Perceptual Analysis of Speech Signals from People with Parkinson's Disease

J.R. Orozco-Arroyave[1,2], J.D. Arias-Londoño[1],
J.F. Vargas-Bonilla[1], and Elmar Nöth[2]

[1] Universidad de Antioquia, Medellín
[2] Friedrich Alexander Universität, Erlangen-Nürnberg

Abstract. Parkinson's disease (PD) is a neurodegenerative disorder of the nervous central system and it affects the limbs motor control and the communication skills of the patients. The evolution of the disease can get to the point of affecting the intelligibility of the patient's speech.

The treatments of the PD are mainly focused on improving limb symptoms and their impact on speech production is still unclear. Considering the impact of the PD in the intelligibility of the patients, this paper explores the discrimination capability of different perceptual features in the task of automatic classification of speech signals from people with Parkinson's disease (PPD) and healthy controls (HC). The experiments presented in this paper are performed considering the five Spanish vowels uttered by 20 PPD and 20 HC.

The considered set of features includes linear prediction coefficients (LPC), linear prediction cepstral Coefficients (LPCC), Mel-frequency cepstral coefficients (MFCC), perceptual linear prediction coefficients (PLP) and two versions of the relative spectra coefficients (RASTA).

Accordin the results for vowels /e/ and /o/ it is not enough to consider one kind of perceptual features, it is required to perform combination of different coefficients such as PLP, MFCC and RASTA. For the case of the remaining vowels, the best results are obtained considering only one kind of perceptual features, PLP for vowel /a/ and MFCC for vowels /i/ and /u/.

Keywords: Perceptual analysis, Parkinson's disease, linear prediction, relative spectra analysis.

1 Introduction

Parkinson's disease (PD) is a neurodegenerative disorder that results from the death of dopaminergic cells in the substantia nigra, a region of the mid-brain. PD is the second more prevalent neurological disorder after the Alzheimer's disease [1]. About 1% of the people older than 65 suffer from this disease and in Colombia the prevalence of PD is around 172.4 per each 100.000 inhabitants [2]. People with Parkinson's disease (PPD) commonly develop speech impairments affecting different aspects such as respiration, phonation, articulation and prosody [3]. It is already demonstrated that the phonation problems in PPD are related to the

J.M. Ferrández et al. (Eds.): IWINAC 2013, Part I, LNCS 7930, pp. 201–211, 2013.

vocal fold bowing and incomplete vocal fold closure [4], [5]. This behavior also generates intelligibility problems for the speech of the patients, affecting their communications skills and their capability for interacting with other people [6].

Medical treatments such as neuropharmacological and neurosurgical are focused on improving limb symptoms, but their impact on speech production is still unclear [7]. Due to this fact, in the last few years several published works have been focused on the automatic classification of speech recordings from PPD and from healthy controls (HC). In [8] the authors employ acoustic and prosodic features along with features derived from a two-mass model of the vocal folds. For the acoustic modeling of the speech signals, the authors consider a speech recognition model based on Gaussian Mixture Models (GMM) with 13 Mel-frequency Cepstral Coefficients (MFCC) and for the prosodic analysis they include fundamental frequency (F_0), energy, voiced and unvoiced segments and pitch periods all calculated on the voiced segments of a running speech recordings. The reported recognition rates are 88% with the acoustic model (MFCC) and 90.5% with the prosodic features.

In [9], the authors perform the automatic classification of PPD and HC considering four features: Harmonics to Noise Ratio (HNR), Recurrence Period Density Entropy (RPDE), Detrended Fluctuation Analysis (DFA) and Pitch Period Entropy (PPE). Their results indicate that, considering this set of features, it is possible to achieve classification rates of up to 91.4%.

On the other hand, the evolution of the PD through the time has been also studied. In [10] the authors form a features set composed by different dysphonia measures and analyze their correlation with the evolution of the Unified Parkinson Disease Rating Scale (UPDRS) in a period of six months. According to their results, the UPDRS scale can be mapped with a precision of up to 6 points, which is very close to the clinician's observations.

Considering that the PPD exhibit a loss of intelligibility in their speech [11], our hypothesis is that including perceptual information in the speech modeling processes is possible to achieve good results in the automatic classification of speech from PPD and HC. The perceptual analysis of pathological speakers have been addressed typically using different kind of coefficients. Some of them have been already used for speech signals from PD [8], but there is still a lack of understanding about the discrimination capabilities that these kind of features can provide for the automatic assessment of speech signals from PPD. Bearing this in mind, this work studies the contribution of six well known representation coefficients estimated over the five Spanish vowels and tries to establish which features are more suitable for the automatic classification of speech from PPD and HC.

The performed experiments include Linear Prediction Coefficients (LPC) [12], Linear Prediction Cepstral Coefficients (LPCC) [13], Mel-frequency Cepstral Coefficients (MFCC) [14], Perceptual Linear Prediction coefficients (PLP) [15] and two versions of the Relative Spectra coefficients (RASTA), those with cepstral filtering (RASTA-PLP-CEPS) and those without cepstral filtering (RASTA-PLP-SPEC) [16].

The rest of the paper is organized as follows. The section 2 includes the description of the methodology tha is being proposed in this work. In the section 3 gives the details of the experiments that are presented in the paper. The section 4 presents the results obtained on each experiment and finally, the section 5 exposed the conclusions derived from the presented work.

2 Methodology

Figure 1 depicts a block diagram of the steps carried out in the methodology presented in this work. The right side of the figure illustrates each stage of the methodology and the left side of the figure shows a brief explanation of that stages. The voice signal is first divided and windowed (with Hamming windows) into frames to perform a short-time analysis. After, the characterization stage takes place. In this work, six different sets of perceptual features are considered: LPC, LPCC, PLP MFCC and two versions of the RASTA-PLP coefficients, one with cepstral filtering and other one without such filtering. Once the features are calculated, a features selection process is performed for each set of features as in [17].

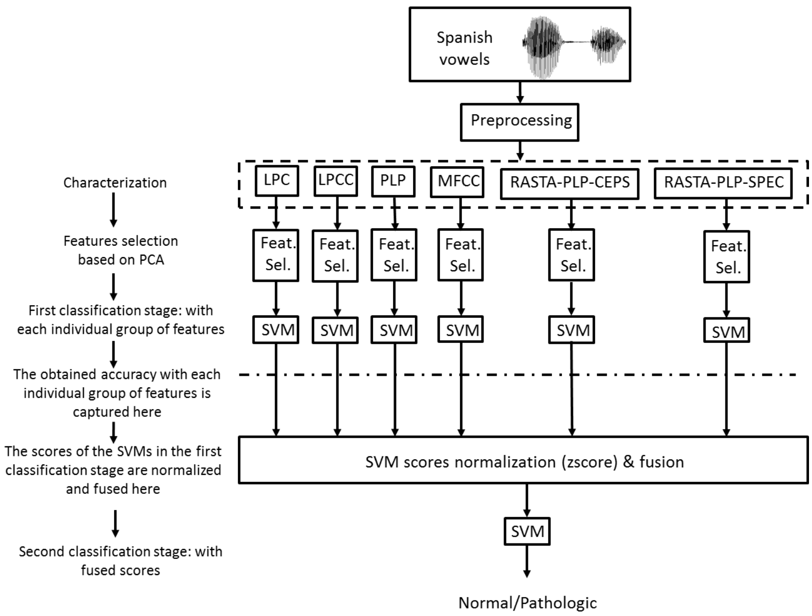

Fig. 1. Methodology for the perceptual analysis of speech from PPD

After the features selection a two layer classification scheme is used. The first stage of the classification process is performed per each kind of perceptual features by means of a support vector machine (SVM) with a radial basis Gaussian kernel. After, the results obtained with each SVM are combined into a new feature space. Once that results are combined, they are normalized using a zscore strategy and finally, the last stage of the classification process is done with another SVM.

In the following subsections, some details of each part of the methodology are presented.

2.1 Perceptual Modeling

As it was pointed out, in this work different perceptual coefficients are implemented with the aim of establish which of them are the more appropriate for the automatic classification of speech from PPD and HC. Considering their wide usage in speech modeling, the LPCs coefficients are included in this work. Linear predictive techniques have been applied not only for the LPCs calculation but also for the formants estimation. The information provided by the LPCs allows to perform articulation analysis in the speech of PPD [18]. This coefficients are able to model the vocal tract as a filter, thus considering that PPD are unable to have a total motor control of their vocal tract, the frequency response of that filter will be also abnormally changed.

A similar modeling to those that is performed by the LPCs in the frequency domain can be made in the cepstral domain. Such analysis is made through the LPCC coefficients. The LPCCs have demonstrated to be more robust in several speech modeling tasks mainly oriented to speech recognition processes [13]. In this work we want to validate their robustness in the automatic classification of speech from PPD and HC.

Additionally the MFCC coefficients have been included in the experiments presented here. This kind of coefficients can model irregular movements in the vocal tract [14] and have demonstrated to be efficient for modeling pathological speech signals, not only in the case of dysphonia detection [14], [19] but also for the analysis of dysarthric speech signals [8].

On the other hand, with the aim of including perceptual information to the modeling that is performed with LPCs, we decided to include the PLP coefficients estimated as in [15]. The LPC analysis assumes the same number of resonances on every frequency bands; however, there are evidences that demonstrate that beyond about 800Hz, the spectral resolution of hearing decreases with frequency [15]. Considering this drawback of the LPC analysis, Hynek Hermansky proposed to use a critical-band filtering over the linear predictive analysis in order to improve the resolution of the analyzed bands.

Other kind of features that are included in this work are the RASTA-PLP coefficients. This technique was presented by Hermansky and Morgan in [16].

The main assumption of this modeling strategy is that the human perception is less sensitive to slowly varying stimuli, thus it is possible to make the speech analysis less sensitive to slowly changing factors. For doing that, the critical-band filter that is used for PLP modeling is replaced by a filter bank with a sharp spectral zero at the zero frequency. The RASTA-PLP coefficients can be represented in spectral or in cepstral domains and the difference between both approaches is that in the spectral domain the resulting set of coefficients is according to the number of filtered bands and in the cepstral domain the resulting set of coefficients is according to the number of cepstral coefficients. In this work we use both representations to analyze their influence in the process of automatic classification of speech from PPD and HC.

2.2 Features Selection and Classification

The features selection strategy that is used in this work is based on principal component analysis and it was implemented as in [17]. After the application of the features selection algorithm the system has a sub set of features that is optimal in terms of its variance content.

The classification between recordings from PPD and HC is performed with a SVM that is trained with a Gaussian kernel [20]. This classifier is used because of its extensive usage in the state of the art for the automatic classification of pathological and healthy voices [9], [8], [17].

3 Experiments

3.1 Recording and Corpus of Speakers

The data for this study consists of speech recordings from 20 PPD and 20 HC sampled at $44.100Hz$ with 16 quantization bits. All of the recordings were captured in a sound proof booth. The people that participated in the recording sessions are balanced by gender and age: the ages of the men patients ranged from 56 to 70 (mean 62.9 ± 6.39) and the ages of the women patients ranged from 57 to 75 (mean 64.6 ± 5.62). For the case of the healthy people, the ages of men ranged from 51 to 68 (mean 62.6 ± 5.48) and the ages of the women ranged from 57 to 75 (mean 64.8 ± 5.65). All of the PPD have been diagnosed by neurologist experts and none of the people in the HC group has history of symptoms related to Parkinson's disease or any other kind of movement disorder syndrome.

The recordings consist of sustained utterances of the five Spanish vowels, every person repeated three times the five vowels, thus in total the database is composed of 60 recordings per vowel on each class. This database is built by *Universidad de Antioquia* in Medellín, Colombia.

3.2 Experimental Setup

The experiments performed in this work are carried out following the methodology exposed in figure 1. The voice recordings are preprocessed by Hamming windows of 40ms with and overlap of 20ms. The perceptual features are calculated for each windowed frame of speech. Each voice signal is represented by the sets of feature vectors, each vector contains the values of the parameters for each frame. The final feature vector per voice signal is composed by the estimations of mean, standard deviation, skewness and kurtosis of the values obtained per feature trough the frames.

The number of coefficients is fixed in 12 for every kind of perceptual features but for the case of RASTA-PLP-SPEC a total of 27 coefficients are estimated because this is the number of frequency bands that are filtered. For RASTA-PLP-CEPS the number of coefficients is 12 because it is the number of cepstral coefficients used. Considering that four statistics are taken from each kind of features, the sets of parameters contain a total of 48 measures for the case of LPC, LPCC, PLP, MFCC and RASTA-PLP-CEPS. For the case of RASTA-PLP-SPEC the number of measures is 108.

The tests performed over the proposed system have been made following the strategy indicated in [21]. The 70% of the data are used for the feature selection and for training the classifier and the remaining 30% is for testing; the different subsets for training and testing are randomly formed. As it was exposed in section 2, for each pair of training and testing subsets the two stages of classification are made: the first is when only each set of perceptual coefficients are considered individually and the second is when the scores obtained in the first classification stage are combined. Each stage of the classification process is repeated ten times per each pair of subsets (training and testing), forming a total of 100 independent realizations of the experiment.

In order to look for the best performance of the system, the scores obtained in the first classification stage are incrementally combined. The order of that fusion is according to the classification rate obtained in the individual classification stage.

4 Results and Discussion

The results obtained in the first stage of the classification process, when each subset of perceptual features are used per each vowel, are presented in table 1. The highlighted items correspond to the best results obtained over all of experiments. Note that for the vowels /a/, /i/ and /u/ the best results correspond to those obtained with only one kind of perceptual coefficients. For vowel /a/ the PLP parameters exhibited the best results, while for vowels /i/ and /u/ the best features were the MFCCs in both cases. The obtained results for the

Table 1. Results obtained in the automatic classification of speech signals from PPD and HC using each kind of perceptual features individually

Vowel	Features	Individual accuracy	Specificity	Sensitivity
/a/	LPC	60,58 % 6,98	65,44 % 12,20	55,72 % 13,97
	LPCC	64,86 % 8,08	60,83 % 14,10	68,89 % 7,49
	MFCC	59,72 % 9,04	66,33 % 15,95	53,11 % 7,60
	PLP	**76,19 % 9,04**	**80,22 % 10,69**	**72,17 % 11,05**
	RASTA-PLP-CEPS	56,72 % 6,17	41,22 % 8,98	72,22 % 7,33
	RASTA-PLP-SPEC	59,47 % 7,58	60,11 % 12,31	58,83 % 10,47
/e/	LPC	66,25 % 6,82	61,56 % 12,85	70,94 % 9,53
	LPCC	63,41 % 8,45	60,17 % 14,37	66,67 % 12,48
	MFCC	66,97 % 11,98	64,61 % 16,46	69,33 % 14,87
	PLP	66,39 % 7,23	67,17 % 9,57	65,61 % 13,93
	RASTA-PLP-CEPS	66,08 % 6,04	64,67 % 9,03	67,50 % 10,60
	RASTA-PLP-SPEC	71,28 % 7,38	66,06 % 13,14	76,50 % 9,51
/i/	LPC	58,17 % 8,76	53,72 % 12,14	62,61 % 14,83
	LPCC	71,61 % 6,9	71,44 % 10,58	71,78 % 11,32
	MFCC	**75,30 % 8,43**	**78,78 % 10,20**	**71,83 % 10,53**
	PLP	70,83 % 7,49	70,50 % 7,07	71,17 % 12,10
	RASTA-PLP-CEPS	66,33 % 5,39	59,94 % 14,71	72,72 % 12,97
	RASTA-PLP-SPEC	69,64 % 5,59	67,44 % 6,59	71,83 % 8,22
/o/	LPC	59,22 % 8,42	64,89 % 12,30	53,56 % 11,55
	LPCC	71,83 % 5,88	82,78 % 6,62	60,89 % 8,39
	MFCC	78,31 % 5,32	87,17 % 8,72	69,44 % 9,39
	PLP	69,97 % 6,84	71,06 % 12,88	68,89 % 9,65
	RASTA-PLP-CEPS	71,11 % 6,89	70,22 % 19,27	72,00 % 15,90
	RASTA-PLP-SPEC	68,61 % 7,62	61,17 % 7,70	76,06 % 10,24
/u/	LPC	62,61 % 8,37	63,94 % 9,31	61,28 % 8,19
	LPCC	64,86 % 7,67	62,44 % 10,36	67,28 % 10,84
	MFCC	**76,28 % 6,11**	**82,44 % 10,75**	**70,11 % 7,14**
	PLP	73,14 % 11,11	76,89 % 9,44	69,39 % 13,16
	RASTA-PLP-CEPS	67,78 % 6,75	52,89 % 12,86	82,67 % 9,27
	RASTA-PLP-SPEC	62,94 % 5,89	52,61 % 11,17	73,28 % 10,26

vowels /e/ and /o/ have not any highlighted row due to the best results for that vowels were obtained when several perceptual coefficients are combined. Such results are presented in table 2. The results obtained in the second stage of the classification process are presented in table 2. Note that in this case the best results are obtained for vowels /e/ and /o/ when five subsets of features are combined. For the case of vowel /e/ the considered features are RASTA-PLP-SPEC, MFCC, PLP, LPC and RASTA-PLP-CEPS while for vowel /o/ the set of features include MFCC, LPCC, RASTA-PLP-CEPS, PLP and RASTA-PLP-SPEC.

It is interesting to note that the best results in vowels /e/ and /o/ include MFCC, PLP and both versions of the RASTA-PLP coefficients for both vowels. It indicates that with the aim of achieving better results in the automatic classification of speech signals from PPD and HC, the characterization with perceptual features must consider more than one kind of coefficients for vowels /e/ and /o/. In general, the performance obtained with the vowels /e/ and /o/ are higher than in the other cases; however, note that for reaching such results the inclusion of five features were required.

Table 2. Results obtained when each subset of features are incrementally combined

Vowel		Two subsets	Three subsets	Four subsets	Five subsets	six subsets
/a/	**Accuracy**	70,67 % 10,63	71,89 % 7,33	70,39 % 6,53	70,06 % 7,58	68,28 % 6,32
	Specificity	72,94 % 14,17	78,67 % 11,31	76,67 % 11,70	75,17 % 9,59	66,50 % 70,65
	Sensitivity	68,39 % 11,53	65,00 % 10,75	64,50 % 8,81	64,64 % 8,80	70,06 % 8,67
/e/	**Accuracy**	73,11 % 7,33	74,11 % 7,67	75,86 % 6,16	**77,22** % **6,28**	77,28 % 7,00
	Specificity	62,28 % 13,64	64,39 % 12,47	65,33 % 14,47	**66,39** % **13,78**	67,94 % 14,56
	Sensitivity	81,50 % 10,67	83,56 % 12,44	86,50 % 9,48	**88,06** % **7,87**	86,78 % 6,50
/i/	**Accuracy**	70,47 % 11,00	72,53 % 8,66	74,86 % 8,78	75,00 % 6,69	74,78 % 5,69
	Specificity	68,78 % 11,40	74,50 % 9,29	74,06 % 10,68	69,44 % 11,27	69,00 % 11,71
	Sensitivity	72,06 % 18,14	70,44 % 20,21	75,67 % 18,67	80,56 % 13,86	80,78 % 13,93
/o/	**Accuracy**	78,92 % 8,89	80,81 % 7,39	79,77 % 6,77	**81,08** % **6,82**	80,78 % 6,91
	Specificity	83,44 % 6,48	80,56 % 9,51	81,72 % 8,27	**80,22** % **8,02**	73,39 % 7,74
	Sensitivity	74,39 % 14,56	81,06 % 9,39	77,83 % 7,93	**81,94** % **7,71**	82,11 % 7,73
/u/	**Accuracy**	76,08 % 10,76	75,64 % 7,97	76,5 % 7,36	76,25 % 6,48	76,14 % 7,13
	Specificity	77,33 % 12,80	67,28 % 9,50	68,50 % 8,85	67,22 % 7,22	65,72 % 8,61
	Sensitivity	74,17 % 9,68	83,72 % 8,02	84,56 % 8,36	85,33 % 6,71	86,67 % 7,89

The table 3 shows the best results obtained for each vowel in terms of sensitivity and specificity. The best sensitivity and specificity are obtained with vowels /e/ and /u/ respectively; however, note that the sensitivity and specificity are more balanced for the vowel /o/. In order to show the best results per vowel in a more compact way and following the methodology presented in [21], the figure 2 with the detection error tradeoff (DET) curve is included. Note that the more balanced behavior is exhibited with the vowel /o/, while the vowels /a/, /e/ and /i/ show very unbalanced results.

Table 3. Best results obtained per vowel

Vowels	**/a/**	**/e/**	**/i/**	**/o/**	**/u/**
Accuracy	76,19 ± 9,04	77.22 ± 6,28	75.30 ± 8,43	81.08 ± 6,82	76.28 ± 6,11
Sensitivity	72,17 ± 11,05	88,06 ± 7,87	71,83 ± 10,53	81,94 ± 7,71	70,11 ± 7,14
Specificity	80,22 ± 10,69	66,39 ± 13,78	78,78 ± 10,20	80,72 ± 8,02	82,44 ± 10,75

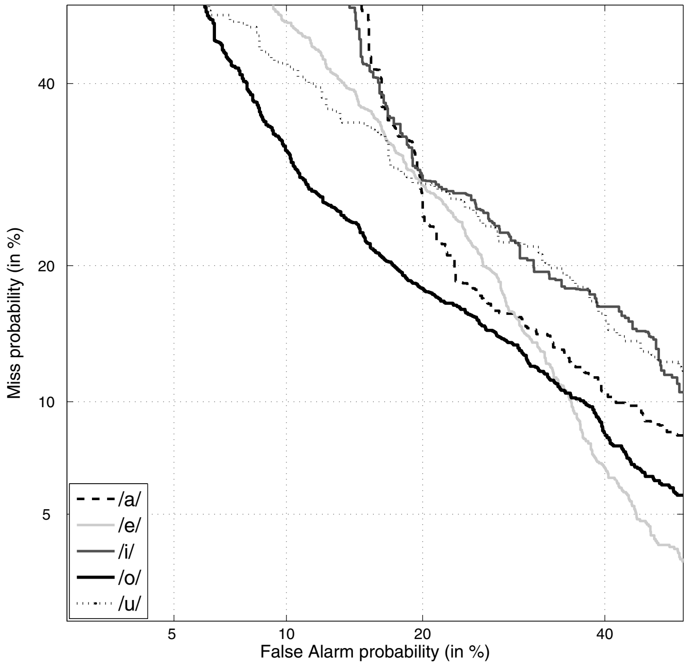

Fig. 2. DET curve for each vowel

5 Conclusions

The perceptual analysis of speech signals from people with Parkinson's disease is performed and the effectiveness, in terms of the discrimination capability, of each kind of perceptual coefficients in the problem of automatic classification of speech from PPD and HC is analyzed.

According to the results, PLP coefficients considered alone, offer good performance just for the case of the vowel /a/ and the MFCCs exhibit the higher performance in the vowels /i/ and /u/. However, for all experiments, the best performance is obtained with the vowel /o/ when MFCC, LPCC, RASTA-PLP-CEPS, PLP and RASTA-PLP-SPEC are merged and considered together.

The RASTA analysis has exposed good results in the task of automatic speech recognition; however, according to our findings, in order to achieve better results with vowels /e/ and /o/, those coefficients must be combined with other perceptual features such as MFCC and LPCC.

The main finding of this work indicates that for vowels /e/ and /o/ it is not enough to consider one kind of perceptual features, it is required to perform combination of different parameters to achieve good results in the task of automatic classification of speech signals from PPD and HC.

Acknowledgments. Juan Rafael Orozco Arroyave is under grants of "Convocatoria 528 para estudios de doctorado en Colombia, generación del bicentenario, 2011" financed by COLCIENCIAS. The authors give a special thanks to all of the patients and collaborators in the foundation "Fundalianza Parkinson-Colombia". Without their valuable support it would be impossible to address this research.

References

1. de Rijk, M.: Prevalence of parkinson's disease in europe: A collaborative study of population-based cohorts. Neurology 54, 21–23 (2000)
2. Sánchez, J., Buriticá, O., Pineda, D., Uribe, C., Palacio, L.: Prevalence of parkinson's disease and parkinsonism in a colombian population using the capture-recapture method. International Journal of Neuroscience 113, 175–182 (2004)
3. Skodda, S., Visser, W., Schlegel, U.: Vowel articulation in parkinson's diease. Journal of Voice 25(4), 467–472 (2011)
4. Hanson, D., Gerratt, B., Ward, P.: Cinegraphic observations of laryngeal function in parkinson's disease. Laryngoscope 94(3), 348–353 (1984)
5. Perez, K., Ramig, L., Smith, M., Dromery, C.: The parkinson larynx: tremor and videostroboscopic findings. Journal of Voice 10(4), 353–361 (1996)
6. Weismer, G., Jeng, Y., Laures, J., Kent, R., Kent, J.: Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniatrica et Logopaedica 53, 1–18 (2001)
7. Ramig, L., Fox, C., Shimon, S.: Speech treatment for parkinson's disease. Expert Review Neurotherapeutics 8(2), 297–309 (2008)
8. Bocklet, T., Nöth, E., Stemmer, G., Ruzickova, H., Rusz, J.: Detection of persons with parkinson's disease by acoustic, vocal and prosodic analysis. In: Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), pp. 478–483 (2011)
9. Little, M.A., McSharry, P., Hunter, E., Spielman, J., Ramig, L.: Suitability of dysphonia measurements for telemonitoring of parkinson's disease. IEEE Transactions on Bio-Medical Engineering 56(4), 1015–1022 (2009)
10. Tsanas, A., Little, M., McSharry, P., Ramig, L.: Accurate telemonitoring of parkinson's disease progression by noninvasive speech tests. IEEE Transactions on Biomedical Engineering 57(4), 884–893 (2010)
11. Falk, T., Chan, W., Shein, F.: Characterization of atypical vocal source excitation, temporal dynamics and prosody for objective measurement of dysarthric word intelligibility. Speech Communication 54(5), 622–631 (2012)
12. Buzo, A., Gray, A., Gray, R., Markel, J.: Speech coding based upon vector quantization. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 15–18 (1980)
13. Kim, H., Choi, S., Lee, H.: On approximating line spectral frequencies to lPC cepstral coefficients. IEEE Transactions on Speech and Audio Processing 8(2), 195–199 (2000)
14. Godino-Llorente, J., Gómez-Vilda, P., Blanco-Velasco, M.: Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters. IEEE Transactions on Biomedical Engineering 53(10), 1943–1953 (2006)
15. Hermansky, H.: Perceptual linear predictive (plp) analysis of speech. Journal of the Acoustical Society of America 87(4), 1738–1752 (1990)

16. Hermansky, H., Morgan, N.: Rasta processing of speech. IEEE Transactions on Speech and Audio Processing 2(4), 578–589 (1994)
17. Orozco-Arroyave, J., Vargas-Bonilla, J., Arias-Londoño, J., Murillo-Rendón, S., Castellanos-Domínguez, G., Garcés, J.: Nonlinear dynamics for hypernasality detection in spanish vowels and words. Cognitive Computation 4(2), 1–10 (2012)
18. Rusz, J., Cmejla, R., Ruzickova, H., Ruzicka, E.: 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 (2011)
19. Arias-Londoño, J., Godino-Llorente, J., Sáenz-Lechón, N., Osma-Ruiz, V., Castellanos-Domínguez, G.: An improved method for voice pathology detection by means of a hmm-based feature space transformation. Pattern Recognition 42, 3100–3112 (2010)
20. Scholköpf, B., Smola, A.: Learning with Kernel. The MIT Press (2002)
21. Sáenz-Lechón, N., Godino-Llorente, J., Osma-Ruiz, V., Gómez-Vilda, P.: Methodological issues in the development of automatic systems for voice pathology detection. Biomedical Signal Processing and Control 1, 120–128 (2006)