoustic features following the methodology proposed in [9]. These features are designed to capture the distinct manifestations of hypokinetic dysarthria, which affects precision of articulation, prosodic variability, and phonemic identifiability. These features were extracted using the Disvoice[6] Toolkit introduced by GITA Lab years ago.

- **Articulation:** These features aim to capture deficits in speech motor control that affect the clarity and sharpness of phoneme transitions. These features are computed around voiced-unvoiced boundaries and include Bark-scale spectral energies and MFCCs extracted from fixed-size windows centered on transition points. Such features reflect articulatory undershoot and reduced precision, which are common in PD [20, 26].
- **Prosody:** These features describe the suprasegmental characteristics of speech, including pitch variation, jitter, shimmer, speech rate, and pause duration. These features quantify the rhythmic and melodic aspects of speech, which are commonly affected in PD due to reduced respiratory and laryngeal control [5, 26].
- **Phonemic indetifiability:** This feature set assess the precision of phonological classes to be pronounced in the acoustic signal. They are computed

[6] https://github.com/jcvasquezc/DisVoice

using a Gated Recurrent Unit (GRU) architecture trained on speech from healthy Latin American Spanish speakers. The model generates posterior probabilities for 18 predefined phonological classes over time. The core assumption is that individuals with PD exhibit lower confidence scores or incorrect predictions, in contrast to HCs which tend to produce clearer and more identifiable phonological targets [25].

2.4 Motor Content Modeling

First we applied part-of-speech tagging to each transliteration to extract all verbs, which were then lemmatized. Each verb is mapped to a normalized motor content score using a reference database composed of 4,565 Spanish verbs [23], where a motility score was assigned to each verb to indicate the degree of motor content. We applied min-max normalization to standardize the scores in the reference dataset. Only verbs of the transliteration found in the reference database were retained, resulting in a subject-specific motor-content vector. From this motor content we extracted 46 linguistic features, which can be grouped into three categories: (i) Motor content distribution, (ii) dynamic structure, and (iii) segmented motor patterns.

- **Motor content distribution:** This feature set aims to capture the overall distribution of motor content across the retelling. Nine statistical descriptors are computed over the raw motor content vector: mean, standard deviation, maximum, minimum, skewness, kurtosis, inter quartile range, number of peaks, and number of valleys, as well as the length of the vector and the slope of a first-degree polynomial fit.
- **Dynamic structure:** This set aims to model how motor content changes over time, including fluctuations, transitions, and irregularities. We computed an entropy vector using a sliding window of three verbs with a step size of one verb. For each window, we calculated the Shannon entropy of the motility values, resulting in a new time series that captures short-term information content of the motility dynamics. We then extracted the same nine statistical functionals mentioned above. In addition, we computed the first and second derivatives of the original motor content vector and extracted the nine functionals from each time series.
- **Motor content segmentation:** In this feature subgroup we defined motility blocks as sequences of consecutive verbs with a motility score greater than 0.75. From these blocks, we computed eight features, including the number of blocks, average block size, the position of the blocks with highest and lowest motility, motility of the first, central, and last verbs, and the difference between the first and last motility values.
  A summary of all features is presented in Table 2.

2.5 Bimodal analysis

We combined the representations from each modality using a late fusion strategy, similar to the one used in [7]. In this fusion strategy, the classification scores from

Table 2. Features extracted from the motor content vector.

Category	Feature description	Count
Motor content distribution	9 statistics over raw motor content vector	9
	Length of motor vector	1
	Slope of 1st-degree polynomial fit	1
Dynamic structure	9 statistics over sliding-window entropy vector	9
	9 statistics over first derivative	9
	9 statistics over second derivative	9
Motor content segmentation	Number of high-motility blocks, average block size	2
	Position of max/min block, initial/central/final motility, edge difference	6
Total		**46**

the individual speech and language models are used as inputs to a second-level classifier, which performs the final bimodal classification.

3 Experiments and results

This section presents the experiments and results obtained in the four binary classification scenarios considered in this paper. For each experiment, HC participants were matched with the corresponding PD patients subgroup to maintain demographic and clinical balance[7]. We used an SVM with a radial basis function (RBF) and a linear kernel for classification. The regularization parameter C and kernel width γ were tuned in a two-stage process. First, a coarse grid search was conducted over $C \in \{0.001, 0.01, .., 100\}$ and $\gamma \in \{0.0001, 0.001, ..., 100\}$ to identify a candidate region of the hyperparameter space. Then, a randomized search was performed within this region, using exponentially scaled distributions to improve generalization while mitigating overfitting.

Figure 2 shows the ROC curves and corresponding AUCs using the speech representation described in Section 2.3. As expected, in the experiment between PD patients (with mixed cognitive profiles) and HCs, the combination of all speech dimensions yielded the highest AUC (0.79), mainly driven by articulatory features (AUC = 0.79). In the HC vs PD-MCI scenario, the phonemic identifiability representation achieved the best performance with an AUC of 0.95 followed by the combination of all features with an AUC of 0.89. For the PD-nMCI vs. HC classification, the combination of all speech dimensions achieved an AUC of 0.69, mainly driven by articulatory and phonological features. Finally, in the PD-nMCI vs. PD-MCI scenario, articulation features achieved the best results, slightly outperforming the phonological representation and the representation base on the combination of all dimensions.

These results align with the findings reported in [9], which showed that the retelling task yielded better performance in cognitively focused comparisons, such as PD-MCI vs. HC and PD-nMCI vs. PD-MCI, while the reading task performed better in broader motor speech comparisons such as PD vs. HC and

[7] One participant classified as HC-nMCI was excluded due to a missing recording.

HC vs. PD-MCI. This pattern reinforces the notion that retelling tasks are more effective in capturing cognitive-linguistic impairments due to their demands on memory, language organization, and spontaneous verbal expression. In contrast, reading tasks are more sensitive to motor speech alterations through their structured and repetitive nature. Our results further confirm that articulatory features consistently contribute to the best performance in scenarios involving patients with cognitively intact or mixed profile PD patients, supporting their reliability as robust markers of dysarthric speech. In contrast, phonemic identifiability features, a subset of phonological features, emerge as good biomarkers in tasks involving cognitive impairment, suggesting that changes in speech production and planning are strong indicators of cognitive decline in PD. In particular, in the cognitively nuanced comparison of PD-nMCI vs. PD-MCI, both articulatory and phonological features were found to be the most effective, underscoring their combined value in capturing the intersection of motor and cognitive symptoms during disease progression.

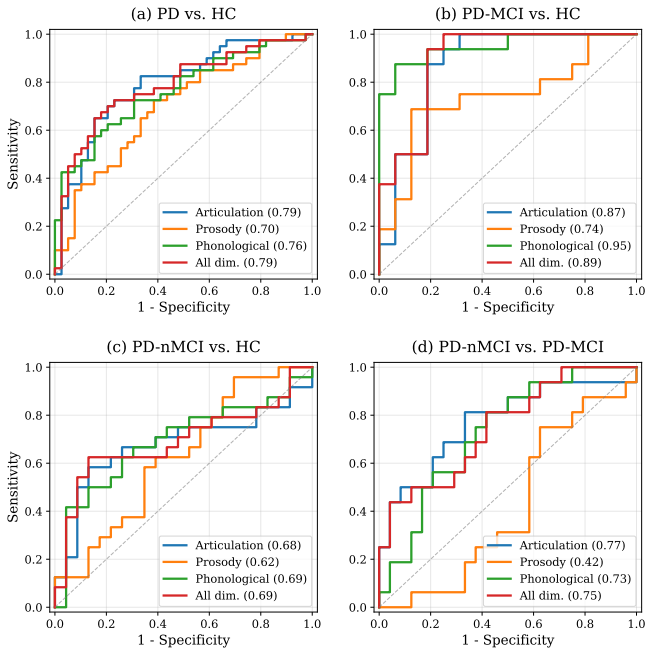

Fig. 2. ROC curves comparing different feature sets for four classification tasks: (a) PD vs. HC, (b) PD-MCI vs. HC, (c) PD-nMCI vs. HC, (d) PD-nMCI vs. PD-MCI. Each curve includes its corresponding AUC in the legend.

Table 3 show the best classification results obtained for each modality independently and for the bimodal approach based on late fusion. The language modality corresponds to the representation based on the dynamic analysis of

Table 3. Classification between PD patients vs. HC subjects for the uni-modal and multimodal approaches.

Modality	Accuracy	Sensitivity	Specificity	F1-score
PD vs. HC				
Speech(all dim)	74.7	67.5	82.1	73.0
Language	59.5	60.0	59.3	60.0
Fusion(all dim + language)	65.8	67.5	64.1	66.7
PD-MCI vs. HC				
Speech(Phonological)	87.5	81.2	93.8	86.7
Language	75.0	75.0	75.0	75.0
Fusion	90.6	87.8	93.8	90.3
PD-nMCI vs. HC				
Speech(Articulation)	70.2	62.5	78.3	68.2
Language	59.6	58.3	60.9	59.6
Fusion(all dim + language)	66.0	70.8	60.9	68.0
PD-nMCI vs PD-MCI				
Speech(Phonological)	72.5	50.0	87.5	59.3
Language	70.0	62.5	75.0	66.7
Fusion(all dim + language)	75.0	68.8	79.2	68.8

All dim: the speech representation obtained concatenating prosody, articulation, and phonological features.

verb-related motor content. The biomodal approach consists in combining each acoustic representation with the language representation. For each modality, we only report the best result according to accuracy.

These results show a limited performance of the language modality in the classification scenario that involves cognitively heterogeneous patients or in experiments that do not involve PD patients with cognitive impairment. Conversely, classification scenarios involving PD patients with cognitive impairment show acceptable results with the language modality. An accuracy of 75% is obtained in the PD-MCI vs. HC scenario. Although this is lower than the accuracy obtained using acoustic features (phonological), the language-based representation provided a more balanced sensitivity and specificity. In the PD-nMCI vs. PD-MCI classification, the language modality outperformed the sensitivity obtained with speech, leading to a higher F1-score.

Results from the language modality contrast with findings from previous studies, where features based on the production of motor-related verbs were primarily associated with PD related linguistic impairment rather than cognitive decline. In our case, the dynamic analysis of verb motor content does not appear to capture the linguistic deterioration typically associated with PD. However, our language representation shows promising patterns as a biomarker of cognitive impairments. We hypothesize that features such as the variability of motor content and the transitions in motor content throughout the narrative reflect lexical richness and the subject's ability to vary or sustain motor-semantic in-

formation, which we expect to be cognitive-linguistic functions more sensitive to cognitive decline than linguistic impairments associated with PD.

Finally, the bimodal approach showed a improved performance in the PD-MCI vs. HC and PD-nMCI vs. PD-MCI classification tasks. The first scenario yielded accuracies of up to 90%, with a more balanced trade-off between sensitivity and specificity. To the best of our knowledge, this represents the highest reported performance for this classification scenario using this dataset. A similar behavior was observed in the PD-nMCI vs. PD-MCI task, where accuracy improved and a better balance between sensitivity and specificity was observed.

4 Discussion and conclusions

This work demonstrates the effectiveness of using retelling speech tasks to detect cognitive decline in PD patients. Acoustic features improve sensitivity, while language features, particularly those capturing motility dynamics, enhance specificity. Fusion strategies strengthen overall classification, highlighting the potential of combining speech and language in naturalistic tasks for early, non-invasive screening of neurodegenerative conditions.

Our novel motility-based language feature, derived from action-oriented verbs in the retelling task, improved classification in cognitively focused tasks, especially distinguishing PD-MCI from PD-nMCI. These results suggest that cognitive impairment in PD alters not only speech but also language structure, particularly in the use of action-related verbs.

Although language modeling alone was less effective for PD detection, it improved specificity in combination with speech. Early fusion improved general PD detection, while late fusion highlighted fine-grained cognitive classifications. The combination of speech and language provides a more robust representation, capturing both motor and cognitive linguistic dimensions of PD and MCI.

5 Acknowledgements

This work was funded by UdeA grant # PI2023-58010 and IAPFI23-1-01.

References

1. Abrevaya, S., Sedeño, L., Fitipaldi, S., Pineda, D., Lopera, F., Buritica, O., Villegas, A., Bustamante, C., Gomez, D., Trujillo, N., et al.: The road less traveled: alternative pathways for action-verb processing in parkinson's disease. Journal of Alzheimer's disease **55**(4), 1429–1435 (2016)
2. Arias-Vergara, T., Klumpp, P., Vasquez-Correa, J.C., Nöth, E., Orozco-Arroyave, J.R., Schuster, M.: Multi-channel spectrograms for speech processing applications using deep learning methods. Pattern Analysis and Applications **24**, 423–431 (2021)
3. Cañete, J., Chaperon, G., Fuentes, R., Ho, J.H., Kang, H., Pérez, J.: Spanish pre-trained bert model and evaluation data. arXiv preprint arXiv:2308.02976 (2023)

4. Costantini, G., Cesarini, V., Di Leo, P., Amato, F., Suppa, A., Asci, F., Pisani, A., Calculli, A., Saggio, G.: Artificial intelligence-based voice assessment of patients with parkinson's disease off and on treatment: machine vs. deep-learning comparison. Sensors **23**(4), 2293 (2023)
5. Dehak, N., Dumouchel, P., Kenny, P.: Modeling prosodic features with joint factor analysis for speaker verification. IEEE Transactions on Audio, Speech, and Language Processing **15**(7), 2095–2103 (2007)
6. Escobar-Grisales, D., Arias-Vergara, T., Rios-Urrego, C.D., Nöth, E., García, A.M., Orozco-Arroyave, J.R.: An automatic multimodal approach to analyze linguistic and acoustic cues on parkinson's disease patients. In: Proceedings of the INTERSPEECH (2023)
7. Escobar-Grisales, D., Ríos-Urrego, C.D., Orozco-Arroyave, J.R.: Deep learning and artificial intelligence applied to model speech and language in parkinson's disease. Diagnostics **13**(13), 2163 (2023)
8. Eyigoz, E., Courson, M., Sedeño, L., Rogg, K., Orozco-Arroyave, J.R., Nöth, E., Skodda, S., Trujillo, N., Rodríguez, M., Rusz, J., et al.: From discourse to pathology: Automatic identification of parkinson's disease patients via morphological measures across three languages. Cortex **132**, 191–205 (2020)
9. García, A.M., Arias-Vergara, T., C Vasquez-Correa, J., Nöth, E., Schuster, M., Welch, A.E., Bocanegra, Y., Baena, A., Orozco-Arroyave, J.R.: Cognitive determinants of dysarthria in parkinson's disease: an automated machine learning approach. Movement Disorders **36**(12), 2862–2873 (2021)
10. García, A.M., Escobar-Grisales, D., Vásquez Correa, J.C., Bocanegra, Y., Moreno, L., Carmona, J., Orozco-Arroyave, J.R.: Detecting parkinson's disease and its cognitive phenotypes via automated semantic analyses of action stories. npj Parkinson's Disease **8**(1), 163 (2022)
11. García, A.M., DeLeon, J., Tee, B.L.: Neurodegenerative disorders of speech and language: Non-language-dominant diseases. In: Della Sala, S. (ed.) Encyclopedia of Behavioral Neuroscience, 2nd edition (Second Edition), pp. 66–80. Elsevier, Oxford, second edition edn. (2022). https://doi.org/https://doi.org/10.1016/B978-0-12-819641-0.00042-6, https://www.sciencedirect.com/science/article/pii/B9780128196410000426
12. Jessiman, L., Murray, G., Braley, M.: Language-based automatic assessment of cognitive and communicative functions related to parkinson's disease. In: Proceedings of the First International Workshop on Language Cognition and Computational Models. pp. 63–74 (2018)
13. Litvan, I., Goldman, J.G., Tröster, A.I., et al.: Diagnostic criteria for mild cognitive impairment in parkinson's disease: Movement disorder society task force guidelines. Movement disorders **27**(3), 349–356 (2012)
14. Liu, L., Luo, X.G., Dy, C.L., Ren, Y., Feng, Y., Yu, H.M., Shang, H., He, Z.Y.: Characteristics of language impairment in parkinson's disease and its influencing factors. Translational Neurodegeneration **4**, 1–8 (2015)
15. Marinus, J., Zhu, K., Marras, C., Aarsland, D., van Hilten, J.J.: Risk factors for non-motor symptoms in parkinson's disease. The Lancet Neurology **17**(6), 559–568 (2018)
16. Mekyska, J., Smekal, Z., Galaz, Z., Mzourek, Z., Rektorova, I., Faundez-Zanuy, M., López-de Ipiña, K.: Perceptual features as markers of parkinson's disease: the issue of clinical interpretability. Recent advances in nonlinear speech processing pp. 83–91 (2016)

17. Oliveira, G.C., Pah, N.D., Ngo, Q.C., Yoshida, A., Gomes, N.B., Papa, J.P., Kumar, D.: A pilot study for speech assessment to detect the severity of parkinson's disease: An ensemble approach. Computers in Biology and Medicine **185**, 109565 (2025)
18. Orozco-Arroyave, J.R., Arias-Londoño, J.D., Vargas-Bonilla, J.F., Gonzalez-Rátiva, M.C., Nöth, E.: New spanish speech corpus database for the analysis of people suffering from parkinson's disease. In: Lrec. pp. 342–347 (2014)
19. Orozco-Arroyave, J.R., Vásquez-Correa, J.C., Nöth, E.: Current methods and new trends in signal processing and pattern recognition for the automatic assessment of motor impairments: The case of parkinson's disease. In: Neurological Disorders and Imaging Physics, Volume 5: Applications in dyslexia, epilepsy and Parkinson's, pp. 8–1. IOP Publishing Bristol, UK (2020)
20. Orozco-Arroyave, J.R., Vásquez-Correa, J.C., Vargas-Bonilla, J.F., Arora, R., Dehak, N., Nidadavolu, P.S., Christensen, H., Rudzicz, F., Yancheva, M., Chinaei, H., et al.: Neurospeech: An open-source software for parkinson's speech analysis. Digital Signal Processing **77**, 207–221 (2018)
21. Pérez-Toro, P.A., Vásquez-Correa, J.C., Strauss, M., Orozco-Arroyave, J.R., Nöth, E.: Natural language analysis to detect parkinson's disease. In: Text, Speech, and Dialogue: 22nd International Conference, TSD 2019, Ljubljana, Slovenia, September 11–13, 2019, Proceedings 22. pp. 82–90. Springer (2019)
22. Rusz, J., Cmejla, R., Tykalova, T., Ruzickova, H., Klempir, J., Majerova, V., Picmausova, J., Roth, J., Ruzicka, E.: Imprecise vowel articulation as a potential early marker of parkinson's disease: effect of speaking task. The Journal of the Acoustical Society of America **134**(3), 2171–2181 (2013)
23. San Miguel Abella, R.A., González-Nosti, M.: Motor content norms for 4,565 verbs in spanish. Behavior Research Methods **52**, 447–454 (2020)
24. Skodda, S., Grönheit, W., Schlegel, U.: Intonation and speech rate in parkinson's disease: General and dynamic aspects and responsiveness to levodopa admission. Journal of Voice **25**(4), e199–e205 (2011)
25. Vásquez-Correa, J.C., Klumpp, P., Orozco-Arroyave, J.R., Nöth, E.: Phonet: A tool based on gated recurrent neural networks to extract phonological posteriors from speech. In: Interspeech. vol. 60, p. 61 (2019)
26. Vásquez-Correa, J.C., Orozco-Arroyave, J., Bocklet, T., Nöth, E.: Towards an automatic evaluation of the dysarthria level of patients with parkinson's disease. Journal of communication disorders **76**, 21–36 (2018)
27. Wang, M., Zhao, X., Li, F., Wu, L., Li, Y., Tang, R., Yao, J., Lin, S., Zheng, Y., Ling, Y., et al.: Using sustained vowels to identify patients with mild parkinson's disease in a chinese dataset. Frontiers in Aging Neuroscience **16**, 1377442 (2024)