

Automated Cross-language Intelligibility Analysis of Parkinson's Disease Patients Using Speech Recognition Technologies

Nina Hosseini-Kivanani¹, Juan Camilo Vásquez-Correa^{2,3}, Manfred Stede⁴, Elmar Nöth²

¹International Experimental and Clinical Linguistics (IECL), University of Potsdam, Germany

²Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Germany

³Faculty of Engineering, University of Antioquia UdeA, Medellín, Colombia

⁴Applied Computational Linguistics, University of Potsdam, Germany

{hosseinikiva, stede}@uni-potsdam.de

jcamillo.vasquez@udea.edu.co

elmar.noeth@fau.de

Abstract

Speech deficits are common symptoms among Parkinson's Disease (PD) patients. The automatic assessment of speech signals is promising for the evaluation of the neurological state and the speech quality of the patients. Recently, progress has been made in applying machine learning and computational methods to automatically evaluate the speech of PD patients. In the present study, we plan to analyze the speech signals of PD patients and healthy control (HC) subjects in three different languages: German, Spanish, and Czech, with the aim to identify biomarkers to discriminate between PD patients and HC subjects and to evaluate the neurological state of the patients. Therefore, the main contribution of this study is the automatic classification of PD patients and HC subjects in different languages with focusing on phonation, articulation, and prosody. We will focus on an intelligibility analysis based on automatic speech recognition systems trained on these three languages. This is one of the first studies done that considers the evaluation of the speech of PD patients in different languages. The purpose of this research proposal is to build a model that can discriminate PD and HC subjects even when the language used for train and test is different.

1 Introduction

Parkinsons disease (PD) (i.e. Shaking palsy (Parkinson, 2002)) is the second most common neurodegenerative disorder after Alzheimers disease. PD displays a great prevalence in individuals of advanced age (Dexter and Jenner, 2013), especially, over the age of fifty (Fahn, 2003). The signs and symptoms of PD can significantly influence the quality of life of patients. They are grouped

into two categories: motor and non-motor symptoms. Speech impairments are one of the earliest manifestations in PD patients.

Early diagnosis of PD is a vital challenge in this field. The first step in analyzing this disease is the development of markers of PD progression through collecting data from several cohorts. To reach this aim, different clinical rating scales have been developed, such as the Unified Parkinson's Disease Rating Scale (UPDRS), Movement Disorders Society - UPDRS (MDS-UPDRS)¹ (Goetz et al., 2008) and Hoehn & Yahr (H & Y) staging, (Visser et al., 2006).

The UPDRS is the most widely used rating tool for the clinical evaluation of PD patients. The examination requires observation and interview by a professional clinician. The scale is distributed into 4 sections: (i) Mentation, behavior and mood, (ii) Activities of daily living (ADL), (iii) Motor sections, and (iv) Motor complications.

One of the most common motor problems is related to speech impairments in PD (Jankovic, 2008). Most of the patients with PD show disabilities in speech production. The most common speech disturbances are monotonic speech, hypophonia (a speech weakness in the vocal musculature and vocal sounds) and hypokinetic dysarthria. These symptoms reduce the intelligibility of the patients, and affect different aspects of the speech production such as articulation, phonation, nasality, and prosody (Little et al., 2009; Goetz et al., 2008; Ramig et al., 2001). Therefore, there is a great interest to develop tools or methods to evaluate and improve the speech production of PD patients.

¹MDS-UPDRS is the most updated version of the UPDRS.

Recently, there has been a proliferation of new speech recognition-based tools for the acoustic analysis of PD. The use of speech recognition software in clinical examinations could make a powerful supplement to the state-of-the-art subjective reports of experts and clinicians that are costly and time-consuming (e.g., Little et al., 2009; Hernandez-Espinosa et al., 2002). In the clinical field, the detection of PD is a complex task due to the fact that the symptoms of this disease are more related to clinicians' observations and perception of the way patients move and speak.

Recently, machine learning tools are used to develop speech recognition systems that make the whole process of objective evaluation and recognition faster and more accurate than analytical clinicians' methods (Yu and Deng, 2016; Hernandez-Espinosa et al., 2002). Using machine learning techniques to extract acoustic features for detecting the PD has become widely used in recent studies (e.g., Dahl et al., 2012; Little et al., 2009).

Automatic speech recognition (ASR) systems are used to decode and transcribe oral speech. In other words, the goal of ASR systems is to find and recognize the words that best represent the acoustic signal. For example, automatic speech recognition systems are used to evaluate how speech intelligibility is affected by the disease.

This study will seek to further investigate the speech patterns of HC and PD groups using recordings from patients speaking in German, Spanish, and Czech. Most of the previous studies only considered recordings in one language and focused on it for detecting PD, but in this study, we plan to evaluate the effect of the PD in three different languages.

2 Related work: ASR for detecting PD symptoms

Speech can be measured by acoustic tools simply using aperiodic vibrations in the voice. The field of speech recognition has been improved in recent years by research in computer-assisted speech training system for therapy (Beijer and Ritvel, 2012) machine learning techniques, which can lead to establish efficient biomarkers to discriminate HC from people with PD (e.g., Orozco-Arroyave et al., 2013).

There are a vast number of advanced techniques to design ASR systems: hybrid Deep Neu-

ral Networks-Hidden Markov Models (DNN²-HMM) (Hinton et al., 2012) and Convolutional Neural Networks (CNN) (Abdel-Hamid et al., 2014). Deep neural networks have recently received increasing attention in speech recognition (Canevari et al., 2013). Other studies have highlighted the strength of the DNN-HMM framework for speech recognition (e.g., Dahl et al., 2012).

On the other hand, regarding the assessment of PD from speech, Skodda et al. (2011) assessed the progression of speech impairments of PD from 2002 to 2012 in a longitudinal study by only using a statistical test to evaluate changes in aspects related to voice, articulation, prosody, and fluency of the recorded speech.

Orozco-Arroyave et al. (2016) considered more than one language for producing isolated words for discriminating PDs from HCs. The characterization and classification processes were based on a method on the systematic separation of voiced and unvoiced segments of speech in their study. Vásquez-Correa et al. (2017) analyzed the effect of acoustic conditions on different algorithms. The obtained detection accuracy of PD speech was reported and shown that the background noise affect the performance of the different algorithms. However, most of these systems are not yet capable of automatically detecting speech impairment of individual speech sounds, which are known to have an impact on speech intelligibility (Zhao et al., 2010; Ramaker et al., 2002).

Our goal is to develop robust ASR systems for pathological speech and incorporate the ASR technology to detect their speech intelligibility problems. A major interest is to investigate acoustic features in the mentioned languages (differences and similarities), including gender differences between subject (HC & PD) groups. The overall purpose of this project is to address the question of whether cross-lingual speech intelligibility among PDs and HCs can help the recognition system to detect the correct disease. The classification of PD from speech in different languages has to be carefully conducted to avoid bias towards the linguistic content present in each language. For instance, Czech and German languages are richer than Spanish language in terms of consonant production, which may cause that it is easier to produce consonant sounds by Czech PD patients than

²DNN is a feed-forward neural network that has more than one layer of hidden units between its inputs and its outputs (Hinton et al., 2012).

by Spanish PD patients. In addition, with the use of transfer learning strategies, a model trained with utterances from one language can be used as a base model to train a model in a different language.

After reviewing the aforementioned literature, the main contribution of our research for modeling speech signals in PD patients is twofold:

- This is one of the first cross-lingual studies done to evaluate speech of people with PD. This work requires a database consisting of recordings of different languages. There is currently a lack of cross-lingual research, which provides reliable classification methods for assessing PDs' speech available in the literature.
- Using speech data is expected to help the development of a diagnostic of PD patients.

This project seeks to bridge the gap in speech recognition for speech of PD, with the hope of moving towards a higher adoption rate of ASR-based technologies in the diagnosis of patients.

3 The set-up of the ASR system

In this work, we will build an ASR system to recognize the speech of patients of Parkinson's Disease. The task of ASR is to convert this raw audio into text. The ASR system is created based on three models: acoustic model (i.e. to recognize phonemes), pronunciation model (i.e. to map sequence of phonemes into word sequences), and language model (i.e. to estimate probabilities of word sequences). We place particular emphasis on the acoustic model portion of the system. We also provide some acoustic models output features that could be used in future speech recognition of PD severity in the clinical field. Ravanelli et al. (2019) stated that along with the improvement of ASR systems, several deep learning frameworks (e.g., TensorFlow (Abadi et al., 2016)) in machine learning are also widely used.

The next section describes the process for modeling the intelligibility of PD speech followed by the description of processes whether the speech signal is from PD patient or from HC subjects.

3.1 Training

The proposed ASR system will be developed using a standard state-of-the-art architecture hybrid DNN-HMM (see Nassif et al., 2019 for more

information about the existing models in ASR), built using the Kaldi speech recognition toolkit³. The preprocessing (i.e. Feature extraction) of the acoustic signal into usable parameters (i.e. label computation) tries to remove any acoustic information that is not useful for the task; it will be done before beginning to train the acoustic model. In this study, we will use Mel-frequency Cepstral coefficients (MFCC) and Mel filter bank energies (e.g., compute-mfcc-feats and compute-fbank-feats) to train the acoustic models of the ASR systems. The task of calculating MFCCs from acoustic features is to characterize an underlying signal using spectrograms and represent the shape of the vocal tract including tongue, teeth etc.

It was observed that filter bank (fbank), one of the most popular speech recognition features, performs better than MFCCs when used with deep neural networks (Hinton et al., 2012). The purpose of using acoustic model is to help us to get boundaries of the phoneme labels. The acoustic models will be trained based on different acoustic features extracted in Kaldi "nnet3" recipes. The extracted acoustic features and the observation probabilities of our ASR system will be used to train the hybrid DNN-HMM acoustic model. The performance of an ASR system will be measured by Word Error Rate (WER) of the transcript produced by the system against the target transcript.

PyTorch: PyTorch is one of the most well known deep learning toolkit that facilitates the design of neural architectures. This tool will be used to design new DNN architectures to improve the performance of the ASR system. We will additionally use PyTorch-Kaldi (Ravanelli et al., 2019), to train⁴ deep neural network based models (e.g., DNNs, CNNs, and Recurrent Neural Networks (RNNs) models) and traditional machine learning classifier. Ravanelli et al. (2019) stated that this PyTorch-Kaldi toolkit acts like an interface with different speech recognition features in it and can be used as a state-of-the-art in the field of ASR (See Figure 1). Figure 1 is shown the general methodology that will be applied in this research.

³Kaldi: <http://kaldi-asr.org/doc/>

⁴PyTorch-Kaldi: <https://github.com/mravanelli/pytorch-kaldi>

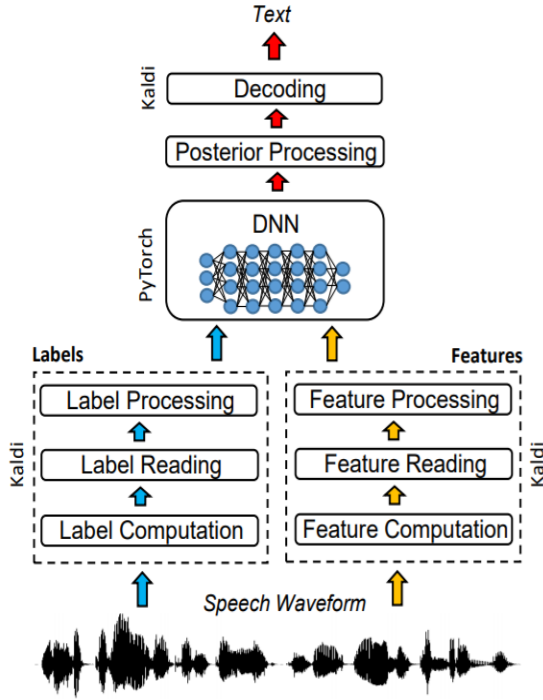


Figure 1: ASR system architecture that will be used in this study (Ravanelli et al., 2019).

4 Methods

4.1 Data

The data of this study comes from an extended version of PC-GITA database for Spanish (Orozco-Arroyave et al., 2014), German (Skodda et al., 2011), and Czech (Rusz et al., 2011) with more recordings from PDs and HCs. The database consists of both PD and HC subjects.

All subjects were asked to do multiple types of speaking tasks to understand how speech changes in different conditions, due to the fact that voice variation is difficult to identify without human experience (Jeancolas et al., 2017). The speech dimensions considered in this project are phonation, articulation and prosody (See Figure 2).

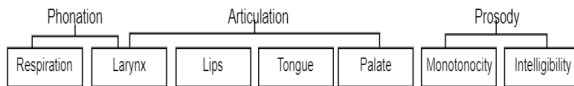


Figure 2: Speech dimensions: Phonation, Articulation and Prosody.

For each subject, speech material includes (i) sustained vowel phonation; participants were asked to phonate vowels for several seconds; (ii) rapid syllable repetition (ideally Diadochokinetic (DDK)); participants were asked to produce such as /pa-ta-ka/, /pa-ka-ta/, /pe-ta-ka/, /pa/, /ta/,

and /ka/, (iii) connected speech, consisting of: (iv) reading a specific text, and (v) spontaneous speech.

This dataset consists of speech samples recorded from 88 PD and 88 HC German speaking participants, 50 PD and 50 HC Spanish speaking participants (balanced in age and gender), and 20 PD and 16 HC Czech speaking participants. These speech samples were assessed by expert neurologists using UPDRS-III and H & Y scales. Their neurological labels were reported based on the UPDRS-III and H & Y scales (*mean ± SD*) for each PD group:

- PD-German: UPDRS-III (22.7 ± 10.9), H&Y (2.4 ± 0.6)
- PD-Spanish: UPDRS-III (36.7 ± 18.7), H&Y (2.3 ± 0.8)
- PD-Czech: UPDRS-III (17.9 ± 7.3), H&Y (2.2 ± 0.5)

Further details are shown in Table 1:

Language	HC		PD	
	Female	Male	Female	Male
German	n= 44 (63.8 ± 12.7)	n= 44 (62.6 ± 15.2)	n= 41 (66.2 ± 9.7)	n= 47 (66.7 ± 8.4)
	n= 25 (61.4 ± 7.0)	n= 25 (60.3 ± 11.6)	n= 25 (60.7 ± 7.3)	n= 25 (61.6 ± 11.2)
Czech	—	n= 16 (61.8 ± 13.3)	—	n= 20 (61 ± 12)

Table 1: Age information of HC and PD subjects in the study (n = number of participant) & the mean and standard deviation are in the parenthesis (*Mean ± SD*).

Although the size of the data is not big enough, the vocabulary that was used by the patients in the capture process of the speech was fixed. This aspect compensates the need to have a huge corpus to evaluate a vocabulary dependent task like the assessment of pathological speech (see Parra-Gallego et al., 2018).

4.2 Sample

Praat software (Boersma and Weenink, 2016) is used for segmenting speech, extracting acoustics features, removing silence from beginning and end of speech file and visualization of speech data. Generally, spoken words, represented as sound waves, have two axes: time on the x-axis and amplitude on the y-axis. Figure 3 illustrates the example of input feature maps extracted from the speech signal which shows the selected spectrograms (the audio waveform is encoded as a representation) of PD and HC subjects when they pronounce the syllable /pa-ta-ka/ that convey 3-dimensional information in 2 dimensions

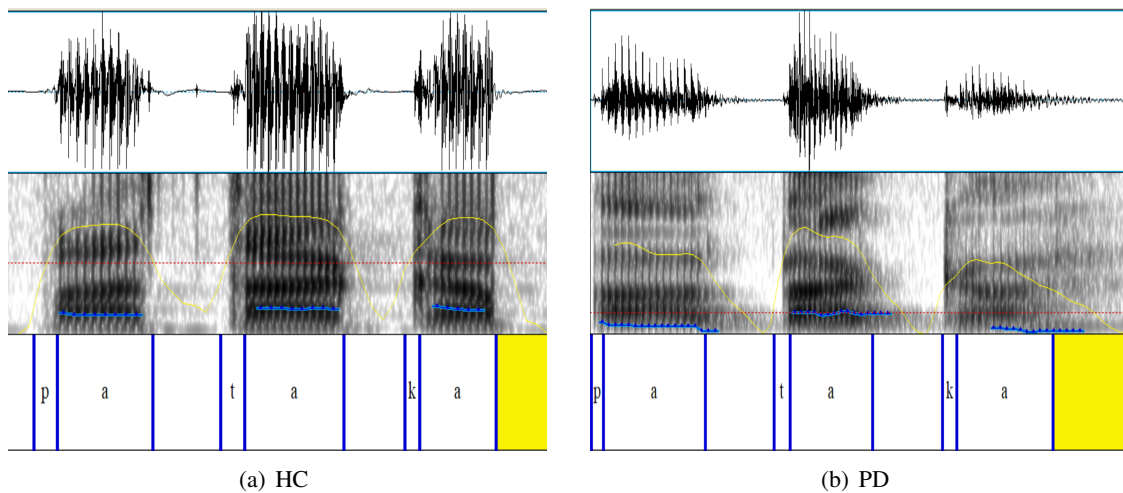


Figure 3: **Top:** Raw waveforms of /pa-ta-ka/ (x-axis: time; y-axis: amplitude). **Middle:** Spectrograms (x-axis: time; y-axis: frequency; shading: amplitude (energy), darker means higher). **Bottom:** the word-level annotation of the signal.

(i.e. Time: x-axis, Frequency: y-axis, and Amplitude: color intensity). The proposed model will be able to identify specific aspects in the speech related to the pronunciation of consonants, which are the most affected aspects of the speech of the patients due to the disease. The segmentation process will be performed using a trained model to detect phonological classes, like those ones used in the previous studies (Vásquez-Correa et al., 2019; Cernak et al., 2017). Figure 3 shows the possible differences in articulation and phonation in PD and HC subjects. By using Praat, the speech samples of syllable /pa-ta-ka/ will be segmented into vowel and consonant frames. The contour of HC sample is more stable than the obtained contour from PD sample. In each sample, silences will be removed from the beginning and the end of each token, using Praat.

5 Conclusion

In this research proposal, we introduced and described the background for speech recognition of PD patients. The focus is on Parkinsons disease speech recognition based on the acoustic analysis of their voice. A brief overview of clinical and machine learning research in this field was provided. The goal is to improve the ASR system to be able to model and detect PD patients independently from their language by taking speech as an input and using machine learning and natural language processing technologies to advance healthcare and provide an overview of the patients men-

tal health. All in all, the proposed method should be able to detect the patient with PD and discriminate them from HC subjects.

References

- Martn Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, Xiaoqiang Zheng, and Google Brain. 2016. *TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning*. In *In 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, pages 265–283.
- Ossama Abdel-Hamid, Abdel Rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu. 2014. *Convolutional neural networks for speech recognition*. *IEEE Transactions on Audio, Speech and Language Processing*, 22(10):1533–1545.
- Lilian Beijer and Toni Rietvel. 2012. *Potentials of Telehealth Devices for Speech Therapy in Parkinson’s Disease*. In *Diagnostics and Rehabilitation of Parkinson’s Disease*. *InTech*, pages 379–402.
- Paul Boersma and David Weenink. 2016. Praat: Doing phonetics by computer (Version 6.0. 14). *Retrieved from (last access: 29.04. 2018)*.
- Claudia Canevari, Leonardo Badino, Luciano Fadiga, and Giorgio Metta Rbcs. 2013. *Cross-corpus and cross-linguistic evaluation of a speaker-dependent DNN-HMM ASR system using EMA data*. In *In Speech Production in Automatic Speech Recognition (In SPASR-2013)*, pages 29–33.

- Milos Cernak, Juan Rafael Orozco-Arroyave, Frank Rudzicz, Heidi Christensen, Juan Camilo Vásquez-Correa, and Elmar Nöth. 2017. [Characterisation of voice quality of Parkinson’s disease using differential phonological posterior features](#). *Computer Speech and Language*, 46(June):196–208.
- George E. Dahl, Dong Yu, Li Deng, and Alex Acero. 2012. [Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition](#). *IEEE Transactions on Audio, Speech and Language Processing*, 20(1):30–42.
- David T. Dexter and Peter Jenner. 2013. [Parkinson disease: from pathology to molecular disease mechanisms](#). *Free Radical Biology and Medicine*, 62:132–144.
- Stanley Fahn. 2003. [Description of Parkinson’s Disease as a Clinical Syndrome](#). *Annals of the New York Academy of Sciences*, 991(1):1–14.
- Christopher G. Goetz, Barbara C. Tilley, Stephanie R. Shaftman, Glenn T. Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B. Stern, Richard Dodel, Bruno Dubois, Robert Holloway, Joseph Jankovic, Jaime Kulisevsky, Anthony E. Lang, Andrew Lees, Sue Leurgans, Peter A. LeWitt, David Nyenhuis, C. Warren Olanow, Olivier Rascol, Anette Schrag, Jeanne A. Teresi, Jacobus J. van Hilten, and Nancy LaPelle. 2008. [Movement Disorder Society-sponsored revision of the Unified Parkinson’s Disease Rating Scale \(MDS-UPDRS\): Scale presentation and clinimetric testing results](#). *Movement Disorders*, 23(15):2129–2170.
- Carlos Hernandez-Espinosa, Pedro Gomez-Vilda, Juan I. Godino-Llorente, and Santiago Aguilera-Navarro. 2002. [Diagnosis of vocal and voice disorders by the speech signal](#). In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, pages 253–258. IEEE.
- Geoffrey Hinton, Li Deng, Dong Yu, George Dahl, Abdel Rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara Sainath, and Brian Kingsbury. 2012. [Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups](#). *IEEE Signal Processing Magazine*, 29(6):82–97.
- Joseph Jankovic. 2008. [Parkinsons disease: clinical features and diagnosis](#). *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4):368–376.
- Laetitia Jeancolas, Habib Benali, Badr-Eddine Benkelfat, Graziella Mangone, Jean-Christophe Corvol, Marie Vidailhet, Stephane Lehericy, and Dijana Petrovska-Delacretaz. 2017. [Automatic detection of early stages of Parkinson’s disease through acoustic voice analysis with mel-frequency cepstral coefficients](#). In *International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–6. IEEE.
- Max A Little, Patrick E McSharry, Eric J Hunter, Jennifer Spielman, and Lorraine O Ramig. 2009. [Suitability of dysphonia measurements for telemonitoring of Parkinson’s disease](#). *IEEE transactions on bio-medical engineering*, 56(4):1015.
- Ali Bou Nassif, Ismail Shahin, Imtinan Attili, Mohammad Azzeh, and Khaled Shaalan. 2019. [Speech Recognition Using Deep Neural Networks: A Systematic Review](#). *IEEE Access*, 7:19143–19165.
- Juan R. Orozco-Arroyave, Julin D. Arias-Londoño, Jesus Francisco V. Bonilla, Mara C. Gonzalez-Rátiva, and Elmar Nöth. 2014. [New Spanish speech corpus database for the analysis of people suffering from Parkinson’s disease](#). In *Language Resources and Evaluation Conference, LREC*, pages 342–347.
- Juan R. Orozco-Arroyave, Julin D. Arias-Londoño, Jess F. Vargas-Bonilla, and Elmar Nöth. 2013. [Analysis of Speech from People with Parkinsons Disease through Nonlinear Dynamics](#). In *International conference on nonlinear speech processing*, pages 112–119. Springer, Berlin, Heidelberg.
- Juan R. Orozco-Arroyave, Juan C. Vásquez-Correa, Florian Honig, Julin D. Arias-Londono, Jess F. Vargas-Bonilla, Sabine Skodda, Jan Ruzs, and Elmar Nöth. 2016. [Towards an automatic monitoring of the neurological state of Parkinson’s patients from speech](#). *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2016-May(March):6490–6494.
- James Parkinson. 2002. [An Essay on the Shaking Palsy](#). *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(2):223–236.
- Luis F. Parra-Gallego, Toms Arias-Vergara, Juan C. Vásquez-Correa, Nicanor Garcia-Ospina, Juan R. Orozco-Arroyave, and Elmar Nöth. 2018. [Automatic Intelligibility Assessment of Parkinsons Disease with Diadochokinetic Exercises](#). pages 223–230. Springer, Cham.
- Claudia Ramaker, Johan Marinus, Anne Margarethe Stiggelbout, and Bob Johannes van Hilten. 2002. [Systematic evaluation of rating scales for impairment and disability in Parkinson’s disease](#). *Movement Disorders*, 17(5):867–876.
- Lorraine Olson Ramig, Steven Gray, Kristin Baker, Kim Corbin-Lewis, Eugene Buder, Erich Luschei, Hillary Coon, and Marshall Smith. 2001. [The aging voice: A review, treatment data and familial and genetic perspectives](#). *Folia Phoniatica et Logopaedica*, 53(5):252–265.
- Mirco Ravanelli, Titouan Parcollet, and Yoshua Bengio. 2019. [The PyTorch-Kaldi Speech Recognition Toolkit](#). *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6465–6469.

- Jan Ruzs, Roman Cmejla, Hana Ruzickova, and Evzen Ruzicka. 2011. [Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease](#). *The Journal of the Acoustical Society of America*, 129(1):350–367.
- Sabine Skodda, Wenke Visser, and Uwe Schlegel. 2011. [Vowel Articulation in Parkinson's Disease](#). *Journal of Voice*, 25(4):467–472.
- Juan C. Vásquez-Correa, Philipp Klumpp, Juan R. Orozco-Arroyave, and Elmar Nöth. 2019. [Phonet: a Tool Based on Gated Recurrent Neural Networks to Extract Phonological Posteriors from Speech](#). In *Proceedings of INTERSPEECH conference (Under review)*.
- Juan C. Vásquez-Correa, Joan Serrà, Juan R. Orozco-Arroyave, Jess F. Vargas-Bonilla, and Elmar Nöth. 2017. [Effect of acoustic conditions on algorithms to detect Parkinson's disease from speech](#). *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 5065–5069.
- Martine Visser, Albert F.G. Leentjens, Johan Marinus, Anne M. Stiggelbout, and Jacobus J. van Hilten. 2006. [Reliability and validity of the Beck depression inventory in patients with Parkinson's disease](#). *Movement Disorders*, 21(5):668–672.
- Dong Yu and Li Deng. 2016. *Automatic Speech Recognition-A Deep Learning Approach*. Springer london limited.
- Ying J. Zhao, Hwee L. Wee, Yiong H. Chan, Soo H. Seah, Wing L. Au, Puay Ngoh Lau, Emmanuel C. Pica, Shu Ch. Li, Nan Luo, and Louis C S Tan. 2010. [Progression of Parkinson's disease as evaluated by Hoehn and Yahr stage transition times](#). *Movement Disorders*, 25(6):710–716.