Distortion of voiced obstruents for differential diagnosis between parkinson's disease and multiple system atrophy

Khalid Daoudi[1], Biswajit Das[1], Solange Milhé de Saint Victor [2], Alexandra Foubert-Samier[3] ,
Anne Pavy-Le Traon[4], Olivier Rascol[4], Wassilios G. Meissner[3], Virginie Woisard[5]

[1]INRIA Bordeaux Sud-Ouest, GeoStat team, France
[2] ENT Department, University Hospital of Bordeaux, France
[3] Department of Neurology for Neurodegenerative diseases, University Hospital of Bordeaux
and University of Bordeaux, CNRS, IMN, UMR 5293, France
[4] MSA Reference Center, Clinical Investigation Center CIC1436,
Dep. Neurosciences & Clinical Pharmacology, University Hospital of Toulouse, France
[5]Voice and deglutition unit , ENT department, University Hospital of Toulouse, France.

khalid.daoudi@inria.fr, biswajit.das@inria.fr,
solange.milhe-de-saint-victor@chu-bordeaux.fr, alexandra.samier-foubert@u-bordeaux.fr,
pavy-letraon.a@chu-toulouse.fr, olivier.rascol@univ-tlse3.fr,
wassilios.meissner@chu-bordeaux.fr, woisard.v@chu-toulouse.fr

Abstract

Parkinson's disease (PD) and the parkinsonian variant of Multiple System Atrophy (MSA-P) are two neurodegenerative diseases which share similar clinical features, particularly in early disease stages. The differential diagnosis can be thus very challenging. Dysarthria is known to be a frequent and early clinical feature of PD and MSA. It can be thus used as a vehicle to provide a vocal biomarker which could help in the differential diagnosis. In particular, distortion of consonants is known to be a frequent impairment in these diseases. The aim of this study is to investigate distinctive patterns in the distortion of voiced obstruents (plosives and fricatives). It is the first study which attempts to examine such distortions in the French language for the purpose of the differential diagnosis between PD and MSA-P (and among the very few studies if we consider all languages). We carry out a perceptual and objective analysis of voiced obstruents extracted from isolated pseudo-words initials. We first show that devoicing is a significant impairment which predominates in MSA-P. We then show that voice onset time (VOT) of voiced plosives (prevoicing duration) can be a complementary feature to improve the accuracy in discrimination between PD and MSA-P.

Index Terms: Parkinson's disease, Multiple system atrophy, Dysarthria, Pathological speech processing, Consonant distortion

1. Introduction

Parkinson's disease (PD) is a neurological disorder principally caused by the degeneration of midbrain dopaminergic neurons, leading to slowness of movement, muscle rigidity and resting tremor. Multiple system atrophy (MSA) belongs to the group of atypical parkinsonian disorders (APD) with a poor prognosis. MSA differs from PD by a more widespread neurodegenerative process, resulting in more rapid disease progression and poor response to dopamine replacement therapy [1, 2]. MSA has two variants, MSA-P and MSA-C, where parkinsonism and cerebellar features predominate, respectively. The majority of PD and MSA-P patients manifest similar clinical features which renders very challenging a correct differential diagnosis [2]. There exists criteria for the diagnosis of "probable" and "possible" MSA, based on clinical or/and imaging features, but the definite MSA diagnosis requires postmortem confirmation by a neuropathological examination [1]. Despite recent efforts, no validated biomarker is currently available for the differential diagnosis. There exists thus a strong need for such markers to improve diagnostic accuracy, particularly in early disease stages. An accurate early diagnosis is indeed essential not only in assessing prognosis and for treatment decisions, but also for understanding the underlying pathophysiology and for the development of new therapies [3].

It is now well established that dysarthria, a class of motor speech impairments resulting from neurological disorders, is an early clinical feature of PD and APD. Dysarthria is mostly caused by control or execution impairment of one or more sensorimotors. PD patients develop essentially hypokinetic dysarthria [4, 5] while MSA patients typically exhibit mixed dysarthria with various combination of hypokinetic and ataxic components [6, 5]. Dysarthria can manifest in all the levels of speech production [7]. In particular, the articulatory mechanism can be affected which causes deficits in range, strength, timing, stability and precision of articulators [8]. One of the most common manifestation of such deficits is imprecise consonant articulation. In the pioneer work [4], consonant realization was perceptually found to be one of the most deviant speech dimensions in PD.

Consonants distortion across various diseases have been typically assessed using perceptual evaluation [8, 9, 10, 11, 12, 13, 14]. During the last 2 decades, a considerable effort has been produced to develop objective measures that assess consonant distortions in PD [15, 16, 17, 18, 19, 20, 21, 22]. In these studies, voice onset time (VOT) has been the most analyzed feature but with rather contradictory outcomes [23, 24, 17, 20]. On the other hand, only few studies have addressed consonants distortion in differential diagnosis between PD and APD [13, 25, 20, 26, 21]. The same statement actually holds for dysarthria-based differential diagnosis in general. Indeed, while

there exists a large amount of work on comparing PD and HC speech, there is only few studies on comparison/discrimination between PD and APD or between APD subgroups [27, 28, 29, 30, 31, 32, 33, 34].

As for the French language, to the best of our knowledge, there exists no study comparing consonant production between PD and MSA. In [13], a comparison has been subjectively performed (spectrogram visual inspection) but between PD, Amyotrophic Lateral Sclerosis and Cerebellar Ataxia. In this paper, we present the first study on French consonant distortion for the purpose of differential diagnosis between PD and MSA-P. We carry out a subjective and objective analysis of word-initial consonants using pseudo-words, called Logatomes [35]. Among all the consonants, obstruents (plosives and fricatives) yield the most interesting results. In particular, we show that voiced obstruents manifest appealing distinctive impairments, in term of devoicing and VOT duration. We then provide a 2-dimensional analysis over these two deviant speech dimensions. This leads us to build a decision model which discriminates between PD and MSA-P with a good accuracy.

The paper is organized as follows. Section 2 describes the speech database. Section 3 presents a brief description of the analysis tools. Results are presented in Section 4. Discussion and conclusion are given in the last section.

2. Database

From 2018 to the time of writing this paper, a total of 43 French speakers were recruited in the framework of a research project involving the neurology and ENT departments of 2 French university hospitals (recruitment is continuing). 20 patients (5 females and 15 males) were diagnosed with idiopathic PD, with a mean age of 60 and a mean symptom duration of 4 years. 12 subjects (8 females and 4 males) were diagnosed with MSA-P, with a mean age of 67 and a mean symptom duration of 3.5 years. 11 healthy controls (HC) with a mean age of 56 (6 female and 5 male) with no history of neurological or communication disorders were recruited. ENTs carried out all the recording sessions for all participants. Each participant performed several speech tasks including sustained phonations, syllables repetition, a reading task, a monologue and a set of 25 isolated pseudo-words called Logatomes (other non-speech biosignals are also recorded). We used only the Logatomes dataset in this study. The speech signals were recorded with 48kHz sampling frequency and 16 bit resolution by a headmount condenser microphone (t.bone HC 444 TWS) placed at a distance of approximately 5cm from the speaker's mouth. Ethics approval was obtained prior to recruitment and all participants gave written informed consent.

3. Analysis tools

This paper is dedicated to the analysis of the production of consonants extracted from the Logatomes. The latter have the advantage of being an easy speech task for patients (independently on their native language) and easy to process (even manually). We performed an auditory and visual examination of the 25 Logatomes produced by each participant. Auditory examination was performed by listening carefully several times to all the audio files. Visual examination was carried out by inspection of waveform and wide-band spectrogram using Praat [36]. As a consequence of this processing, we manually annotated all the speech signals. We followed the criteria of [24] to set the boundaries of the different phonetic units, we refer to [24] for

details.

During the examination, we assessed the distortion of consonants (and vowels) when auditorily or/and visually perceived. This analysis showed that, among the 25 Logatomes, voiced obstruents were significantly distorted, in term of the manifestation of devoicing. We mention that this assessment was based on word-initials only (in CV format) in order to avoid co-articulation and speaking rate effects. We thus focused on the 3 voiced plosives /b/, /d/ and /g/ extracted from the Logatomes "berdo", "dirou" and "guizant", and the 3 voiced fricatives /v/, /z/, and /ʒ/ extracted from the Logatomes 'vonia", "zacu", and "jinin". After devoicing analysis, we performed the traditional voice onset time (VOT) analysis of the plosives using the manually segmented units. Finally we combined the two analysis, devoicing and VOT, in order to build a decision model for differential diagnosis.

For the statistical analysis, data normality of each acoustic feature was evaluated by the one-sample Kolmogorov-Smirnov test. No data was found Gaussian, the pairwise Kruskal-Wallis test to measure difference between groups was used. Statistical significance was set at a p-value $p < 0.05$.

4. Results

4.1. Devoicing analysis

4.1.1. Visual inspection of spectrograms

By visual inspection of the spectrograms, we assessed devoicing by the total or partial absence of voicing bars in the realization of voiced obstruents. Figure 1 shows an example of spectrograms of the consonant /g/ pronounced normally by a HC, with a partial devoicing by a PD and with a total devoicing by an MSA-P patient. We observed an other phenomenon which could be considered as partial devoicing, the occurrence of voicing bars with weak energy. However, for sake of clarity, reproducibility and to reduce subjectivity effects, we did not use this criterion in our assessment of devoicing.

We found that 67% of MSA-P and 15% of PD presented devoicing in at least one obstruent. In particular, we did not observe any devoicing of /b/ nor /d/ in PD while 42% of MSA-P showed devoicing in these consonants. This suggests that devoicing of /b/ or/and /d/ could be a signature of MSA-P. We mention here however that [13] reported that 37% of PD presented devoicing in /d/ or /g/. In our data, only 10% of PD showed devoicing in /g/ (and thus in /d/ or /g/), while 33% of MSA-P showed devoicing in /d/ or /g/. This difference in PD might be due to the relatively small size of our dataset as compared to the one of [13]. As for voiced fricatives, 33% of MSA-P and 15% of PD presented devoicing in at least one of /v/, /z/ and /ʒ/.

4.1.2. Objective assessment

We now provide an objective measure to detect devoicing. Given a labeled consonant, a simple way to assess total or partial absence of voicing is to consider the degree of voicing measure:

$$DV = \frac{number\ of\ voiced\ frames}{total\ number\ of\ frames}(\%)$$

We emphasize here that other objective criteria can be used to define and detect devoicing, our measure is however easy to interpret and to reproduce. Given a speaker, we then defined his/her degree of devoicing DVT as the minimal DV value over all the voiced obstruents. We then heuristically assessed
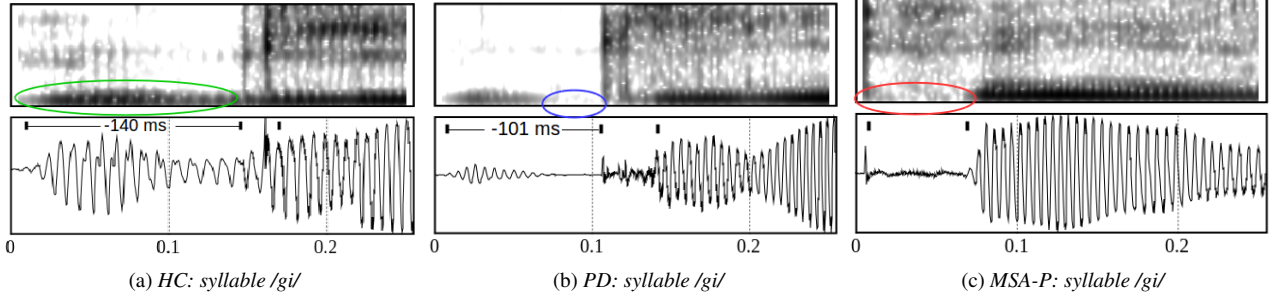
Figure 1: *Example of no/partial/total devoicing of /g/ in a HC/PD/MSA-P (top). Example of normal/shorter/vanishing VOT of /g/ for the same HC/PD/MSA-P (bottom)*

(a) *HC: syllable /gi/* (b) *PD: syllable /gi/* (c) *MSA-P: syllable /gi/*

the presence of devoicing if $DVT < 50\%$. Using this criterion, the assessment of devoicing matched perfectly with our visual observations, that is, the same 67% of MSA-P and 15% of PD presented a $DVT < 50\%$. The value of DVT is shown in Figure 2 for each speaker (see the projection over the DVT dimension). Along this dimension, one can note the large margin between subjects manifesting devoicing and the others. This is confirmed by the statistical test of group difference, as shown in Table 1. This suggests that devoicing is generally strong when it occurs, and thus easy to detect objectively by the standard tool Praat.

Overall, these results show that devoicing can be a valuable cue for differential diagnosis between PD and MSA-P. However, this cue alone is not sufficient to achieve this diagnosis with a high accuracy.

4.2. VOT analysis of voiced plosives

As mentioned earlier, VOT is among the most studied features in consonant distortion. VOT is generally associated with plosives and is defined as the duration between the vocal fold vibration starts relative to the release of the plosive (there exist however VOT definitions for other consonant types [37]). In the case of voiced plosive, vibration begins before the release, VOT is thus considered as negative. When negative VOT tends to 0, it actually corresponds to a total devoicing. In order to avoid a potential dependency on speaking rate, VOT ratio (VOTR) is sometimes considered. VOTR is defined as VOT divided by the duration of whole syllable [24].

The purpose of this section is to determine whether VOT analysis of voiced plosives (/b/, /d/ and /g/) can yield another distinctive cue (hopefully complementary to devoicing). Using our manual segmentation, we computed the VOT and VOTR statistical group difference between HC and PD, HC and MSA-P, PD and MSA-P. Table 1 shows the obtained p-value of each group difference. We first observed that, for VOT, statistical significance between MSA and the other groups was achieved only for /g/. More interestingly, this impairment was more severe in MSA-P than in PD. The waveforms of Figure 1 shows an example of such a distortion. This trend was confirmed by VOTR with an additional group difference between PD and HC. Globally, this is in accordance with the findings of [20] which reported shorter VOT and lower VOTR for MSA averaged on all voiced plosives (with Czech patients). We cannot however confirm the same statement for /b/ and /d/. On the other hand, we can confidently consider that VOT/VOTR of /g/ is a valuable cue for the differential diagnosis. However, as devoicing, this cue alone is not sufficient to achieve this diagnosis with a high accuracy.

Table 1: *Univariate statistical group difference of DVT and VOT/VOTR (for the 3 voiced plosives /b/,/d/ and /g/). Bold values indicate significant group difference ($p < 0.05$)*

Feature	Consonant	HC vs PD	HC vs MSA-P	PD vs MSA-P
		p-value		
DVT	All	0.12	**0.0008**	**0.009**
VOT	/b/	0.36	0.15	0.41
	/d/	0.8	0.9	0.66
	/g/	0.07	**0.004**	**0.03**
VOTR	/b/	0.08	**0.001**	**0.029**
	/d/	0.26	**0.03**	0.161
	/g/	**0.002**	**0.0009**	**0.014**

4.3. Classification

Given the findings of the previous sections, it is natural to proceed with an analysis over the 2 deviant speech dimensions, devoicing and VOT of /g/ ($VOT_{/g/}$). Using a decision tree classifier with a leave-one-speaker-out (LOSO) cross-validation, we obtained an accuracy of 69%, with a high sensitivity (correctly classified MSA-P) of 83% but a low specificity of 60%. The tree obtained with all data in training is shown in Figure 3 (Train accuracy 72%). This means that a mis-diagnosis of MSA-P presenting devoicing or short VOT of /g/ is unlikely. This statement does not hold for PD.

Figure 2a shows the biplot of DVT w.r.t to $VOT_{/g/}$ with decision thresholds obtained using all data in training. We noted that the estimated thresholds, $h_1 = 45\%$ for DVT and $h_2 = -60ms$ for $VOT_{/g/}$, might be considered as *not* overfitted to our (small size) data. A possible justification is the following. First, it would be reasonable to consider that devoicing occurs when a patient presents a DV less that 45% in a voiced obstruent. Second, using our HC data, the mean/standard deviation of the VOT of /g/ is $-103/22(ms)$. This is in accordance with the $-109/32(ms)$ reported in [20] and [24] (the latter reported the mean only). Thus h_2 could be reasonably considered as a threshold above which one can confidently consider that a VOT impairment of /g/ is occurring. These two arguments need however to be confirmed by additional data and analysis.

We then carried out the same analysis as above but using VOTR. Using again a LOSO decision tree classifier, the specificity significantly increased to 83% and the sensitivity to 85%. The resulting accuracy score was **84%**. The tree obtained with all data in training is shown in Figure 3 (Train accuracy 87%). Figure 2b shows the biplot of DVT w.r.t to $VOTR_{/g/}$ with decision thresholds obtained using all data in training. We see now that, along the $VOTR_{/g/}$ dimension, a separation appears between the 3 MSA-P and 5 PD which were confused using $VOT_{/g/}$ (right top rectangle of 2b).
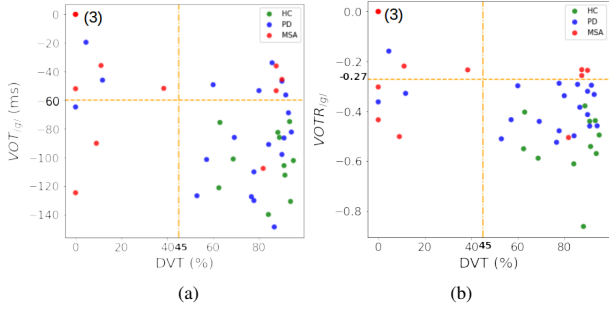
Figure 2: *Biplot of $DVT(\%)$ w.r.t to $VOT_{/g/}$ and $VOTR_{/g/}$ (dotted line represent decision thresholds using all data in training); (3) means that 3 MSA-P patients have same coordinates (total devoicing)*
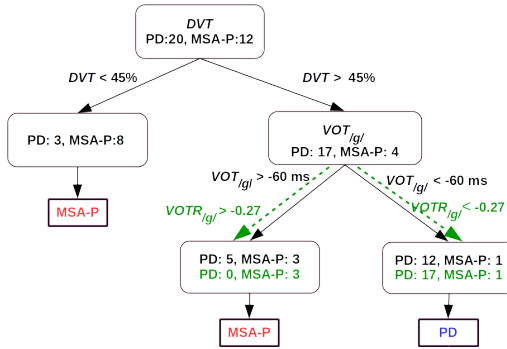

Figure 3: *Decision tree using DVT and $VOT_{/g/}$ or $VOTR_{/g/}$ (in green) dimensions for discrimination between PD and MSA-P*

This time however we can not argue that the VOTR threshold (-0.27) could be consider as reasonable decision threshold as we do not have reference values of VOTR. This threshold might thus be overfitted to our data. Therefore, given the small amount of instances of /g/ and its following vowels, we cannot confidently claim that VOTR is a better feature for discrimination than VOT. These results show however that the prevoicing duration of /g/ (and probably all voiced plosives) could be a complementary cue to devoicing of obstruents in order to achieve a high accuracy differential diagnosis between PD and MSA-P.

Overall, the results (along with literature reporting) show that devoicing of voiced obstruents and VOT of /g/ are 2 distinctive and deviant speech dimensions which are worth considering in the differential diagnosis between PD and MSA-P.

5. Discussion and conclusion

This work constitutes the first study that attempts to highlight distinctive cues in the distortion of French voiced obstruents realization in PD and MSA-P. Our results partially confirmed previous findings on VOT with other languages [25, 20]. Indeed, we found that VOT of the voiced plosive /g/ was significantly reduced in MSA-P while it was natural for most PD. On the other hand, VOT in not the only factor of the distortion of voiced plosives (and obviously fricatives). We showed indeed that the absence of voicing leads was the main factor of

voiced obstruents distortions and is the most distinctive cue between PD and MSA-P (in the production of voiced obstruents). Moreover, there was a perfect match in devoicing assessment between perceptual and objective evaluations. This supports a potential use of devoicing in clinical practice as an additional tool for the examination of patients with a suspicion of MSA-P. We also showed that the combination of VOT and devoicing can significantly improve the differential diagnosis accuracy.

VOT impairment can be explained by a difficulty in initiating articulation resulting from a deficit in maintaining the speech motor program [20]. The latter is a characteristic of hypokinetic dysarthria which is a known feature in both PD and MSA. An accurate production of word-initial voiced plosives requires a precise coordination between glottal opening and articulatory closure. Devoicing is a manifestation of an impairment of such coordination. This is a characteristic of ataxic dysarthria which is known to manifest in MSA. Our results are thus in accordance with the consensus that PD develop essentially hypokinetic dysarthria while MSA develop a mixed type dysarthria. More importantly, since ataxia seems to be responsible for devoicing, the latter might manifest in early disease stages. If proven, devoicing would thus constitute a valuable deviant speech dimension to consider in early differential diagnosis.

There are some limitations to our study. The most significant one is obviously the relatively small size of the dataset due to the difficulty of recruiting patients, particularly with a rare disease such as MSA-P. We are however still continuing the effort of recruitment. Moreover, the dataset is unbalanced in gender, we cannot thus exclude that our findings are biased by gender-specific effects. Another limitation is that we used only one consonant instance per speaker, we do not thus know how the results stand to intra-speaker pronunciation variability. We can expect however that the effect of such variability is reduced by the restriction to word-initials. From this perspective, our study should be considered as a promising first step in the analysis of French voiced obstruents in PD and MSA-P. Our findings need to be confirmed by additional data. This is the purpose of our on going research.

6. Acknowledgements

Author contribution:
1) Research project: A. Conception, B. Organisation, C. Execution
2) Acoustic Analysis: A. Design, B. Execution, C. Review&Critique
3) Manuscript: A. Writing of the first draft, B. Review & Critique
KD: 1A-C, 2A+B, 3A. **BD**: 2B, 3B. **APL, AFS, OR**: 1A, 3B. **SM, VW**: 1A+C, 3B. **WM**: 1A-C, 3B.

7. References

[1] S. G. et al., "Second consensus statement on the diagnosis of multiple system atrophy," *Neurology*, vol. 71, no. 9, pp. 670–676, 8 2008.

[2] A. Fanciulli and G. K. Wenning, "Multiple-system atrophy," *New England Journal of Medicine*, vol. 372, no. 3, pp. 249–263, 2015.

[3] W. G. Meissner, P.-O. Fernagut, B. Dehay, P. Péran, A. Pavy-Le Traon, A. Foubert-Samier, M. Lopez Cuina, E. Bezard, F. Tison, and . Rascol, "Multiple system atrophy: Recent developments and future perspectives," *Movement Disorders*, vol. 34, no. 11, pp. 1629–1642, 2019.

[4] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246–269, 1969.

[5] J. Rusz, C. Bonnet, J. Klempir, T. Tykalova, E. Baborov, M. Novotny, A. Rulseh, and E. Ruzicka, "Speech disorders reflect differing pathophysiology in parkinson's disease, progressive supranuclear palsy and multiple system atrophy," *Journal of neurology*, vol. 262, pp. 992–1001, 2015.

[6] K. Kluin, S. Gilman, M. Lohman, and L. Junck, "Characteristics of the dysarthria of multiple system atrophy," *Archives of Neurology*, vol. 53, no. 6, pp. 545–548, 1996.

[7] J. Duffy, *Motor Speech Disorders Substrates, Differential Diagnosis, and Management*. Elsevier, 2013.

[8] J. A. Logemann and H. B. Fisher, "Vocal tract control in parkinson's disease: Phonetic feature analysis of misarticulations," *The Journal of speech and hearing disorders*, vol. 46, pp. 348–52, 12 1981.

[9] L. Hartelius, H. Gustavsson, M. Astrand, and B. Holmberg, "Perceptual analysis of speech in multiple system atrophy and progressive supranuclear palsy," *Journal of Medical Speech-Language Pathology*, vol. 14, pp. 241–248, 2006.

[10] D. Duez, "Acoustic analysis of occlusive weakening in parkinsonian french speech," *International Congress of Phonetic Sciences*, 08 2007.

[11] S. Sachin, G. Shukla, V. Goyal, S. Singh, V. Aggarwal, M. Behari *et al.*, "Clinical speech impairment in parkinson's disease, progressive supranuclear palsy, and multiple system atrophy," *Neurology India*, vol. 56, p. 122, 2008.

[12] Y. Kim, R. Kent, J. Kent, and J. Duffy, "Perceptual and acoustic features of dysarthria in multiple system atrophy," *Journal of Medical Speech-Language Pathology*, vol. 18, pp. 66–71, 2010.

[13] T. Antolík and C. Fougeron, "Consonant distortions in dysarthria due to parkinson's disease, amyotrophic lateral sclerosis and cerebellar ataxia," pp. 2152–2156, 2013.

[14] N. Miller, U. Nath, E. Noble, and D. Burn, "Utility and accuracy of perceptual voice and speech distinctions in the diagnosis of parkinson's disease, psp and msa-p," *Neurodegenerative Disease Management*, vol. 7, pp. 191–203, June 2017.

[15] H. Ackermann and I. Hertrich, "Voice onset time in ataxic dysarthria," *Brain and Language*, vol. 56, no. 3, pp. 321 – 333, 1997.

[16] H. Ackermann, S. Gräber, I. Hertrich, and I. Daum, "Phonemic vowel length contrasts in cerebellar disorders," *Brain and Language*, vol. 67, no. 2, pp. 95 – 109, 1999.

[17] M. Novotny, J. Rusz, R. Cmejla, and E. Ruzicka, "Automatic evaluation of articulatory disorders in parkinson's disease," *IEEE-ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 22, pp. 1366–1378, 2014.

[18] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. V. Bonilla, S. Skodda, J. Rusz, and E. Nöth, "Voiced/unvoiced transitions in speech as a potential bio-marker to detect parkinson's disease." in *INTERSPEECH*. ISCA, 2015, pp. 95–99.

[19] Y. Kim, "Acoustic characteristics of fricatives /s/ and /ʃ/ produced by speakers with parkinson's disease," *Clinical Archives of Communication Disorders*, vol. 2, 01 2017.

[20] T. Tykalova, J. Rusz, J. Klempir, R. Cmejla, and E. Ruzicka, "Distinct patterns of imprecise consonant articulation among parkinson's disease, progressive supranuclear palsy and multiple system atrophy," *Brain and language*, vol. 165, pp. 1–9, February 2017.

[21] J. Rusz, T. Tykalová, G. Salerno, S. Bancone, J. Scarpelli, and M. Pellecchia, "Distinctive speech signature in cerebellar and parkinsonian subtypes of multiple system atrophy," *Journal of Neurology*, vol. 266, 03 2019.

[22] U. Goswami, S. Nirmala, V. C M, S. Kalita, and S. Prasanna, "Analysis of articulation errors in dysarthric speech," *Journal of Psycholinguistic Research*, vol. 49, 10 2019.

[23] A. Flint, S. Black, I. Campbell-Taylor, G. Gailey, and C. Levinton, "Acoustic analysis in the differentiation of parkinson's disease and major depression," *Journal of Psycholinguistic Research*, vol. 21, no. 5, pp. 383–399, 1992.

[24] E. Fischer and A. M. Goberman, "Voice onset time in parkinson disease," *Journal of Communication Disorders*, vol. 43, no. 1, pp. 21 – 34, 2010.

[25] M. Saxena, M. Behari, S. S. Kumaran, V. Goyal, and V. Narang, "Assessing speech dysfunction using bold and acoustic analysis in parkinsonism," *Parkinsonism & Related Disorders*, vol. 20, no. 8, pp. 855 – 861, 2014.

[26] J. Hlavnička, T. Tykalová, R. Čmejla, J. Klempir, E. Růžička, and J. Rusz, "Dysprosody differentiate between parkinson's disease, progressive supranuclear palsy, and multiple system atrophy," in *Proceedings of Interspeech*, 2017, pp. 1844–1848.

[27] N. Chakraborty, T. Roy, A. Hazra, A. Biswas, and K. Bhattacharya, "Dysarthric bengali speech: A neurolinguistic study," *Journal of postgraduate medicine*, vol. 54, pp. 268–72, 10 2008.

[28] S. Skodda, W. Visser, and U. Schlegel, "Acoustical analysis of speech in progressive supranuclear palsy," *Journal of Voice*, vol. 25, pp. 725–731, November 2011.

[29] Y. Huh, J. Park, M. Suh, S. Lee, J. Kim, Y. Jeong, H. Kim, and J. Cho, "Differences in early speech patterns between parkinson variant of multiple system atrophy and parkinson's disease," *Brain and language*, vol. 147, pp. 14–20, May 2015.

[30] J. Hlavnička, R. Čmejla, T. Tykalová, K. Sonka, E. Růžička, and J. Rusz, "Automated analysis of connected speech reveals early biomarkers of parkinson's disease in patients with rapid eye movement sleep behaviour disorder," *Scientific Reports*, vol. 7, 12 2017.

[31] M. Brückl, A. Ghio, and F. Viallet, "Measurement of tremor in the voices of speakers with parkinson's disease," *Procedia Computer Science*, vol. 128, pp. 47 – 54, 2018.

[32] G. Li, K. Daoudi, J. Klempír, and J. Rusz, "Linear classification in speech-based objective differential diagnosis of parkinsonism," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, 2018, pp. 5999–6003.

[33] B. Das, K. Daoudi, J. Klempír, and J. Rusz, "Towards disease-specific speech markers for differential diagnosis in parkinsonism," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, 2019, pp. 5846–5850.

[34] R. Kowalska-Taczanowska, A. Friedman, and D. Koziorowski, "Parkinson's disease or atypical parkinsonism? the importance of acoustic voice analysis in differential diagnosis of speech disorders," *Brain and Behavior*, vol. 10, no. 8, 2020.

[35] M. Lalain, A. Ghio, L. Giusti, D. Robert, C. Fredouille, and V. Woisard, "Design and development of a speech intelligibility test based on pseudowords in french: Why and how?" *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 7, pp. 2070–2083, 2020.

[36] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]," *Version 5.3.51*, 2013.

[37] A. Abramson and D. Whalen, "Voice onset time (vot) at 50: Theoretical and practical issues in measuring voicing distinctions," *Journal of Phonetics*, vol. 63, pp. 75 – 86, 2017.