INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

An Automatic Multimodal Approach to Analyze Linguistic and Acoustic Cues on Parkinson's Disease Patients

Daniel Escobar-Grisales[1], *Tomás Arias-Vergara*[1,2], *Cristian David Rios-Urrego*[1], *Elmar Nöth* [2], *Adolfo M. García*[3,4,5], *Juan Rafael Orozco-Arroyave*[1,2]

[1]GITA Lab, Faculty of Engineering, University of Antioquia UdeA, Medellín, Colombia
[2]Pattern Recognition Lab. Friedrich-Alexander Universität, Erlangen-Nürnberg, Erlangen, Germany
[3]Global Brain Health Institute, University of California, San Francisco, USA
[4]Cognitive Neuroscience Center, Universidad de San Andrés, Buenos Aires, Argentina
[5]Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile

daniel.esobar@udea.edu.co

Abstract

Early detection and monitoring of Parkinson's disease are crucial for properly treating and managing the symptoms. Automatic speech and language analysis has emerged as a promising non-invasive method to monitor the patient's state. This study analyzed different speech and language representations for automatic classification between Parkinson's disease patients and healthy controls. First, each modality is analyzed independently. General representations such as Wav2vec or BETO are used together with representations oriented to model disease traits such as phonemic identifiability in speech modality and grammatical units analysis in language modality. The best speech and language representations were combined using a fusion strategy based on Gated Multimodal Units. The best results are achieved with the multimodal approach, outperforming all results obtained with unimodal representations and the traditional fusion strategy.

Index Terms: Speech analysis, language analysis, multimodal

1. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder that is characterized by the gradual loss of dopaminergic neurons in the mid-brain [1], resulting in various motor and non-motor symptoms, including tremors, bradykinesia, cognitive decline, depression, and others [1, 2]. The early detection of PD and the monitoring constant is crucial for properly treating and managing the disease. While there are various methods for detecting PD, the analysis of speech and language has emerged as a promising approach due to its non-invasive and cost-effective nature. Typically, PD patients develop hypokinetic dysarthria, which is a speech disorder characterized by reduced voice quality, mono loudness, monotonicity, imprecise pronunciation of both consonants and vowels, and others [3, 4]. Furthermore, PD is characterized by various lexico-semantic abnormalities, including reduced verbal fluency, significant impairment in producing spontaneous speech [5], alterations in the usage and production of motor verbs (i.e., verbs denoting bodily movements)[6, 7], and deficits related to learning new verbs[8]. The automatic analysis of PD has focused more extensively on speech than on language, mainly due to the fact that speech impairments are a hallmark symptom of the disease, and they tend to be more noticeable and have a more significant impact on the patient's quality of life. However, recent studies have shown that language analysis can provide a better understanding of the cognitive difficulties experienced by people with PD.

Pathological speech in PD has generated significant interest in recent years. Different studies have proposed approaches based on Deep Learning (DL) models to detect speech impairments and predict PD progression. Some studies considered models based on Convolutional Neural Networks (CNN) using spectrograms as input to classify PD patients and Healthy Controls (HCs) or to detect dysarthria and predict its severity level [9, 10]. Recent works implemented a model combining unidimensional-CNN (1D-CNN) and bidimensional-CNN (2D-CNN) to capture frequency and time information [11, 12]. Other works have focused on representations that aim to model acoustic cues of PD; in [13, 14], authors used different feature sets to model different speech dimensions such as phonation, prosody, and articulation, and then this features sets are used to classify PD patients and HCs. A similar approach was proposed in [15], where the authors modeled the phoneme articulation precision in PD patients using features set of phonetic information. Finally, in a recent work [16], prosody, articulation, and phonemic information features were combined to classify PD patients and HCs. The features based on phonemic information were used to discriminate PD patients with cognitive impairment from control subjects, with an accuracy of 87%. The approach also distinguished between cognitively spared and impaired patients with accuracies of up to 72%.

Several studies have explored the automatic analysis of language abnormalities in PD using Natural Language Processing (NLP) techniques such as: Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), Global Vector Representations (GloVe), Word2vec, and others. Although these general representations were not developed explicitly for mapping PD traits, some of them have demonstrated correlations with certain cognitive tests and achieved to classify PD patients and HCs with accuracies of up to 72% [17, 18]. Other studies aim to create more specific representations to model linguistic cues of the pathology. In [19] 17 PD patients and 15 HCs listened to verbs and nouns during functional Magnetic Resonance Imaging (MRI) scans. The authors found no connectivity differences between PD and HC during noun listening, but functional connectivity differences were found during action-verb processing. Thus, verb production analysis in PD patients can be useful to model different traits of the pathology. Similarly, in [20], the authors aimed to classify PD patients and HC using retellings of action and non-action stories produced by 80 participants (40 PD patients). In each retelling transliteration, authors weighted action and non-action concepts using a Proximity-to-Reference-Semantic-Field (P-RSF) metric, which was computed using Latent Semantic Analysis (LSA). These features were used to train an SVM as a classifier, and accura-

10.21437/Interspeech.2023-2287

cies of up to 85% were obtained using the retelling of action stories. Results of that work indicate that PD patients exhibit an impaired ability to process action concepts compared with non-action concepts. The main limitation of this work was that the P-RSF metric depends on a text-specific task. Therefore, this approach cannot be generalized to tasks such as monologue. The fusion between speech and language has been poorly explored in the automatic PD analysis. In [21], depression in people with PD was modeled using speech and language representations. The authors utilized Bidirectional Encoder Representations from Transformer (BERT) to model the transliterations of monologues and acoustic features such as Bark band energies and Mel frequency cepstral coefficients to model the corresponding speech signal. Both representations were combined using an early fusion strategy, where F1-scores of up to 77% were obtained in the classification of depressed and non-depressed PD patients.

This paper compares unimodal and multimodal approaches to classify PD patients and HCs based on speech and language analysis. For each modality, we consider two approaches: general and uninterpretable representations and pathology-oriented representations, which aim to characterize typical traits of PD. Speech recordings are analyzed using representations extracted from Wav2vec 2.0 model and representations that consider different speech dimensions, such as prosody, articulation, and phonemic information. In language, we explored typical representations based on BETO (BERT model trained with Spanish corpus), such as the statistical functionals of the word-embeddings that compose the document and representations pathology-oriented, where a 1D-CNN is used to analyze the deficit of PD patients to process grammatical units such as verbs and nouns. Finally, we combine both information sources using Gated Multimodal Units (GMUs), which combines the advantages of early and late fusion to find an intermediate representation based on a combination of data from both modalities. The remainder of this paper is organized as follows. Section 2 presents the materials and methods used in this study. Section 3 describes the experimental setup and the results. Finally, Section 4 contains the conclusions and future work.

2. Materials and methods

2.1. Data

The study includes 165 participants, divided into two groups: 80 PD patients and 85 HCs. All participants are native Spanish speakers from Colombia, and the groups are balanced in age and gender. These recordings are part of an extended version of the PC-GITA corpus [22]; they were normalized using a GSM full-rate compression technique and down-sampled to 8 kHz [23]. A neurologist expert evaluated PD patients according to the third section of the Movement Disorders Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS-III) to determine disease severity [24]. The recordings included in the study were obtained by asking participants to talk about their daily routine for approximately 90 seconds. Transliterations of the recordings were generated using the Amazon transcribe service. Table 1 contains additional information about the speakers.

Table 1: *Clinical and demographic information of the subjects. [F/M]: Female/Male.*

	PD patients	**HC subjects**
Gender* [F/M]	38/42	43/42
Age** [F/M]	63.7±7.3/64.5±10.2	60.9±8.2/64.8±10.5
Range of age [F/M]	51–81/45–86	49–83/42–86
MDS-UPDRS-III [F/M]	34.6±19.9/38.5±19.6	
Range of MDS-UPDRS-III [F/M]	9–106/7–92	

*Gender matching between PD and HC subjects with a p–value=0.81 calculated through a Chi–square test. **Age matching between PD and HC subjects with a p–value=0.38 calculated through a t-test. Values as mean ± standard deviation.

2.2. Methodology

Figure 1 summarizes the methodology proposed in this work, where speech and language are processed independently to classify PD patients and HCs. In the speech approach, we considered representations obtained using Wav2Vec 2.0 model and representations based on speech dimensions such as prosody, articulation, and phonemic information. In the language approach, we use a state-of-the-art word-embedding model called BETO to get a numerical representation from each word; these embeddings are analyzed using statistical functionals and a strategy where a 1D-CNN analyzes only verbs and nouns from each transcription. Then, the best representations from each modality are used to set a multimodal approach, where both information sources are combined. The combination is developed using two strategies the traditional early fusion and the fusion based on GMUs.

2.3. Speech analysis

To get speech representations, we include two approaches, general representations obtained from a pre-trained model called Wav2Vec 2.0, and pathology-oriented representations to model three speech dimensions: prosody, articulation, and phonemic information.

2.3.1. Wav2vec

This architecture is based on transformers and it was proposed in [25]. The main idea is to encode speech audio via a multi-layer CNN and then mask spans of the resulting latent speech representations, which feed a transformer network to build contextualized representations. In this work, we used a pre-trained Wav2Vec 2.0 model available in Pytorch to get a speech representation for each recording from scratch. This architecture was pre-trained on 960 hours of unlabeled audio from the LibriSpeech dataset. The temporal mean was computed to get a final representation of 768 dimensions.

2.3.2. Speech dimensions

We computed prosody, articulation, and phonemic information features to model the speech signals. This work defines prosody as the variation of loudness, pitch, and timing to produce natural speech. Articulation is defined as the spectral information obtained from voiceless-to-voiced (onset) and voiced-to-voiceless (offset) speech transitions [26]. Finally, phonemic information features aim to evaluate phoneme precision. These features are based on the posterior probability of a speech frame belonging to one (or more) of 18 phonological classes. These features were extracted using the Disvoice toolkit [1].

2.4. Language analysis

The BETO pre-trained model is used to obtain a numerical representation of each word in the transcription. This word-embedding model, which is a Spanish version of BERT proposed in [27], was trained using Spanish data from Wikipedia

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

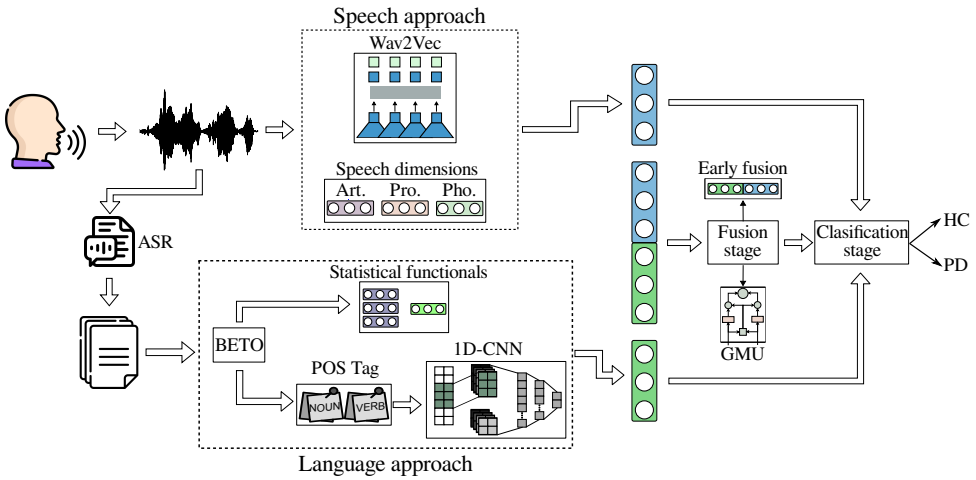

Figure 1: *General methodology proposed in this study.*

and all of the resources of the OPUS project [28]. The source code required to compute BETO embeddings is available online[2] [29]. In order to obtain a static representation of each transcription, two approaches are utilized: statistical functionals and verbs and nouns analysis.

2.4.1. Statistical functionals:

We obtained a static representation of each document by computing the mean, standard deviation, skewness, and kurtosis of the word embeddings, except for the stop words; thus, each document is represented as a 1-dimensional feature vector with 3072 elements.

2.4.2. Verbs and nouns analysis:

Given that different studies show that PD patients exhibit severely impaired capability processing verbs, specifically action verbs, compared to other grammatical units such as nouns [30, 19], we proposed an approach based on 1D-CNN to analyze the word-embeddings of both grammatical units. Initially, the word embeddings of each grammatical unit are vertically concatenated in the transcription order to form a real-valued embedding matrix $\mathbf{X}^{n\times d}$, where n is the number of verbs or nouns in the transcription, and d is the word-embedding dimension. Then, this embedding matrix is used to feed a 1D-CNN, which includes three parallel filters to capture different relationships among the words-embeddings of each grammatical unit. Specifically, $2 \times d$, $3 \times d$, and $4 \times d$ filters are used in the convolutional layers to analyze bi-gram, tri-gram, and four-gram relations, respectively. Finally, a fully-conected layer is used to classify PD and HCs. Details about this architecture can be found in [31].

2.5. Multimodal analysis

Two fusion strategies are employed to combine the best representations from each modality: concatenation and GMU. The first strategy uses early fusion, concatenating both representations, and then an SVM is used as a classifier. The second strategy involves a DL model known as GMU. This model was proposed in [32]. The main idea is to combine early and late fusion aspects using multiplicative gates to find an intermediate representation based on a combination of data from different modalities. First, the representation of each modality feeds a fully-connected layer with an activation function $\tanh$; then, a gate layer with σ activation function controls the contribution of each modality. Finally, the representation obtained is used as input to a fully-connected layer, which develops the classification using a Softmax activation function.

[2] https://github.com/PauPerezT/WEBERT

3. Experiments and Results

Representations of each modality are used independently to classify PD patients and HCs. Best representations from each modality are used to set the multimodal approach. In experiments where an SVM was implemented, we considered Linear and Gaussian kernels, and its hyper-parameters were optimized using a grid-search with values for $C \in \{0.001, 0.01, \cdots, 100\}$ and $\gamma \in \{0.0001, 0.001, \cdots, 100\}$. In DL models, the convolutional filter and the number of units in the fully-connected layer were optimized using a grid search with values of 2^k with $3 \leq k \geq 6$. In addition, we use regularization strategies such as early stopping, dropout, and l_2 regularization. Architecture parameters are optimized upon the validation loss. Finally, all experiments were developed using the same stratified k-fold cross-validation strategy with 10 folds.

3.1. Speech

In the speech approach, we tested five representations: Wav2vec, which is used as a general representation that is not pathology-oriented and features pathology-oriented such as prosody, articulation, phonemic information, and the combination of all of them. For the Wav2vec strategy, we performed different segmentations of the raw signal at its input (1sec, 2sec, 5sec, and the full recording); however, we only reported the best result, which was obtained with an input size of 2 seconds. Table 2 summarizes the performance measures of the approaches tested. Although the Wav2vec strategy achieved an accuracy of 79.4%, the combination of all speech dimensions shows an accuracy slightly better and a balanced sensitivity and specificity. The individual speech dimension with the highest accuracy is the feature set related to phonemic information, which aims to evaluate the phoneme articulation precision of the patient.

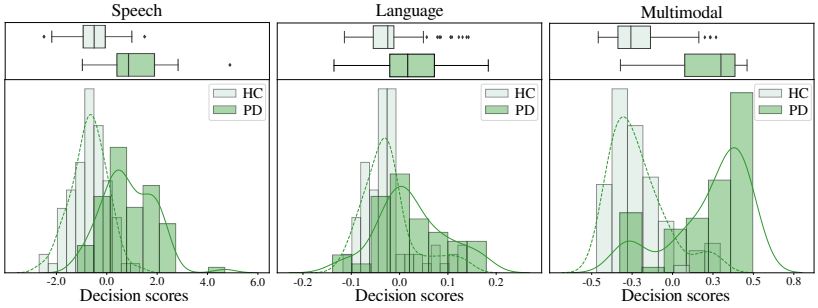

Figure 2: *Distributions for the best representation in unimodal and multimodal approaches.*

Table 2: *Classification between PD patients vs. HC subjects for the uni-modal and multimodal approaches.*

	Accuracy	Sensitivity	Specificity	F1-score
Speech				
Wav2vec	79.4± 9.6	71.3±14.8	87.1±12.1	76.6±11.1
Prosody	77.2±12.7	70.0±19.5	83.9±11.4	73.8±15.6
Articulation	74.5± 7.7	68.8±11.5	80.0± 9.2	72.1± 8.8
Phonemic	78.2±11.7	75.0±14.8	81.3±20.2	77.0±11.5
All dim.	**81.9±10.3**	**80.0±12.7**	**83.9±14.3**	**81.0±10.9**
Language				
Statistical functionals	63.1± 9.8	42.5±16.0	82.5±11.6	51.8±14.6
Nouns-CNN	73.3± 9.3	68.8±23.2	77.8±18.7	69.5±15.0
Verbs-CNN	**76.4± 5.8**	**71.3±17.7**	**81.1±11.0**	**73.3± 9.3**
Multimodal				
All dim. ‖ verbs-CNN	80.7±10.3	80.0±13.9	81.1±14.9	79.9±11.1
GMU (All dim. , verbs-CNN)	**87.3± 6.8**	**83.8±14.8**	**90.7±13.0**	**86.1± 7.5**

Values reported in terms of mean ± standard deviation. ‖ denotes the concatenation operation. **All dim**: the representation obtained concatenating prosody, articulation, and phonemic information features.

3.2. Language

In the language modality, we tested three representations: statistical functionals, Nouns+CNN, and Verbs+CNN. In this case, statistical functionals representation is the static representation of each document by computing the statistical functionals of the BETO embeddings in the transcription. Representation based on statistical functionals is used as the general language representation that is not pathology-oriented. Noun-CNN and Verbs-CNN are the two representations obtained when the 1D-CNN analyzes only the nouns or verbs in the transcription, respectively. Table 2 shows the performance for these representations where the best result is obtained with Verbs+CNN representation. This result is coherent with the literature, where different not automatic approaches have reported that PD patients show difficulties processing verbs compared to other grammatical units such as nouns [30].

3.3. Multimodal

The best speech and language representations are combined using two fusion strategies: traditional early fusion and fusion based on GMU. In traditional early fusion, representations of each modality are concatenated to feed an SVM classifier. In GMU-based fusion, a DL model learns an intermediate representation based on both modalities' information. Table 2 shows that GMU-based fusion outperforms traditional early fusion by up to 7% in PD patients and HCs classification. Furthermore, multimodal results surpass unimodal results, indicating complementary information between language and speech representations. Figure 2 shows the distribution of the scores of the best models from each modality and for the multimodal approach using GMUs.

4. Conclusions

This study aims to classify PD patients and HCs considering unimodal and multimodal approaches from speech and language analysis. We consider non-pathology-oriented and pathology-oriented representations from speech and language. In the speech approach, we used features based on Wav2vec and representations based on different speech dimensions, such as prosody, articulation, and phonemic information. In the language approach, we used BETO and proposed an approach where verbs and nouns were analyzed using a 1D-CNN. Best unimodal representations were used to set a multimodal approach, where two fusion strategies are evaluated: traditional early fusion and a strategy based on GMU.

Unimodal results showed that pathology-oriented representations outperformed general representations in both speech and language modalities. In speech modality, the best result was obtained when all speech dimensions were concatenated, and in language modality, the best result was achieved with the representation obtained when the 1D-CNN analyzed only the verbs in the transcription; accuracies of up to 81% and 76% were obtained for speech and language modality, respectively. In the multimodal, the strategy based on GMU achieved accuracies of up to 87%, outperforming the traditional fusion strategy by up to 7%, and in up to 6%, the best result obtained in the unimodal approaches.

Results show that pathology-oriented representations outperform the general and uninterpretable representations obtained with complex DL models. On the other hand, relevant information about the impaired processing of verbs in PD patients was extracted using the 1D-CNN strategy. Furthermore, results in the language approach are consistent with other studies where PD patients show more difficulties processing verbs than other grammatical units such as nouns. Finally, multimodal results suggest that combining information from speech and language can help improve the automatic analysis of PD. The main limitation of this work is that we did not consider synchronous fusion strategies, which can be addressed by forced alignment techniques. Therefore, future work will include synchronous fusion between speech and language for automatic disease monitoring.

5. Acknowledgements

This work was funded by UdeA grants # ES92210001 and PRG2020-34068. Adolfo García is supported by GBHI, Alzheimer's Association, and Alzheimer's Society (Alzheimer's Association GBHI ALZ UK-22-865742); ANID (FONDECYT Regular 1210176); and PIIECC, Facultad de Humanidades, USACH.

6. References

[1] O. Hornykiewicz, "Biochemical aspects of Parkinson's disease," *Neurology*, vol. 51, no. 2 Suppl 2, pp. S2–S9, 1998.

[2] J. Mu *et al.*, "Parkinson's disease subtypes identified from cluster analysis of motor and non-motor symptoms," *Frontiers in aging neuroscience*, vol. 9, p. 301, 2017.

[3] S. Pinto *et al.*, "Treatments for dysarthria in parkinson's disease," *The Lancet Neurology*, vol. 3, no. 9, pp. 547–556, 2004.

[4] R. J. *et al.*, "Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task," *Journal of the Acoustical Society of America*, vol. 134, no. 3, pp. 2171–2181, 2013.

[5] L. Liu *et al.*, "Characteristics of language impairment in Parkinson's disease and its influencing factors," *Translational Neurodegeneration*, vol. 4, no. 1, pp. 1–8, 2015.

[6] E. Eyigoz *et al.*, "From discourse to pathology: automatic identification of Parkinson's disease patients via morphological measures across three languages," *Cortex*, vol. 132, pp. 191–205, 2020.

[7] A. García *et al.*, "Cognitive determinants of dysarthria in Parkinson's disease: An automated machine learning approach," *Movement disorders*, vol. 36, no. 12, pp. 2862–2873, 2021.

[8] M. Grossman *et al.*, "Verb learning in parkinson's disease." *Neuropsychology*, vol. 8, no. 3, p. 413, 1994.

[9] B. Sonawane and P. Sharma, "Speech-based solution to Parkinson's disease management," *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29 437–29 451, 2021.

[10] C. Rios-Urrego *et al.*, "End-to-end Parkinson's disease detection using a deep convolutional recurrent network," in *International Conference on Text, Speech, and Dialogue*. Springer, 2022, pp. 326–338.

[11] J. Vásquez-Correa *et al.*, "Transfer learning helps to improve the accuracy to classify patients with different speech disorders in different languages," *Pattern Recognition Letters*, vol. 150, pp. 272–279, 2021.

[12] C. Quan *et al.*, "End-to-end deep learning approach for Parkinson's disease detection from speech signals," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 2, pp. 556–574, 2022.

[13] J. C. Vásquez-Correa *et al.*, "Towards an automatic evaluation of the dysarthria level of patients with parkinson's disease," *Journal of communication disorders*, vol. 76, pp. 21–36, 2018.

[14] Y. Liu *et al.*, "Automatic assessment of parkinson's disease using speech representations of phonation and articulation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 242–255, 2022.

[15] P. Klumpp *et al.*, "The phonetic footprint of parkinson's disease," *Computer Speech & Language*, vol. 72, p. 101321, 2022.

[16] A. M. García *et al.*, "Cognitive determinants of dysarthria in parkinson's disease: an automated machine learning approach," *Movement Disorders*, vol. 36, no. 12, pp. 2862–2873, 2021.

[17] L. Jessiman, G. Murray, and M. Braley, "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*, 2018, pp. 63–74.

[18] P. A. Pérez-Toro *et al.*, "Natural language analysis to detect Parkinson's disease," in *International Conference on Text, Speech, and Dialogue*. Springer, 2019, pp. 82–90.

[19] S. Abrevaya *et al.*, "The road less traveled: alternative pathways for action-verb processing in parkinson's disease," *Journal of Alzheimer's Disease*, vol. 55, no. 4, pp. 1429–1435, 2017.

[20] A. García *et al.*, "Detecting Parkinson's disease and its cognitive phenotypes via automated semantic analyses of action stories," *NPJ Parkinson's Disease*, vol. 8, no. 1, pp. 163–10, 2022.

[21] P. A. Pérez-Toro *et al.*, "Depression assessment in people with parkinson's disease: The combination of acoustic features and natural language processing," *Speech Communication*, vol. 145, pp. 10–20, 2022.

[22] J. R. Orozco-Arroyave *et al.*, "New spanish speech corpus database for the analysis of people suffering from parkinson's disease." in *LREC*, 2014, pp. 342–347.

[23] J. M. Huerta and R. M. Stern, "Speech recognition from gsm codec parameters." in *ICSLP*, 1998, pp. 1463–1466.

[24] C. G. Goetz *et al.*, "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement disorders*, vol. 23, no. 15, pp. 2129–2170, 2008.

[25] A. Baevski *et al.*, "Wav2vec 2.0: A framework for self-supervised learning of speech representations," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 449–12 460, 2020.

[26] J. Orozco-Arroyave, J. Vásquez-Correa, and N. 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*. IOP Publishing, 2020, ch. 8. In: Neurological Disorders and Imaging Physics, Volume 5.

[27] J. Canete *et al.*, "Spanish pre-trained bert model and evaluation data," *Pml4dc at iclr*, vol. 2020, pp. 1–10, 2020.

[28] J. Tiedemann, "Parallel data, tools and interfaces in opus." in *Lrec*, vol. 2012, 2012, pp. 2214–2218.

[29] P. A. Perez-Toro, "PauPerezT/WEBERT: Word Embeddings using BERT," Jul. 2020.

[30] A. M. García, J. DeLeon, and B. L. Tee, "Neurodegenerative disorders of speech and language: non-language-dominant diseases," 2022.

[31] D. Escobar-Grisales *et al.*, "Colombian dialect recognition based on information extracted from speech and text signals," in *2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. IEEE, 2021, pp. 556–563.

[32] J. Arevalo *et al.*, "Gated multimodal networks," *Neural Computing and Applications*, vol. 32, pp. 10 209–10 228, 2020.