Czech Technical University in Prague

Faculty of Electrical Engineering

Doctoral Thesis

January 2016

Michal Novotný

Czech Technical University in Prague

Faculty of Electrical Engineering
Department of Circuit Theory

Automated assessment of diadochokinesis and resonance in dysarthrias associated with basal ganglia dysfunction

Doctoral Thesis

Michal Novotný

Prague, (January 2016)

Ph.D. Programme: Electrical Engineering and Information Technology

Branch of study: Electrical Engineering Theory

Supervisor: doc. Ing. Roman Čmejla, Ph.D.
Supervisor-Specialist: Ing. Jan Rusz, Ph.D.

The studies included in this thesis have been financially supported by the Czech Grant Agency under grant No. 102 /12/ 2230, Charles University in Prague under grant No. PRVOUK-P26/LF1/4, Czech Ministry of Health under grant No. NT14181-3/2013 and MZ ČR 15-28038A and Czech Technical University in Prague under grant No. SGS12/185/OHK4/3T/13.

DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

Signed:_____

Date:_____

ABSTRACT

The speech disruption caused by neurodegenerative disease is termed dysarthria. Previous research has identified dysarthria as one of the earliest and the most common subclinical manifestation associated with neurodegenerative diseases such as Parkinson's disease (PD), Huntington's disease (HD), multiple system atrophy (MSA) and progressive supranuclear palsy (PSP). Even though presence of articulatory and resonatory deficits represent an important part of dysarthria manifestations the most widely used methods currently available for the automatic evaluation of speech performance are focused only on the assessment of dysphonia. Therefore the aim of the thesis is to present the automatic assessment of articulatory and resonatory deficits as a useful tool enhancing current approaches of dysarthria analysis. Furthermore, the next goal is to use the objective dysarthria assessment to describe different dysarthria patterns. For this reason, we have examined articulatory deficits from two perspectives: first as the description of fast repetitive moves produced during speech diadochokinetic task (repetition of /pa/, /ta/ and /ka/ syllables) and second one as the description of the velopharyngeal function using a sustained phonation of the vowel /i/. The thesis summarizes the results of several studies and therefore of different participant datasets, including speakers diagnosed with PD, MSA, PSP, HD and healthy control (HC) speakers, which were used for the purposes of articulatory and resonatory deficits estimation.

In the first part aimed at rapid articulatory movements, the fully automatic assessment of several acoustic features describing six dysarthria dimensions (i.e. voice quality, laryngeal and supralaryngeal coordination, precision of consonant articulation, tongue movements, occlusion weakening and speech timing) was proposed. Results presented in this section showed significant differences between healthy and pathological utterances and moreover revealed distinctive patterns of articulatory distortion connected with different neuropathology. The presence of these distinctive patterns has been supported by the results of the performed classification experiment, which was able to distinguish between PD vs. HC with 87% accuracy.

The second part used the 1/3-octave spectrum method to describe the presence of nasal resonance as an acoustic feature reflecting increased nasality. Results of this section showed an overall increased presence of hypernasality in Parkinsonian and Huntingtonian patients. Nevertheless, significant differences were found only among the HD speakers. More importantly the intermittent character of hypernasality was found to be one of the very distinctive features present in 78% of HD speakers and moreover a significant increase in the presence of intermittent hypernasality was also documented in 68 % of MSA speakers.

In general, the results presented in the thesis show that the automatic assessment of articulatory deficits may provide useful clues for the early diagnosis, monitoring of disease progression, monitoring of speech therapy efficacy and adjustment and the differential diagnosis.

ABSTRAKT

Poškození řeči spojené s neurodegenerativními onemocněními se nazývá dysartrie. Dřívější výzkum ukázal, že dysartrie je jedním z nejdříve se vyskytujících a nejčastějším subklinickým projevem neurodegenerativních onemocnění jako jsou Parkinsonova nemoc (PD), Huntingtonova nemoc (HD), mnohočetná systémová atrofie (MSA) a progresivní supranukleární palasa (PSP). I přestože artikulační a rezonanční obtíže patří mezi významné dysartrické projevy, nejčastěji využívané metody automatického hodnocení jsou založeny pouze na měření dysfonie. Proto je hlavním cílem této práce návrh automatického hodnocení artikulačních a rezonančních deficitů jako užitečného nástroje doplňujícího metody používané v současnosti. Navíc je dalším cílem využití navržených objektivních metod pro hodnocení dysartrie k popsání různých dysartrických profilů. Z toho důvodu byly dysartrie zkoumány ze dvou perspektiv první bylo vyšetření rychlých artikulačních pohybů založené na úloze rychlého opakování slabik /pa/,/ta/ a /ka/ a tou druhou bylo hodnocení funkce měkkého patra založené na prodloužené fonaci samohlásky /i/. Dizertační práce shrnuje výsledky několika studií, a proto také pracuje s více různými soubory dat získaných od účastníků diagnostikovaných s PD, MSA, PSP, HD a navíc s kontrolními nahrávkami zdravých dobrovolníků.

V první části zaměřené na rychle artikulační pohyby je představena metodika automatického hodnocení šesti řečových dimenzí zahrnujících kvalita hlasu, koordinaci laryngeálních a supralaryngeálních artikulátorů, přesnost artikulace konsonant pohyby jazyka, slábnutí okluze, časování řeči. Výsledky prezentované v této části ukazují signifikantní rozdíly mezi zdravými a patologickými promluvami a navíc odhalili rozdílné charaktery řečových poruch u rozdílných typů neurodegenerativních onemocnění. Prezentovaný klasifikační experiment ukázal, že je možné rozčlenit de novo PD pacienty od zdravých účastníků s 88% úspěšností.

V druhá část, za použití analýzy 1/3-oktávového spektra, popisuje přítomnost nasální rezonance jako akustického ukazatele reflektujícího hypernasalitu. Výsledky prezentované v této části ukazují zvýšení výskytu hypernasality u PD i HD pacientů. Nicméně pouze u pacientů diagnostikovaných s HD byly nalezeny statisticky signifikantní rozdíly. Navíc se u HD a MSA se ukázala jako významná přítomnost kolísavé hypernasality kterou jsme dokumentovali u 78 % případů HD a u 68 % případů MSA Celkově, výsledky prezentované v této dizertační práci ukazují, že automatické hodnocení artikulační a rezonančních poruch může poskytnout užitečná vodítka pro potřeby brzké a diferenciální diagnózy, monitorování progrese onemocnění, monitorování efektivity léčby a monitorování účinků řečové terapie.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my thesis supervisor, Roman Čmejla and my supervisor-specialist Jan Rusz for providing me with an interesting research topic and moreover for their guidance through the fields of signal processing and speech analysis. Without their steadfast support this thesis would not have been completed. I would also like to Evžen Růžička for his valuable advices from the field of neurology.

As this work is result of team work, I am gratefully indebted to Tereza Tykalová for her methodical approach to the data acquisition, to Hana Růžičková for her valuable help in the field of perceptual analysis of speech, to Vlastimila Čmejlová, Vlastimila Stará, Veronika Stará, Jan Hlavnička and Tomáš Lustyk for their practical help with perceptual analysis. I would also like to thank to rest of the Signal Analysis, Modelling and Interpretation group for the beneficial discussions.

Thank you to Kayla Friedman and Malcolm Morgan of the Centre for Sustainable Development, University of Cambridge, UK for producing the Microsoft Word thesis template used to produce this document.

I am grateful to my family for their material and emotional support and for giving me a home, where I always gladly return. And finally, I owe a lot to my wife Hana for her endless support, persistence and for her belief in me.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

1FT, First formant trend
2FT, Second formant trend
AMC, Amplitude modulation component
AMR, Alternating motion rate
ANOVA, Analysis of variance
APS, Atypical parkinsonian syndromes
BSCD, Bayesian step changepoint detector
CN, Caudate nucleus
CSM, Consonant spectral moment
CST, Consonant spectral trend
CV, Consonant vowel
DDK, Diadochokinesis
EFn mean, Mean energy of nasal resonance peak
EFn SD, Standard deviation of nasal resonance peak energy
EFn trend, Tilt of nasal resonance peak energy through utterance
ES, Cohen's effect size
F0, Fundamental frequency
F1, First formant frequency
F2, Second formant frequency
F3, Third formant frequency
fs, Sampling frequency
GPe, Globus palidus externa
GPi, Globus palidus interna
GUI, Graphical user interface
H&Y, Hoehn & Yahr
HC, Healthy control
HD, Huntington's disease
ICC, Intra class correlation
L-dopa, Levodopa
LOSO, Leave one sample out
LP, Linear prediction
MSA, Multiple system atrophy
MSA-P, Parkinsonian MSA phenotype
MSA-C, Cerebellar MSA phenotype
NNIPPS, Natural History and Neuroprotection in Parkinson Plus Syndromes
p(f), Sound pressure
PD, Parkinson's disease
PD_U, Parkinson's disease – untreated

PD_{SDD}, Parkinson's disease – short disease duration
PSP, Progressive supranuclear palsy
PSP-CBS, Corticobasal PSP phenotype
PSP-P, Parkinsonian PSP phenotype
PSP-PAGF, Pure akinesia with gait freezing PSP phenotype
PSP-PNFA, Progressive nonflueant aphasia PSP phenotype
PSP-RS, Richardson's syndrome PSP phenotype
R(f), Radiation characteristic
RBF, Radial basis function
RMS, Root mean square
S(f), Source characteristic
SD, Standard deviation
SMA, Supplementary motor area
SN, Subthalamic nucleus
SNc, Substantia nigra pars compacta
SNr, Substantia nigra pars reticulata
SNR, Signal to noise ratio
STN, Substantia nigra
SVM, Support vector machine
T(f), Transfer function
TEO, Teager energy operator
UPDRS, Unified Parkinson's Disease Rating Scale
UHDRS, Unified Huntington's Disease Rating Scale
VOT, Voice onset time
VSQ, Vowel similarity ratio
VSQ_{30}, Vowel similarity ratio of first 30 ms from the vowel
VVQ, Vovel variability quotient

FOREWORD

This thesis presented to obtain the Ph. D. degree in Electrical Engineering Theory by the Czech Technical University of Prague; the Faculty of Electrical Engineering is result of four studies performed at the Department of Circuit Theory. The two articles have been published in international journals, one is currently in review to the international journal and one is the study presented during 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2015.

ARTICLES:

"Automatic Evaluation of Articulatory Disorders in Parkinson's Disease."
M. Novotný, J. Rusz, R. Čmejla, and E. Růžička.
EEE/ACM Transactions on Audio, Speech, and Language Processing (2014), 22(9): 1366 – 1378.
Impact Factor (2014): 2.45
Quartile in Category: 1.
Journal Rank in Category: 4/31 (Acoustics) 35/200(Engineering, Electrical & Electronic)

"Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy."

J. Rusz, C. Bonnet, J. Klempíř, T. Tykalová, E. Baborová, M. Novotný, A. Rulseh, and E. Růžička.
Journal of Neurology (2015), 262: 992 - 1001
Impact Factor (2014): 3.37
Quartile in Category: 1.
Journal Rank in Category: 52/192 (Clinical Neurology)

"Hypernasality in Parkinson's disease and Huntington's disease."

M. Novotný, J. Rusz, R. Čmejla, H. Růžičková, J. Klempíř, E. Růžička.
Currently submitted

CONFERENCE PROCEEDINGS:

"Automatic detection of voice onset time in dysarthric speech."
M. Novotný, J. Pospíšil, R. Čmejla, J. Rusz.
ICASSP 2015
SCImago Journal Rank (2014): 0.465

Furthermore, several minor studies related to the topic of the thesis have also been published. These are listed at the end of the thesis.

1 INTRODUCTION

Neurodegenerative diseases rank among the medical disruptions with the most salient impact on the quality of a patient's life. In particular, the neurodegenerative diseases connected with motor skill decline have a profound effect. Among these, the basal ganglia disruptions result in a multitude of motor skill deficits. Furthermore, due to the different pathomechanism basal ganglia disruptions results in several very different diseases (e.g Parkinson's disease (PD), Huntington's disease (HD), Multiple System Atrophy (MSA) and Progressive Supranuclear Palsy (PSP)). In general, these conditions can be divided into the akinetic rigid syndromes (including PD, MSA and PSP) and the hyperkinetic movement disorders (including HD). Diagnosis of these diseases is usually based on the onset of primary motor manifestations. For instance, Parkinson's disease as the most common basal ganglia disorder is diagnosed on the basis of the presence of tremor at rest, bradykinesia, rigidity, postural and gait instability. However, previous research has shown that by the time of diagnosis, extensive damage has been already caused. In the case of PD, the damage corresponds to a decrease of 70% dopaminergic neurons and the 80% dopamine detriment.

Indeed, the prodromal manifestations exist and it has been shown that they can be present several years prior to the diagnosis. Among the most common and the earliest manifestations is dysarthria. Dysarthria, as a speech distortion resulting from distorted neural control of speech, occurs in 90 % of all PD cases. Reason for this prevalence is caused by the complexity of human speech. As a consequence, the speech assessment may be a useful clue for early diagnosis. In addition, different neurodegenerative diseases result in different types of dysarthria which match the

affected brain area. For instance, HD results in hyperkinetic dysarthria, whereas PD results in hypokinetic dysarthria and PSP and MSA manifest mixed dysarthria including features of hypokinetic dysarthria, ataxic dysarthria and spastic dysarthria. For this reason, the speech analysis may also provide useful indices for the differential diagnosis of various neurodegenerations.

Currently, even though, dysarthria represents primarily a disruption of articulation, the majority of recently published studies have been based on phonatory features and sustained phonation task. Therefore, the goal of this study is to provide an automatic assessment of the articulatory and resonatory aspects of speech using speech diadochokinetic (DDK) task and sustained phonation of vowel /i/.

2 STATE OF THE ART

2.1 Basal Ganglia

2.1.1 Anatomy

The basal ganglia system consists of four main nuclei incorporating the striatum, the globus palidus, the subthalamic nucleus (SN) and the substantia nigra (STN). The striatum has a curved shape and consists of the caudate nucleus (CN) and the putamen. The putamen is connected with the globus palidus to form the lentricular nucleus. The globus palidus may be further divided into external (GPe) and internal (GPi) sections. The substantia nigra, which is divided into two parts i.e. the substantia nigra pars compacta (SNc) and the substantia nigra pars reticulata (SNr), is located together with the subthalamic nucleus below the thalamus and the lentricular nucleus Figure 2.1. The system of basal ganglia, however, is not segregated and other structures, e.g. the thalamus and the pendulopontine nucleus, also have a profound effect on basal ganglia function.

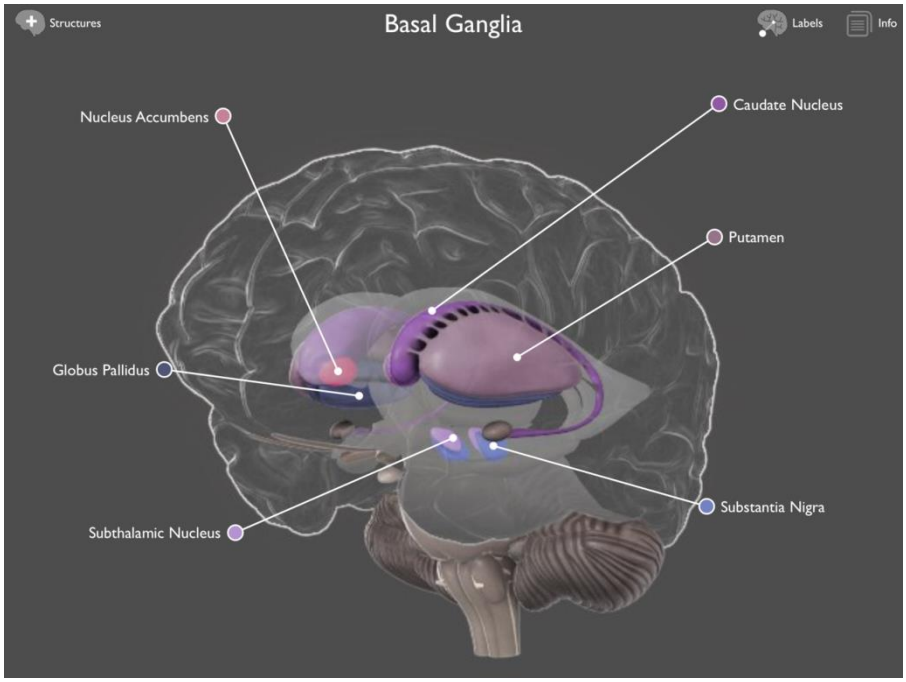

Figure 2.1: Anatomical organization of the basal ganglia depicting the position and shape of the single basal ganglia nuclei (including the striatum formed by the caudate nucleus and the putamen, the lentricular nucleus formed by the putamen and the globus palidus, the substantia nigra and the subthalamic nucleus). Adapted from (Yates, 2013)

2.1.2 Dopaminergic Control

The effort to explain both physiological and pathological processes in basal ganglia during PD led to the creation of a functional model proposed by Albin et al. (1989). Based on the observed striatal heterogeneity, the model presented competitive direct and indirect excitatory pathways. The direct pathway is characterized by the presence of neurons containing the substance P and projects from the striatum to the GPi and the SNr; the indirect pathway is characterized by the presence of enkhephalins and projects from the striatum towards the GPe, see Fig. 2. (Lemoine and Bloch, 1995).

This model has explained the normal functioning of the basal ganglia, describing the excitatory effect of direct pathway and the inhibitory effect of the indirect pathway caused by the STN an excitation modulating the thalamocortical projection (Tripoliti, 2010). Furthermore, this model provides explanation of movement disruptions in the Parkinsonian state. According to this model, the dopamine deficiency caused by dopaminergic cell loss in the SNc leads to disinhibition of the GPi and the SNr and therefore to the increased inhibition of the thalamocortical projection (Yelnik, 2002).

Nevertheless, several concerns have been raised against this model. The study by Yelnik (2008) summarizes four main issues. First the model supposes two different types of dopamine receptors D1 (excited by dopamine) and D2 (inhibited by dopamine) located on different neuron populations. Even though the presumption has been supported by the study Gerfen et al. (1990), the study Aizman et al. (2000) observed that many striatal neurons co-localize both D1 and D2 receptors. Second the direct/indirect pathway model assumes that striatal neurons project either to the GPe or the GPi, yet the study Parent and Hazrati (1995) showed neurons projecting to the GPe, the GPi and the SNr. Third, according to the model, the GPe should be hypoactive, yet this finding has not been supported (Vila et al., 1997). Finally, the model does not consider the effect of other pathways which have a crucial effect on basal ganglia function (Smith et al., 1994, Francois et al., 2000, Jan et al., 2000, Orieux et al., 2000).

2.1.3 Connectivity and Functional Organization

The understanding of how the basal ganglia interact with other parts of the brain or even how single nuclei interact with each other is limited. For this reason, several models have been developed. In general, all models agree that the main information input arises from the cereberal cortex; however there are different models of the information projection.

The direct/indirect pathway model presented by Albin et al. (1989) originally consisted of two segregated pathways, a direct pathway with excitatory effect on movement and an indirect pathway with inhibitory effect on the movement. The direct pathway traverses the striatum, continuing past the GPi and then the SNr. The indirect pathway passes from the striatum across the GPe, then the STN and from there to the GPi and the SNr. In the mid 90’s, the study Mink (1996) extended this model to include a hyperdirect pathway, see Figure 2.2.

Five segregated circuits were proposed in the model presented by DeLong and Wichmann (2007), these being the oculomotor, motor, dorsolateral, prefrontal, lateral orbitofrontal and anterior cingulated circuits. These five circuits originate in the frontal cortex and pass over the basal ganglia by both direct and indirect pathways.

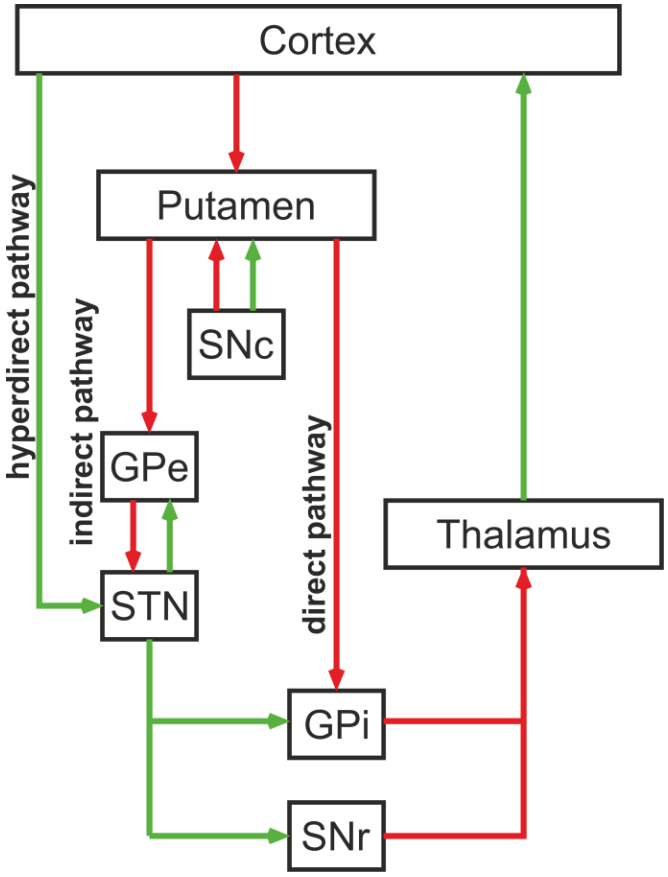

Figure 2.2: The box-arrow model of the basal ganglia consisting of the direct pathway (the putamen and the GPi), the indirect pathway (the putamen, GPe, the STN, the GPi, the SNr) published in Albin et al. (1989) and the hyperdirect pathway connecting the cortex and the STN (Mink, 1996)

The third suggested model was the model of three functional territories presented in Parent (1990). This model posits the sensorimotor, associative and limbic basal ganglia territories. The sensorimotor territory processes motor and somesthetic information, the associative territory covers cognitive information and the limbic territory process emotional and motivational information. This model has been validated by a study based on primate models (Francois et al., 2004) and also by research performed on human participants (Mallet et al., 2007; Fig.3).

In addition to the models with segregated circuits, the study by Yelnik (2008) suggests some integrative properties of the basal ganglia. This model considers the significant decrease in volume of each successive nucleus, as well as the significant decrease in the number of neurons in each successive nucleus. Based on the strong convergence, several studies Yelnik (2008) and Flaherty and Graybiel (1993) suggest that the basal ganglia do not form independent circuits but instead provide ascending connections between the sensorimotor, associative and limbic striatal subdivisions.

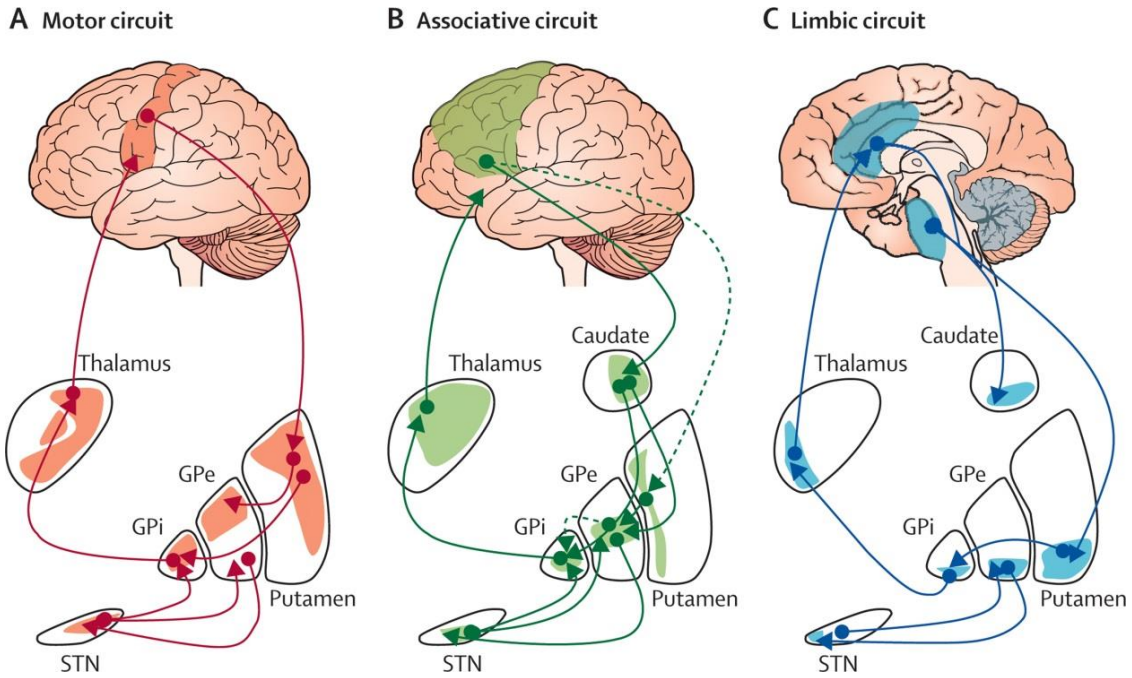

Figure 2.3: The basal ganglia model comprising the A) motor circuit, B) associative circuit and C) limbic circuit, which according to (Yelnik, 2008) are increasingly interconnected in each successive nucleus. Figure adapted from Rodriguez-Oroz et al. (2009)

2.2 Basal Ganglia Disorders

According to the study Hallett (1993), the basal ganglia diseases can be divided into the hyperkinetic movement disorders and the akinetic-rigid syndromes (Figure 2.4). The akinetic-rigid syndromes are characteristic by difficulty in generating voluntary movements and comprise Parkinson's disease and atypical parkinsonian syndromes. The hyperkinetic movement disorders are described as involuntary, undesired movements and comprise chorea, dyskinesia and dystonia.

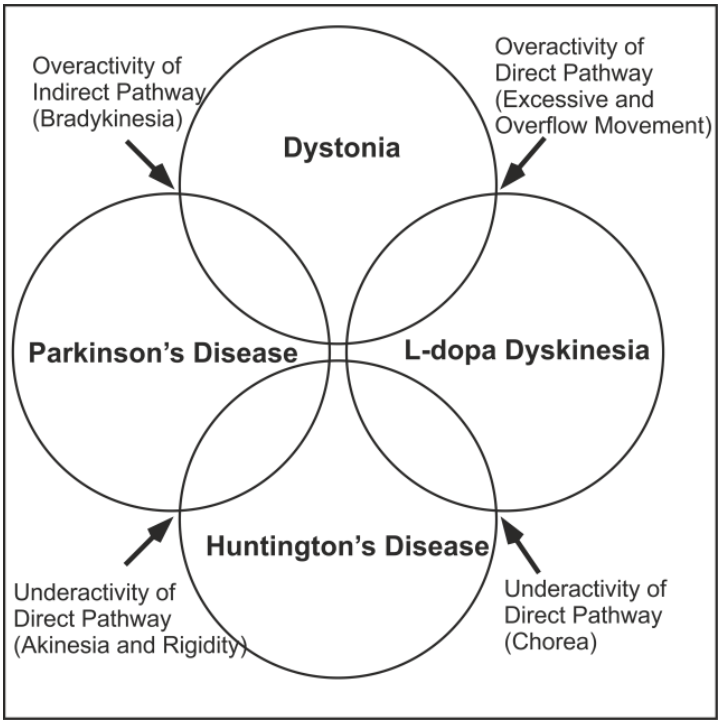

Figure 2.4: Venn diagram of the four main basal ganglia disorders, representing disease overlaps and separations. Figure adapted from Hallett (1993)

2.2.1 Akinetic-rigid Syndromes

The akinetic-rigid syndromes are characterized by reduction of movement and increased tone (rigidity). The reduction of movement can be further divided into the failure to move or to initiate a movement (akinesia) and slowness of movement (bradykinesia). The most common example of disease leading to akinetic-rigid syndrome is Parkinson's disease. Among other diseases resulting in akinetic-rigid syndrome are the atypical parkinsonian syndromes including Progressive Supranuclear Palsy and Multiple System Atrophy.

2.2.1.1 Parkinson's disease

Parkinson's disease is a progressive idiopathic disorder which primarily affects dopaminergic neurons in the SNc and causes dopaminergic striatal loss (Hornykiewicz, 2008). Low levels of dopamine leads to dysfunction of the basal ganglia and primarily account for motor deficits. The cardinal features of PD include tremor at rest, rigidity, bradykinesia, and postural instability. In addition to the most common motor manifestations, other non-motor manifestations such as autonomic dysfunction,

cognitive and neurobehavioral abnormalities, sleep alterations and sensory disruptions may be evident (Jankovic, 2008, Rodriguez-Oroz et al., 2009).

Being the second most common neurodegenerative disease PD usually affects people after age over 50 years and only 10% of all new cases are diagnosed in participants under 40 years of age (de Rijk et al., 2000, Hoehn, 1992). According to the study (de Rijk et al., 1997) PD affects 1.6% people over age of 65 years. Moreover the age is the most important factor, which in combination with population ageing, is assumed to result in the increase disease occurrence (Van Den Eeden et al., 2003).

The diagnosis of PD is based upon the presence of primary motor manifestations, which develop after 60-70% of dopaminergic neurons degenerate and dopamine levels are reduced by 80% (Bernheimer et al., 1973, Postuma et al., 2012). Due to the slow gradually progressive nature of the disorder, PD has a prodromal interval during which the main motor manifestations are not clearly evident. The duration of PD prodromal period has been documented as 3-15 years (Postuma et al., 2012).

In the current time there is no treatment which would heal disease or even stop its progression. Therefore the pharmacotherapy and neurosurgical interventions that are currently available only offer alleviation of certain parkinsonian manifestations. Despite the fact that medication generally prolongs active life expectancy, the effect of treatment depends upon the stage of the disease during which it is initiated. Furthermore, there is no treatment that can cure PD or halt its progression. Therefore the early diagnosis of PD plays a vital role in improving the patient's quality of life (Becker et al., 2002, Rodriguez-Oroz et al., 2009).

2.2.1.2 Atypical Parkinsonian Syndromes

In addition to the idiopathic form of Parkinson's disease, there are atypical parkinsonian syndromes (APS) which may exhibit similar manifestations. However, the disorder's progress is significantly faster, the APS respond poorly to pharmacological treatment by levodopa (L-dopa) and the disease manifestations exhibit other associated features (Quinn, 1995). Among others the two most common APS are progressive supranuclear palsy and multiple-system atrophy (Federico et al., 1997).

Progressive Supranuclear Palsy

Progressive supranuclear palsy is a neurodegenerative disease of middle and late age. The incidence of disease is estimated as 5.3 new cases per 100 000 in population with age over 50 years, and the average life expectancy of 5.3 years after disease onset

(Bower et al., 1997). Nevertheless, Hughes et al. (1992) performed necropsies of 100 patients dying with diagnosis of idiopathic PD documented PSP in 6% of them, which therefore suggests that the incidence reported by Bower et al. (1997) may be underestimated. Furthermore, the study (Schrag et al., 1999) scanned 121 608 people with records mentioning "Parkinson's disease", "parkinsonism", "Shy-Drager syndrome", "parkinsonism with orthostatic hypotension" and "other extrapyramidal disorders, not otherwise specified", along with those ever receiving antiparkinsonian medication, and revealed that the PSP prevalence rate among these patients was 4.9 per 100 000.

According to the most prominent manifestations, four PSP phenotypes are defined. The most common is Richardson's syndrome (PSP-RS), while the second most common is PSP-parkinsonism (PSP-P) characterized by bradykinesia, rigidity and sometimes tremor. The third phenotype is known as pure akinesia, with gait freezing (PSP-PAGF), the fourth is corticobasal syndrome (PSP-CBS) and finally the progressive nonfluent aphasia (PNFA; Boeve, 2012, Williams and Lees, 2009).

PSP diagnosis is based on four mandatory inclusion criteria. First, as the age of main manifestation onset is usually between 50 and 70 years (Schrag et al., 1999), the onset age is at least 40 years. Second, the disease has the character of a gradually progressive disorder. Third, the patient manifests vertical supranuclear gaze abnormalities. The supranuclear gaze may be preceded by slowing of saccades and hypometric saccades and may be followed by horizontal gaze abnormalities (Rehman, 2000). Supranuclear vertical gaze may occasionally be absent (Daniel et al., 1995). And finally, there is prominent postural instability with tendency to falls in the first year after symptom onset (Rehman, 2000). Furthermore, one of the clues for the PSP diagnosis is that patients experience bradykinesia and rigidity, which in PSP are usually nonresponsive to L-dopa.

No effective treatment for PSP is currently available. In contrast to PD, the dopaminergic replacement therapy is only transiently or mildly effective in relieving some of the manifestations, even though it may sometimes have beneficial effect (Rehman, 2000).

Multiple System Atrophy

Multiple system atrophy usually starts in the sixth decade of life. After PSP, MSA is the second most common APS with an incidence of 3.0 new cases per 100 000

in population with age over 50 years, the average life expectancy of 8.5 years from the symptom onset (Bower et al., 1997) and the prevalence between 1.9 and 4.5 cases per 100 000 (Tison et al., 2000, Vanacore et al., 2001). However, similarly as in PSP Schrag et al. (1999) reported MSA prevalence of 4.4 per 100 000 among participants who already received other diagnosis and thus suggests that, due to the misdiagnosed patients, the MSA prevalence figures are underestimated.

Among the main MSA features belong autonomic failure, Parkinsonism, cereberal ataxia and pyramidal signs in any combination (Wenning et al., 2004). According to the prominent motor presentations, two major MSA phenotypes can be clinically distinguished (Wenning et al., 2004). The more common parkinsonian (MSA-P) phenotype comprises 80% of all cases, whereas cereberal ataxia (MSA-C) predominates in 20% MSA patients (Wenning et al., 1994). The MSA-P is connected with progressive akinesia and rigidity, jerky postural tremor, and orofacial or craniocervical dystonia. It is difficult to distinguish early stages of MSA-P from PD, however MSA-P is connected with significantly more rapid deterioration. Furthermore, only 30% of MSA-P patients react to L-dopa treatment in short term and only 10% of MSA-P cases are responsive for the long term treatment (Boesch et al., 2002). The MSA-C phenotype is described by presence of gait ataxia, limb kinetic ataxia, scanning dysarthria, and cerebellar oculomotor disturbances (Wenning et al., 2004). The MSA-C in the later stages usually develops additional non-cerebellar manifestations, but in the early stages the MSA-C may be hardly distinguishable from idiopathic late onset cerebellar ataxia (Wenning et al., 2004).

At the current time, there is no treatment which would at least stop the disease progression, and as a result only the symptomatic treatment is administered. Nevertheless, even the symptomatic treatment is limited due to the absence of any effective treatment for the cerebellar manifestations of MSA and failure of parkinsonian features to show more than limited responsiveness to L-dopa. Therefore, treatment of autonomic dysfunctions - namely orthostatic hypotension, urinary symptoms, male impotence and inspiratory stridor - is crucial for the increase of the patient's life quality at present time (Wenning et al., 2004, Jankovic et al., 1993, Bonnet et al., 1997, Hussain et al., 2001).

2.2.2 Hyperkinetic Movement Disorders

The hyperkinetic movement disorders including chorea, dyskinesia and dystonia are characterized by the presence of involuntary undesired movements (Hallett, 1993).

One of the well-known hyperkinetic movement disorders comprising chorea and dystonia is *Huntington's disease* (Walker, 2007).

2.2.2.1 Huntington's disease

Huntington's disease is an autosomal-dominant neurodegenerative disorder with prominent atrophy in the caudate nucleus and the putamen (Vonsattel and DiFiglia, 1998). The neurons containing enkephalin and those projecting to the GPe are more directly involved (Gutekunst et al., 2002) which shows preferential involvement of the indirect basal ganglia pathway (Paulsen et al., 2005). Distortion of the indirect pathway leads to the chorea, dystonia and incoordination. Furthermore, HD is associated with cognitive decline and behavioral difficulties (Walker, 2007).

As an autosomal-dominant disorder, Huntington's disease is caused by the presence of a mutant protein, Huntingtin, which is expanded by an excessive number of CAG triplet repetitions. The present literature states, that the presence of 35 or fewer CAG triplets does not cause HD, yet reaching the number of 36 CAG repeats the incomplete penetrance may happen, and exceeding 40 triplets, the disease becomes fully penetrant (Bates and Benn, 2002, Rubinsztein et al., 1996). Indeed, previous studies have shown that number of CAG triplets does have a negative correlation with the age of symptom onset (Wexler et al., 2004, MacDonald et al., 1999). The mean age of HD becoming symptomatic is 40 years, though this onset is not limited to mid or late age and may be also present in juvenile form, consequetly HD may become symptomatic between age of 1 and 80 (Walker, 2007). The life expectancy after HD symptoms onset is 15 – 20 years (Ross and Tabrizi, 2011).

The prevalence of HD ranges between 5-7 cases per 100 000 people in white populations, whereas Asian and African populations show reduced prevalence of around 0.5 cases per 100 000 (Hayden, 1981, Folstein, 1989, Harper and Jones, 2002). However, places where intermarriage with whites is more common show increased disease frequency (Hayden, 1981, Folstein, 1989, Harper and Jones, 2002).

Although currently only symptomatic treatment is available, several agents have shown promising results in animal HD models. Among others, the coenzyme Q10 and the substance creatine show the most interesting outcomes, and the two phase III clinical trials focus on the administration of these two substances in HD patients (Ross and Tabrizi, 2011). However, with these studies there arises a need for the biomarker to

reflect treatment efficacy and disease progression, because currently used perceptual assessments are subjective and do not display sufficient sensitivity (Walker, 2007).

2.3 Speech

Speech is possibly the most complex human activity, involving the coordination of up to one hundred muscles and several brain centers (Simonyan and Horwitz, 2011). Furthermore, healthy people using this precise muscle coordination are usually able to produce speech at a high rate of five to six syllables per second (Levelt, 1989). Production of speech may be distributed between three anatomical structures, those being the respiratory, the laryngeal and the supralaryngeal systems. The respiratory system consists of the lungs and provides airflow as a source of energy for the speech. The laryngeal system includes vocal folds and controls alterations between voiced unvoiced sounds and intonation. The supralaryngeal system comprises the vocal tract, the velum, the tongue, the jaw and the lips and is responsible for articulation. As for articulation itself, according to Tripoliti (2010) it may be further divided into two articulatory tasks. The first part is to interrupt airflow and therefore produce unvoiced sounds, and the second is to set the parameters of vocal tract which modulates voiced sounds (Stevens, 2000). In contrast to the anatomical division, speech may be also divided into speech production subsystems including respiration, phonation, articulation and prosody. For these terms, we can define respiration as the mechanical exchange of gas via the lungs, phonation as vocalization through vocal fold vibration, articulation characterized as the adjustment of the vocal tract and sound modulation, and prosody as tone, intonation, rhythm and stress in utterance.

2.3.1 Speech Motor Control

Studies by Brown et al. (2009) and Price (2010) used imaging techniques during speech to identify the brain areas responsible for the speech production. These studies illustrated the involvement of the motor and premotor cortex, the cerebellum, the supplementary motor area (SMA) the superior temporal gyri, the tempoparietal cortices and the anterior insula with left lateralized activation in the putamen.

It has been shown that the anterior insula is activated with speech being executed or only imagined, and therefore it is assumed that the anterior insula is responsible for speech planning (Riecker et al., 2002, Watkins et al., 2008). The SMA area is also included in motor planning and motor sequencing (Indefrey and Levelt,

2004), though its precise role is not well understood (Price, 2010). During movement initiation and execution the bilateral premotor and motor cortex, the pre-SMA, and the left putamen are activated (Brown et al., 2009). Furthermore, Brown et al. (2009) was able to distinguish two areas responsible for larynx movements and for tongue and lip movements. The anterior cingulated and bilateral head of the caudate was found to be important for the orofacial speech movements (Chang et al., 2009). Further, the involvement of additional brain areas has been documented during increased complexity of produced speech (Soros et al., 2006). This study, using single /pa/, /ta/, /ka/ syllables and the /pa/-/ta/-/ka/ syllable train and the clustered MRI, detected that monosyllabic utterances lead to almost identical activation in pyramidal and extrapyramidal structures, whereas multisyllabic utterances cause activation in the Broca's area, the left cerebellum, the left caudate nucleus and the bilateral superior and middle temporal gyri (Figure 2.5).

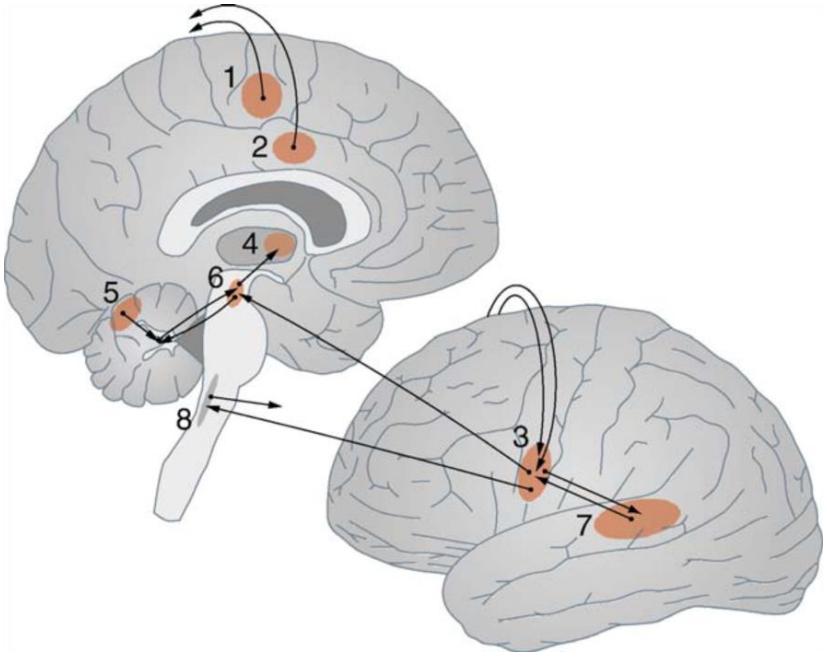

Figure 2.5: The neural network of speech production. Areas activated during speaking are shown in red. Schematic fiber tracts connecting those areas are represented by black arrows. Only main areas of activation and main fiber tracts are shown. The supplementary motor area (1) and the cingulate motor areas (2) are connected with the primary motor cortex (3). Several connections exist between the cortical and the subcortical motor system. Subcortical activation was found in the thalamus (4), the basal ganglia (not shown), the red nucleus (6) and in the vermal and paravermal cerebellum (5). In addition, the bilateral posterior superior temporal gyrus (7) was

activated. The brain stem nuclei innervating the articulatory organs, such as the nucleus hypoglossus, lay outside the field of view (8). Adopted from study Soros et al. (2006).

2.3.1.1 Basal Ganglia Speech Motor Control

The effect of different speech rate on the place of activation has been documented by Wildgruber et al. (2001), who used three rates of silent /ta/ syllable repetitions. This study has described the activation of the left putamen during lower rates and the activation of the cerebellum during higher rates. This finding in accordance with clinical observations of accelerated speech tempo in patients with basal ganglia disruptions (Ackermann et al., 1997), and that maximum speech rate does not drop below 3 Hz in patients with cerebellum-damage-induced speech impairment. Furthermore, the asymmetry of basal ganglia activation is in concordance with the observed articulatory impairment and reduced voice volume after left-sided subcortical infarction (Alexander et al., 1987).

Together with the different area activation, two different time patterns were described (Riecker et al., 2005). The first achieved peak activation happened in three to five seconds after the onset of acoustical stimulation and included the SMA, the left dorsolateral pre-frontal cortex (including the Broca's area), the left anterior insula and the right superior cerebellum. This pattern is thought to be responsible for the motor preparation. Conversely, the second pattern, which was activated eight to nine seconds after onset of speech production, includes the left sensorimotor cortex, left thalamus, left putamen/pallidum, left caudate nucleus and right inferior cerebellum. This loop is thought to be responsible for the movement execution (Figure 2.6). However, due to the complexity of speech, the involvement of different brain areas is complicated. Examination of stuttering showed overactivity of the substantia nigra, subthalamic nucleus, pendulopontine nucleus and red nucleus (Watkins et al., 2008). Furthermore, Brown et al. (2009) detected no putamen activity during phonation and strong left hemispheric activity during connected speech, suggesting that the putamen is more involved in the articulation than in phonation. Nevertheless, damage of the basal ganglia among others leads to severe disruption of phonation.

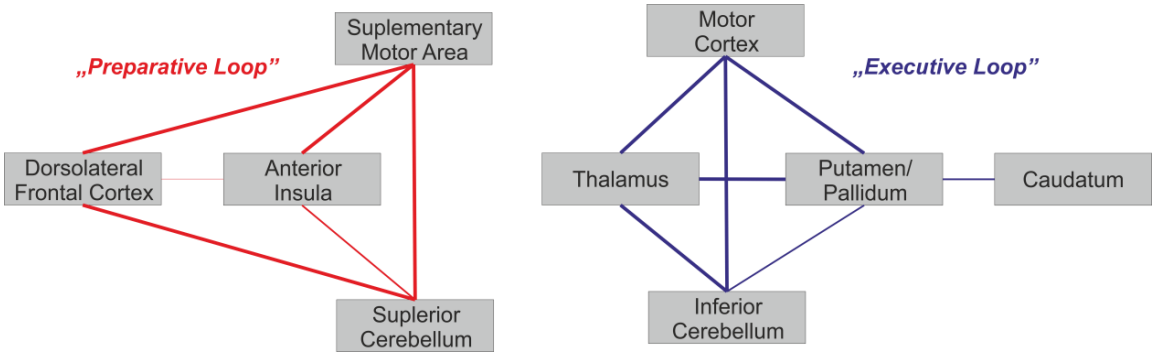

Figure 2.6: The preparative and executive loops obtained by quantitative functional connectivity analysis. The bold lines are connections with very high correlation (r higher than 0.90), whereas thin lines mean high correlation (r between 0.75 to 0.90). Low and intermediate correlations are not included. Adapted from Riecker et al. (2005)

2.3.2 Dysarthria

Dysarthria is commonly defined as a neurologic speech disorder. However, Duffy (2013) provides a narrower description of dysarthria as a neurological speech disorder resulting in alterations of the strength, speed, range, steadiness, tone, or accuracy of movements required for the proper respiration, phonation, articulation or prosody. Furthermore, Duffy (2013) explicitly states that a) dysarthria is of neurological origin, b) dysarthria is a disorder of movement and c) dysarthria can be divided into different subtypes according to its perceptual characteristics and, presumably, according to the underlying neuropathology.

Among neurodegenerative diseases, dysarthria is a common manifestation. Examination of two hundred PD participants revealed that dysarthria is present in up to 90% (Logemann et al., 1978). Furthermore, the high sensitivity of speech implies that dysarthria may be one of the earliest prodromal manifestations of neurodegenerative disease. For instance Postuma et al. (2012) documented the presence of disrupted speech in PD patients 3-15 years prior to disease diagnosis. Six major dysarthria subtypes with their most common degenerative causes are listed in Table 2.1; furthermore Duffy (2013) defines the mixed dysarthria as a combination of previous six dysarthria subtypes. For the more detailed description of PD, PSP, MSA and HD related dysarthria subtypes see further sections concerning ataxic, spastic, hypokinetic and hyperkinetic dysarthrias.

Table 2.1: Neurodegenerative causes and common manifestations of different dysarthria subtypes

Flaccid Dysarthria	
Neuropathology	Manifestation
Amyotrophic lateral sclerosis; motor neuron disease; **multiple system atroph**y; spinomuscular atrophy, Kenedy's disease	Hypernasality; breathiness; diplophonia; nasal emission; audible inspiration; short phrases; rapid deterioration and short recovery with rest; speaking on inhalation
Spastic Dysarthria	
Neuropathology	Manifestation
Amyotrophic lateral sclerosis; motor neuron disease**; multiple system atrophy; progressive supranuclear palsy**; spinocerebelar atrophy; Friedrych's ataxia; primary lateral sclerosis	Harshness; low pitch; slow rate; strained-strangled quality; pitch breaks; slow and regular alternate motion rates;
Ataxic Dysarthria	
Neuropathology	Manifestation
Cerebellar degeneration; Spinocerebellar ataxia; **multiple system atrophy, progressive supranuclear palsy**; olivpontocerebelar atrophy	Excess and equal stress; irregular articulatory breakdowns; irregular alternate motion rates; distorted vowels; excess loudness variation; telescoping of syllables
Hypokinetic Dysarthria	
Neuropathology	Manifestation
Parkinson's disease; parkinsonism; **multiple system atrophy; progressive supranuclear palsy**; Lewy bodies disease; corticobasal degeneration; frontotemporal dementia; parkinsonism + amyotrophic lateral sclerosis	Monopitch; monoloudness; reduced stress; Inappropriate silences; short rushes of speech; variable rate; increased rate i segments; increased overall rate; rapid "blurred" alternate motion rates; repeated phonemes; palilalia

Hyperkinetic Dysarthria

Neuropathology	Manifestation
Huntington's disease	Irregular alternate motion rates; distorted vowels; prolonged intervals; sudden forced inspiration/expiration; voice arrests; transient breathiness; voice tremor; myoclonic vowel prolongation; intermittent hypernasality; slow and irregular alternate motion rates; marked deterioration with increased rate; inappropriate vocal noises; coprolalia; intermittent strained voice/arrests; intermittent breathy/aphonic segments

Unilateral Upper Motor Neuron Dysarthria

Neuropathology	Manifestation
Multiple sclerosis	Poorly sequenced sequential motion rates; articulatory groping; distorted substitutions attempts at self-correction; articulatory complications; automatic > volitional speech; inconsistent articulatory errors; increased errors with increased length

2.3.2.1 Ataxic Dysarthria

Ataxic dysarthria is related to the damage of cerebellar control circuit and manifests in reduced muscle tone and incoordination, respiration, phonation, articulation and prosody. Among others ataxic dysarthria features are present in PSP and MSA patients.

Examination of respiration and laryngeal function found that patients with ataxic dysarthria exhibits reduced vital capacity and respiratory and phonatory incoordination (Abbs et al., 1983, Murdoch et al., 1991, Mcclean et al., 1987). Examination of speech have recorded incoordination between exhalation and phonation, paradoxical or abrupt changes in the rib cage movements, irregular chest wall movements during prolonged phonation and tendency to initiate utterance at lower lung levels (Abbs et al., 1983, Murry, 1983). Studies concerned with the phonatory aspects have described abnormal variability of F0 and intensity, increased jitter, shimmer, pitch level and abnormal voice onset time (VOT; Gentil, 1990, Kent et al., 2003, Kent et al., 1997). Duffy (2013) summarizes these findings as phonatory instability or phonatory-respiratory instability, conditions that are secondary to problems of muscle coordination, timing, control, or tremor.

Regarding articulatory deficits, the literature refers to the presence of slow rate comprising increased sentence and word duration, increased duration of formant transitions, slow alternate motion rates excessive pauses slow supralaryngeal articulation movements and difficulties to initiate movement (Ackermann et al., 1995b, Ackermann and Hertrich, 1997, Ziegler and Wessel, 1996, Kent et al., 2000a, Kent et al., 1997).

2.3.2.2 Spastic Dysarthria

Spastic dysarthria is related to the bilateral damage of direct and indirect pathways. The main features of spastic dysarthria are the combination of spasticity and weakness. Among other neural disruptions, ataxic dysarthria features are present in PSP and MSA.

The effect of spastic dysarthria on respiration is not well understood. However the possible effect of reduced voluntary activation of respiratory muscles has been presented (Gouveia et al., 2006). Moreover, Duffy (2013) suggests that spastic dysarthria may have a similar effect to spastic cerebellar palsy, exhibiting respiratory intake restriction and reduced vital capacity (Aronson, 1990).

Visual studies examining the larynx found hyperadduction of the true and false vocal cords (Aronson, 1990, Ziegler and von Cramon, 1986). Studies based on the acoustic approach have documented deviations in F0, decreased harmonic to noise ratio, decrease F0 and intensity variability and vowel prolongation duration (Darley et al., 1975, Patel and Campellone, 2009, Sherrard et al., 2000). These findings are in line with perceptual features including monopitch, strained-harsh voice quality, and slow rate.

The articulation assessment showed decreased speech rate, imprecise articulation related to weakness, slowness and reduced movement range, inadequate tongue and jaw and preserved jaw movements and reduced vowel space (Darley et al., 1975, Ziegler and von Cramon, 1986, Hirose, 1986, Patel and Campellone, 2009). Furthermore, the palate movements are slowed or absent during vowel prolongation, and therefore may lead to hypernasality (Duffy, 2013).

2.3.2.3 Hypokinetic Dysarthria

Parkinson's disease is the most common cause of hypokinetic dysarthria, which is in general characterized by the lack of movements. Nevertheless, features of hypokinetic dysarthria have been also documented in MSA and PSP. Even though the hypokinetic dysarthria is primarily an articulatory disorder it also has an effect on the patient's respiration, phonation and prosody.

Respiration research of patients with hypokinetic dysarthria has documented reduced maximum vowel duration, reduced number of syllables per breath, increased inspiration duration, increased number of inspirations and reduced airflow during vowel production (Canter, 1965b, De Letter et al., 2007, Huber and Darling, 2011, Huber et al., 2003). Although these characteristics also could be attributed to the laryngeal distortion, the physiological examinations have revealed reduced vital capacity, reduced amplitude of chest movements, reduced respiratory muscle strength and endurance, and irregularities in breathing patterns (De Letter et al., 2007, Huber et al., 2003, Solomon and Hixon, 1993, Weiner et al., 2002). Therefore, respiratory disruption may be attributed at least partially to other hypokinetic manifestations, especially, reduced loudness, short phrases, short rushes of speech and inappropriate pauses (Duffy, 2013).

A large number of studies have focused distortion of fundamental frequency (F0) and its variability (Kent et al., 2000b, Skodda et al., 2009, Rusz et al., 2011a, Rusz et al., 2011b, Tykalova et al., 2014, Goberman and Coelho, 2002, Midi et al., 2008,

Bunton, 2006), voice tremor (Gallena et al., 2001), maximum phonation time, voice quality estimation including jitter, shimmer and signal to noise ratios (Rusz et al., 2011b, Adams, 1997, Gamboa et al., 1997, Ludlow and Basich, 1984), laryngeal motor control (Kent and Rosenbek, 1982, Ludlow and Basich, 1984) and laryngeal structure and movement (Hanson et al., 1984). In general, these studies have documented reduced laryngeal efficiency, flexibility and control, which are responsible for distorted phonation and prosody features. These disruptions may in particular be attributed to the rigidity, reduced range and slowness of movements in the laryngeal muscles (Duffy, 2013).

The examination of articulatory deficits documented several distortions including abnormal nasal airflow (Theodoros et al., 1995, Hoodin and Gilbert, 1989), decreased articulatory precision (Ramig et al., 1988), reduced range of jaws, lips (Caligiuri, 1989, Darling and Huber, 2011, Forrest and Weismer, 1995, Hirose, 1986) and also, using the first and the second formant frequencies (F1, F2), the tongue movements (Forrest et al., 1989, Rosen et al., 2006, Rusz et al., 2013b). Among other changes, there are articulatory abnormalities in speech rate (Canter, 1965a, Canter, 1963, Huber and Darling, 2011, Rusz et al., 2011a, Rusz et al., 2011b, Skodda and Schlegel, 2008, Skodda, 2011, Chenausky et al., 2011), reduced force in upper and lower lips, velum and tongue (Netsell et al., 1975, Gentil et al., 2003, Pinto et al., 2003, Solomon et al., 2000) and the jaw and lip tremor (Hunker and Abbs, 1984).

2.3.2.4 Hyperkinetic Dysarthria

Duffy (2013) associates hyperkinetic dysarthria mostly to basal ganglia dysfunction, though he also notes that the cereberal control circuit or other portions of extrapyramidal circuit may be involved. The most common neurodegenerative disease related to hyperkinetic dysarthria is Huntington's disease; however the neurodegenerative causes of hyperkinetic dysarthria are only a minority of all the hyperkinetic dysarthria cases, and most of the reasons are unknown (Duffy, 2013).

There are only few studies concerning hyperkinetic dysarthria. Based on the description provided in (Duffy, 2013), it is possible to detect sudden, forced involuntary inspirations or expirations. This is a distinctive pattern which occurs only in hyperkinetic dysarthria (Darley et al., 1975). Another hyperkinetic dysarthria manifestation is the voice with strained-strangled, breathy and harsh quality with excess loudness variation. Furthermore, abnormal F0 variability, variable VOT and reduced vowel duration have been detected in HD patients (Hertrich and Ackermann, 1994,

Ramig, 1986, Zwirner et al., 1993, Skodda et al., 2014a, Skodda et al., 2014b, Rusz et al., 2014c, Rusz et al., 2014b). One of the most distinct patterns is the presence of sudden voice arrests (Rusz et al., 2014c).

Articulatory deficits are thought to be the most prominent, yet not the most distinguishing, and comprise hypernasality, distorted vowels and irregular articulatory breakdowns caused by choreiform movements of palate, tongue, jaws and lips. The abnormal F1 and F2 variability in Huntington's disease has been documented and is considered to have negative effect on prosody (Ackermann and Ziegler, 1995, Hertrich and Ackermann, 1994, Duffy, 2013). Prosodic disturbances are the most prominent, reflecting the impact of chorea on speech. Darley et al. (1975) identified prolonged intervals, inappropriate silences, prolonged phonemes, excess and equal stress, monopitch, monoloudness, reduced stress, and short phrases.

2.4 Methodology of Dysarthria Assessment

The gold standard of dysarthria analysis is perceptual assessment (Kent, 2000). Nevertheless, the condition of dysarthria is a complex speech disorder distributed among all speech systems, including respiration, phonation, articulation and prosody, all of which are more or less affected depending on dysarthria subtype. The effect of masking is common when several dysarthria manifestations overlap, which makes perceptual assessment difficult. Furthermore it has been documented that the reliability of perceptual assessment is limited due to its subjective nature (Brancewicz and Reich, 1989). Therefore the objective instrumental parameters are used to provide useful information and support for the perceptual assessment.

One of the most affordable instrumental methods is acoustic analysis, which is based on similar approaches as the perceptual assessment. In accordance to the measured speech feature, a proper speech task is chosen similarly as in the perceptual assessment. These tasks are designed to reflect specific speech features and ideally to diminish the effect of other speech aspects. These four basic speech tasks consist of sustained phonations, speech diadochokinetic tasks, rhythmic tasks and connected speech. Connected speech provides the most complex task enabling assessment of all speech subsystems, though it is the most prone towards effect of masking. The rhythmic tasks are design to address speech timing (Skodda and Schlegel, 2008, Rusz et al., 2015). The DDK task, due to the rapid repetitive character, is sensitive to disruptions of articulatory movements (Rusz et al., 2011b). Sustained phonation is the most common

speech task, even though it mostly reflects dysphonia (Little et al., 2009, Tsanas et al., 2012). The sustained phonation could be used for estimation of articulation to a limited extent describing formant frequencies and the vocalic triangle (Rusz et al., 2013b), and estimating distorted velopharyngeal control (Kataoka et al., 1996).

2.4.1 Rapid Articulatory Moves

The perceptual assessment of articulatory moves is limited by the masking effect, which is the result of the strong acoustic overlap of dysarthria manifestations. For instance the distorted tongue movements are manifested by distorted formant frequencies (Kent et al., 2000b), yet it has been shown that formant frequencies are also affected by disrupted velopharyngeal control (Fant, 1960). Moreover, the presence of a strained, harsh, or noisy voice further distorts perceptual assessment. For this reason, speech tasks designed to highlight articulatory deficits are usually employed.

One of the common speech tasks for the articulatory movement assessment is the diadochokinetic speech task, defined as rapid repetitions of single syllable words or series. For instance, one of the common DDK tasks is the consonant vowel (CV) train of /pa/, /ta/ and /ka/ syllables. During this task the participant is asked to perform the fastest as well as the most steady /pa/-/ta/-/ka/ repetitions which are still intelligible. This series contains transitions between consonant and vowels as well as tongue transitions between bilabial, alveolar and velar positions. For this reason, the DDK task provides a reasonable trade-off between simple sustained phonation, which is not suitable to reflect the variety of different articulation disruptions, and the connected speech, which is more prone towards dysarthria feature overlap. Indeed, it has been shown that in some neurodegenerative diseases, patients are able to willingly compensate for the articulatory disruptions (Ackermann et al., 1995a). However, the effort to produce the utterance as fast as possible limits this ability to compensate deficits to some extent.

The most common parameters estimated on the basis of the DDK task are estimations of articulation rate including speech rate (DDK rate) and variability of speech rate (DDK regularity; Fletcher, 1972, Rusz et al., 2011b, Rusz et al., 2014a, Weismer, 1984). Among other parameters proposed by Rusz et al. (2011b) belong the relative intensity range variation as variations of energy, robust relative intensity slope assessing linear regression of energy, spectral distance change variation as a variations of spectral distance changes in signal spectrum and robust formant periodicity correlation describing the first autocorrelation coefficient of F2 contour. Furthermore,

other parameters estimating distinct articulation dimensions may be applied to the DKK task, namely the voice onset time (Kent et al., 2000b, Fischer and Goberman, 2010), reflecting laryngeal and supralaryngeal coordination, consonant spectral characteristics representing articulatory undershoot (Kent et al., 2000b) and signal-to-noise ratio addressing occlusive weakening (Duez, 2007).

2.4.2 Velopharyngeal Control

Currently the most common method for hypernasality estimation is perceptual rating. However, inter-rater and intra-rater reliability is questionable and perceptual rating requires a trained speech specialist (Kuehn and Moller, 2000). Consequently, more objective methods have been developed to complement perceptual ratings. Invasive methods, such as x-ray tracing with a lead pellet attached to the velum, provide direct observation of velopharyngeal movements (Hirose et al., 1981). Other methods employ indirect estimation based on measurements of nasal airflow, nasal cavity sonography, nasometry comparing nasal and oral acoustic outputs, or the Horii Oral-Nasal Coupling Index (Hardin et al., 1992, Dillenschneider et al., 1973, Horii, 1980). One of the least demanding methods with respect to patients and equipment is the 1/3-octave spectra, which is based on direct, non-invasive analysis of acoustic speech signal and was originally developed for the estimation of velopharyngeal insufficiency in cleft palate (Kataoka et al., 1996) and was later validated by Vogel et al. (2009).

The 1/3-octave spectra method is a type of spectral analysis based on the examination of spectral changes caused by resonatory speech pathologies. This method is based upon the linear source–filter theory of speech, which was first described by Gunnar Fant (Fant, 1960). The linear source–filter theory presents speech as a product of three basic components including source characteristic S(f) connected with fundamental frequency, vocal tract transfer function T(f) associated with formant frequencies and mouth radiation characteristic R(f) (Kent and Read, 2002, Stevens, 2000). The source characteristic approximates the spectrum of vocal folds vibrations, which is filtered by the transfer function of the vocal tract and further modified by the radiation characteristics of the mouth (see Fig. 7; Kent and Read, 2002, Stevens, 2000). Although the transfer function is essentially given by the length of the vocal tract, it may be modified by vocal tract adjustments. Therefore different voiced phonemes are produced by different settings of the cross-sectional area through the vocal tract. Variations in the vocal tract cross-sectional area shape and shift formant frequency

peaks along the frequency axis. The shape of the vocal tract is then approximated by the resonatory model. For instance, the vowel /ae/ may be represented by two connected resonators, where the first representing the pharynx has significantly smaller cross-sectional area than the consecutive tube approximating the oral cavity with low tongue position (Kent and Read, 2002, Stevens, 2000). These settings result in a spectrum with close F1, F2 and distant F3. Conversely, due to constriction in the oral cavity created by the tongue, the vowel /i/ may be approximated by a Helmholtz resonator with a low distant F1 and high close F2 and F3, as is illustrated in Figure 2.7. (Kent and Read, 2002, Stevens, 2000, Fant, 1960).

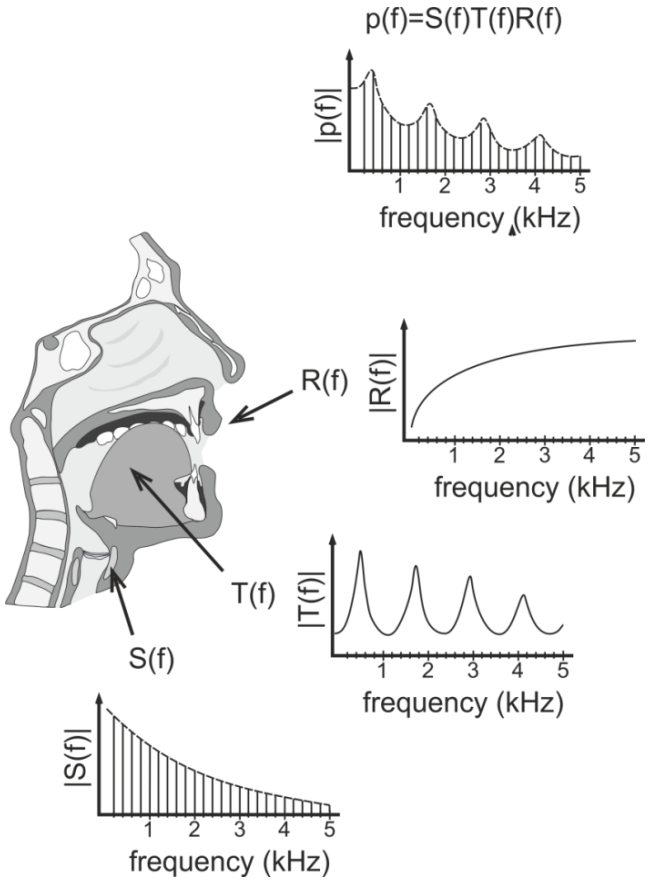

Figure 2.7: Source of voicing S(f), vocal tract transfer function T(f) and radiation characteristic R(f) with final sound pressure p(f) as a product of these three functions with their marked positions along the vocal organs.

When approximating a straight tube, every transfer function is an all-pole filter where the poles represent only formant frequencies, even though different cross-sectional area along the vocal tract leads to differing spectral characteristics (Stevens, 2000). Nevertheless, in the case that nasality occurs, the nasal cavity passage is parallel to the oral cavity and the nasal poles and zeros are introduced to the vocal tract transfer

function (Stevens, 2000). This affects the shape and position of the formant frequencies and implements nasal resonance Fn at an area around 1 kHz (Stevens, 2000, Kataoka et al., 1996). Even if there are some general effects such as a flatter spectrum shape up to 1200 kHz due to the wider and lower F1 and the presence of Fn, the exact effect is dependent on the shape of vowel spectrum (Stevens, 2000). Therefore, F1 is shifted to the lower frequencies in the back vowel /ae/, whereas the prominence of F2 may be significantly decreased due to proximity to nasal zero. In the case of the front vowel /i/, the nasal F1 is shifted to higher frequencies compared to the non-nasal F1. Formants F2 and F3 are not significantly affected by hypernasality as illustrated in Figure 2.8. Furthermore, due to the distance between F1 and F2 in the /i/ spectrum, the Fn peak is more evident than in the /ae/ vowel.

As a matter of this fact, the vowel /i/ appears to have the best potential for hypernasality detection based on spectral analysis. Being the most evident, nasal resonance in the vowel /i/ should be more robust to anatomical variation of the nasal cavity including asymmetrical shape and varying shape of the connected sinuses. Moreover, the vowel /i/ is considered to be the most sensitive to nasal coupling (Fant, 1960) and thus previous studies have focused on the quantitative evaluation of VIS hypernasality through the sustained vowel /i/ (Kataoka et al., 1996, Yoshida et al., 2000, Lee et al., 2003). Based on experiments with experienced listeners and rating of nasalance in artificially generated sounds in patients with cleft palate and those that underwent maxillectomy, previous studies have confirmed the vowel /i/ as an ideal speech task for hypernasality assessment.

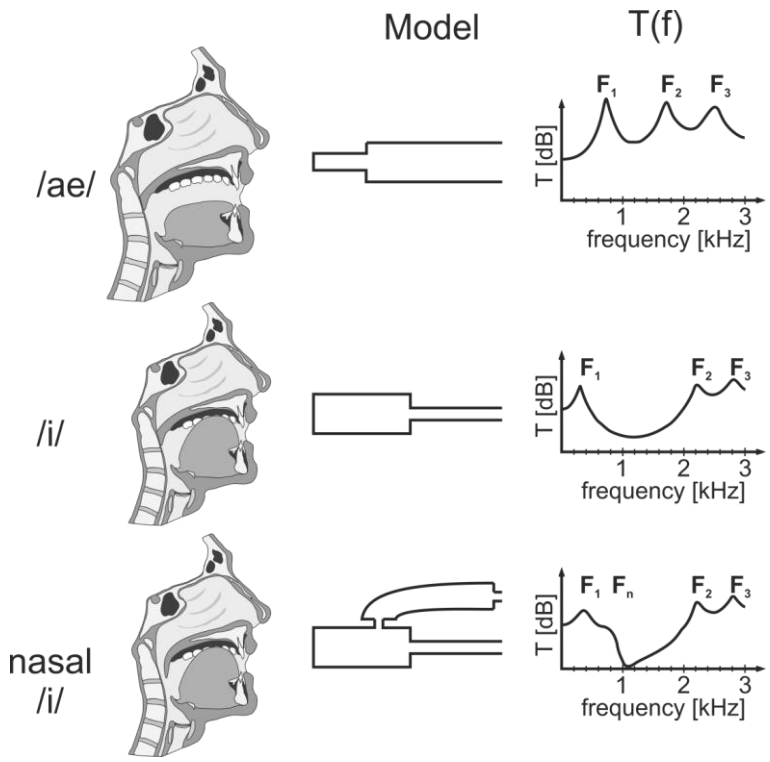

Figure 2.8: Three settings of the vocal tract for the vowels /ae/, /i/ and /i/ with present nasality. Each vowel is represented by the anatomical setting, acoustic model of the vocal tract and shape of the vocal tract transfer function estimated in Stevens (2000).

3 OBJECTIVES AND HYPOTHESES

3.1 Rapid Articulatory Moves

Although articulatory deficits represent an important manifestation of dysarthria in Parkinson's disease, the most widely used methods currently available for the automatic evaluation of speech performance are focused on the assessment of dysphonia.

We hypothesize that the acoustical assessment of articulatory moves during the DDK task may provide useful information about, articulation of laryngeal and supralaryngeal articulators. Furthermore we assume that on the basis of the articulatory assessment it is possible to perform automatic classification of PD vs. healthy speakers.

Furthermore we assume that it is possible to extend application of DDK-based assessment on other dysarthria types including hyperkinetic dysarthria related to HD. And more importantly to apply objective assessment of articulation disruption on atypical parkinsonian syndromes such as progressive supranuclear palsy and multiple system atrophy to verify whether any specific speech characteristics allow differentiation between PD, PSP and MSA.

We hypothesize that due to the different pathophysiology of PD, PSP and MSA the distinctive speech patterns may be detected using acoustical assessment of speech performance and these patterns may provide useful clues for the differential diagnosis.

Therefore aims of the section Rapid articulatory moves are:

- **To develop an automatic segmentation algorithm allowing for the accurate detection of the initial burst, vowel onset, and occlusion in patients diagnosed with PD.**

- **To extend the detection algorithm to be applicable to different dysarthria subtypes, and to demonstrate its applicability by evaluating PD and HD speakers.**

- **Using the proposed segmentation algorithm to introduce several acoustic features sensitive to possible articulatory deficits caused by hypokinetic dysarthria.**

- **To explore the suitability of the designed acoustic features in capturing parkinsonian articulatory disorder, and to perform a classification experiment in order to differentiate PD subjects from controls.**

- **To determine specific patterns of articulatory disruptions and estimate their reliability in differentiating between PD, PSP and MSA**

3.2 Velopharyngeal Control

The purpose of the velopharyngeal control section is to analyze acoustic and perceptual correlates of the velopharyngeal seal closure in Parkinson's disease and Huntington's disease participants in comparison to healthy control (HC) speakers.

We hypothesize that bradykinetic distortion of palatal control in PD may distort articulation of the velopharyngeal seal and accordingly lead to steady air leakage and increased hypernasality. Moreover, distorted neuromuscular control of levator veli palatini in PD may also lead to increase of hypernasality with increased fatigue during speech tasks. Conversely in HD, we hypothesize that choreatic movements of the velopharyngeal seal and velum may lead to varying resonance distortion, which would be in agreement with reported intermittent hyperkinetic dysarthria.

Furthermore we hypothesize that the different underlying pathology of PD, PSP and MSA results in different resonatory disruption patterns and therefore may provide useful clue for the differential diagnosis of PD, PSP and MSA.

Therefore goals of the section Resonance are:

- **To employ methods of objective hypernasality assessment, which may be easily grasped by between the tools of the differential diagnosis.**

- **To evaluate the presence and character of hypernasality in PD and HD speakers.**

- **To examine possible relationships between the severity of hypernasality and disease-specific motor manifestations, to provide more insight into the pathophysiology responsible for development of hypernasality in basal ganglia disorders.**

- **To determine specific resonatory disruption patterns characteristic for PSP, MSA and PD.**

4 METHODS

4.1 Preliminary Note

The thesis comprises four studies which can be divided into two groups. The first three studies describe DDK-based approach of estimation of articulatory movement deficits connected with neurodegenerative diseases. The first two depict the automatic approach for the evaluation of articulation and application in early diagnosis. The third study concerns articulatory differences between idiopathic PD and atypical parkinsonian syndromes (MSA and PSP). The second group describes the effect of neurodegeneration on velopharyngeal control in the PD, the APS and the HD.

Because of the different purposes of each study, different participant groups were formed, as a result of the different study designs and the necessity to match different participant groups. All the participants were Czech native speakers.

The studies were approved by the Ethics Committee of the General University Hospital in Prague and all participants provided written, informed consent.

4.2 General Methods

4.2.1 Clinical Diagnostic Criteria

The inclusion criteria of patients for Parkinson's disease were UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria (Hughes et al., 1992); for progressive supranuclear palsy were used NINDS-PSP (Litvan et al., 1996); for multiple system atrophy were used the consensus diagnostic criteria (Gilman et al., 2008); Huntington's disease patients were genetically confirmed.

UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria consist of: Step 1) presence of bradykinesia (slowness of initiation of voluntary movement with progressive reduction in speed and amplitude of repetitive actions) and at least one of the following: muscular rigidity, 4-6 Hz rest tremor, postural instability not caused by primary visual, vestibular, cerebellar, or proprioceptive dysfunction; Step 2) exclusion criteria for Parkinson's disease; and Step 3) supportive prospective positive criteria for Parkinson's disease including excellent response to L-dopa.

The NINDS-PSP comprises three groups of criteria including mandatory inclusion criteria, mandatory exclusion criteria and supportive criteria. Furthermore, the inclusion criteria are divided into three degrees, these being PSP possible, PSP probable and PSP definite. The condition of possible PSP is described by presence of a gradually progressing disorder with an onset after 40 years of age, vertical supranuclear gaze palsy or both slowing of vertical saccades and prominent postural instability with falls in the first year of onset, as well as the absence of evidence of other diseases that could explain these features. The probable PSP degree requires vertical supranuclear palsy, prominent postural instability and falls in the first year of onset, accompanied as well by the other features of possible PSP. The PSP definite degree is defined by a history of possible or probable PSP and histologic evidence of typical PSP.

The consensual diagnostic criteria for MSA define three criteria groups describing clinical MSA certainty as possible, probable and definite. The definite group requires MSA neuropathological confirmation. The possible MSA degree requires a sporadic, progressive adult-onset disease including parkinsonism or cerebellar ataxia and at least one feature suggesting autonomic dysfunction and one other feature which may be a clinical or a neuroimaging abnormality. The Probable MSA is described by presence of sporadic, progressive adult-onset disease including rigorously defined

autonomic failure and poorly levodopa-responsive parkinsonism or cerebellar ataxia. The definite MSA requires histologic confirmation of MSA.

4.2.2 Motor Manifestation Scales

All participants diagnosed with neurodegenerative disease i.e. PD, HD, PSP, or MSA were rated according to the corresponding rating scale.

The Unified Parkinson's Disease Rating Scale III (UPDRS; Hughes et al., 1992) and Hoehn & Yahr scale (H&Y; Hoehn and Yahr, 1967) were used for the assessment of PD motor manifestations. The UPDRS III scale contains 27 items, each scored from 0 (no disability) to 4 (severe disability). We estimated several UPDRS composite subscores including *rigidity* (sum of UPDRS III item 22), *bradykinesia* (sum of UPDRS III items 23, 24, and 25), *resting tremor* (sum of UPDRS III item 22), and *axial score* (sum of UPDRS III items 18, 19, 27, 28, 29, and 30). Furthermore, UPDRS III item 18 was used to estimate speech motor deficits in PD. The UPDRS III item 18 is concerned with the assessment of speech, and is ranked from 0 to 4, where 0 represents normal speech; 1 slight loss of expression, diction and volume; 2 monotone slurred but understandable speech, moderately impaired; 3 marked speech impairment, difficult to understand; and 4 unintelligible speech. Hoehn & Yahr scale contains five grades of PD severity and is commonly used for the description of PD progression. The scale comprehends severity from a mild unilateral motor disorder as the first grade, to confinement to bed or wheelchair as the fifth grade.

The Huntington's disease was rated according to the Unified Huntington's Disease Rating Scale (UHDRS; Huntington-Study-Group, 1996). The UHDRS scale contains 31 items, each scored from 0 (no disability) to 4 (severe disability). Using compound groups of UHDRS items, we estimated *rigidity*, *bradykinesia*, *dystonia*, and *chorea*.

The motor manifestations described by the UHDRS and the UPDRS are defined as follows:

- *Rigidity* represents elevated muscle tone experienced as muscle tension or spasm by the patient and as increased resistance to passive movement across joints by the examiner.
- *Bradykinesia* describes slowed movements and reaction times and may include difficulties with fine motor control.

- *Resting tremor* involves a unilateral shaking movement with a frequency between 4–6 Hz, mainly affecting the hands, legs, lips, chin, and jaw and subsiding with muscle action and during sleep.
- *Axial score* represents speech intelligibility, extent of facial expressions and reflects the ability to rise from a chair, problems with postural stability related to the loss of reflexes and problems during walking including sudden and transient inability to move.
- *Dystonia* is involuntary sustained muscle contraction that causes twisting and repetitive movements or abnormal postures.
- *Chorea* is a state of excessive, spontaneous movements, irregularly timed, non-repetitive, randomly distributed and abrupt in character.

The MSA and PSP patient were rated using The Natural History and Neuroprotection in Parkinson Plus Syndromes (NNIPPS; Payan et al., 2011). The NNIPPS comprises 85 items scored mostly (65) from 0 to 5 and with the maximal range from 0 to 6, where 0 means "normal" and 6 is "very severe". Similarly to PD participants the MSA and PSP patients were rated according to the UPDRS III item 18 describing speech.

4.2.3 Recording

All recordings took place in a quiet room with a low ambient noise level using a head-mounted condenser microphone (Beyer-dynamic Opus 55-, Heilbronn, Germany) positioned approximately 5 cm from each subject's mouth. The utterances were sampled at 48 kHz with 16 bit quantization. All the voice signals were obtained during single session conducted by a speech specialist, who asked participants to perform several speech tasks including sustained phonation of vowel /i/, speech /pa/-/ta/-/ka/ DDK_task and freely spoken monologue. During the DDK_task participants were instructed to take a deep breath and intelligibly repeat /pa/-/ta/-/ka/ syllable sequence as fast and as steady as possible. During the sustained phonation participants were asked to take a deep breath and perform sustained phonation of vowel /i/ at a comfortable loudness and pitch, as constant and long as possible. The inclusion criteria imposed in study IV were determined as the participant's ability to sustain prolonged phonation for at least three seconds. The DDK task, sustained phonation were performed twice. The freely spoken monologue was on a given topic including family, work or interests, and lasted at least

two minutes. All three tasks were part of a comprehensive dysarthria test battery. No time limits were imposed during recording.

4.3 Rapid Articulatory Moves

4.3.1 Subjects

Data were collected as part of an original studies (Rusz et al., 2011a, Rusz et al., 2014b) and (Rusz et al., 2015) and comprises six different groups including two PD groups, one matched to HC group (PD_U) and one used for the comparison with APS (PD_{SDD}), one HD group, one healthy control group, one PSP group and one MSA group.

The PD_U group consisted of 24 participants (20 men, 4 women), all of whom fulfilled the diagnostic criteria for PD. All PD speakers were examined immediately after the diagnosis was made and before symptomatic treatment was initiated. The mean age of PD participants was 60.9 ± standard deviation (SD) 12.6 years, mean disease duration 31.3 ± SD 22.3 months, disease stage 2.2 ± SD 0.5 according to the Hoehn & Yahr scale, mean motor score 17.4 ± SD 7.1 according to the Unified Parkinson's Disease Rating Scale III. In agreement with perceptual evaluations based on UPDRS III item 18, 13 patients obtained a score of 0 and 11 patients a score of 1, suggesting no-to-mild speech impairment. None of the PD patients reported previous speech disorders unrelated to the present illness.

The healthy control group was comprised of 22 volunteers (15 men, 7 women; mean age 58.7 ± SD 4.6 years) with no history of neurological disease. No differences in age between the PD and HC groups were observed (two-sample *t*-test; $t(44) = -0.89$, $p = 0.38$).

The hyperkinetic dysarthria was represented by 77 utterances from 40 speakers (20 men and 20 women) diagnosed with HD. The mean age of the HD group was 48.6 ± SD 13.4 years. All the participants were assessed by specialist using UHDRS; the HD group received UHDRS motor score of 26.9 ± SD 11.6.

The age of PD_{SDD} participants (9 men 6 women) was 61.1 ± SD 6.5 years with the symptom duration 4.6 ± SD 1.5 years. The H&Y score for the PD_{SDD} group was 2.0± SD 0.4, the UPDRS III score was 15.9 ± SD 7.4 and the UPDRS III item 18 was 0.6 ± SD 0.5. Contrary to the PD_U group, the PD_{SDD} participants were on stable symptomatic treatment and they were recorded during the ON state.

The PSP participants (10 men 2 women) comprised 9 participants diagnosed with PSP-RS, 2 participants were diagnosed with PSP-P and one with PSP-PAGF. The

age of PSP group was 65.8 ± SD 5.4 years with the symptom duration 3.8 ± SD 1.4 years. The H&Y score 3.3 ± SD 0.8 and the NNIPS score of PSP group was 66.3 ± SD 28.7 and the UPDRS III item 18 was 2.0 ± SD 1.0.

The MSA group (6 men and 7 women) included 10 participants diagnosed with MSA-P and 3 participants diagnosed with MSA-C. The age of MSA group was 60.8 ± SD 4.9 years with symptom duration 3.68 ± SD 1.3 years. The H&Y score 3.6 ± SD 0.7 and the NNIPS score of MSA group was 78.5 ± SD 19.9 and the UPDRS III item 18 was 2.0 ± SD 0.7.

No participant had history of speech therapy. All the patient characteristics are summarized in Table 4.1

Table 4.1: Clinical characteristics of participants

	HC mean/SD	PD_U mean/SD	HD mean/SD	PD_{SDD} mean/SD	PSP mean/SD	MSA mean/SD
Age	58.7/4.6	60.9/12.6	48.6/13.4	61.1/6.5	60.8/6.5	65.8/5.4
Symptoms duration		2.6/1.9	6.1/3.4	4.6/1.5	3.6/1.3	3.8/1.4
CAG triplets			44.9/3.6			
L-dopa equivalent		0.0/0.0		615/317	899/394	800/373
Amantadine		0.0/0.0			300/89	200/107
NNIPPS					78.5/19.9	66.3/28.7
UPDRS III		17.4/7.1		15.9/7.4		
UPDRS III speech item 18		0.5/0.5		0.6/0.5	2.0/0.7	2.0/0.1
H&Y		2.2/0.5		2.0/0.4	3.6/0.7	3.3/0.8
UHDRS			26.9/11.6			
UHDRS speech item			0.8/0.5			

4.3.2 Reference Labels

The manually obtained reference labels of /pa/, /ta/ and /ka/ syllables were used for the purposes of performance estimation of the Study I and Study II which concerns rapid articulatory moves. As can be seen in Figure 4.1, each /pa/, /ta/, or /ka/ syllable consists of an initial burst, vowel onset, and occlusion. These three basic events generally describe the timing of articulation, and their positions must be detected in each syllable to analyze articulation deficits. Thus, reference labels must first be established, and this procedure requires the manual segmentation of each utterance. However, segmentation may be challenging even for manual labeling and therefore, the criteria according to which labeling was performed must be stated. Fischer and

Goberman (2010) summarize three basic rules based on previous (Wang et al., 2004, Volaitis and Miller, 1992, Allen et al., 2003, Duez, 2007) which were used as a foundation for our labeling criteria (see Table 4.2).

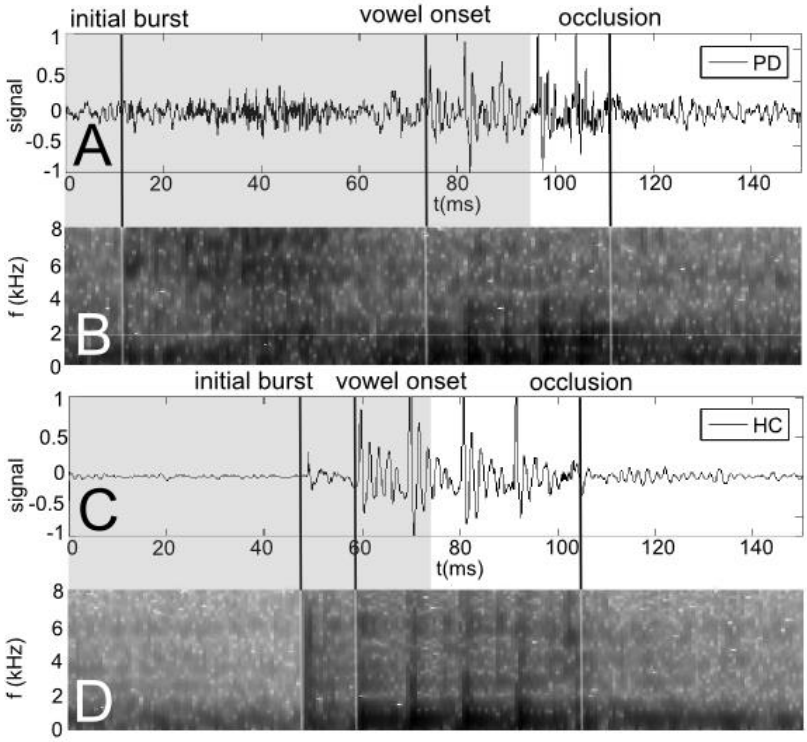

Figure 4.1: Examples of syllable and its wideband spectrogram for Parkinsonian (A), (B) and healthy (C), (D) speakers, with marked positions of the initial burst, vowel onset and occlusion. The gray background shows the front part of the syllable and the white background refers to the rear part of the syllable.

Table 4.2: Labelling Criteria based on (Fischer and Goberman, 2010)

Position	Description	Frequency domain	Time domanin
Initial Burst			
	Abrupt onset of noise energy caused by turbulent airflow during stop release. Good contrast in time and moderate energy contrast.	Moderate excitation of one or a few time windows of the spectrogram over the entire frequency range.	Used for specification of burst onset. In the case of multiple bursts the initial burst is marked (Wang et al., 2004).
Vowel Onset			
	Abrupt onset of periodic signal with highest acoustic energy caused by vocal fold vibration. Good contrast in time and best energy contrast.	Onset of fundamental (F0) and first formant frequencies (F1, F2, F3) (Volaitis and Miller, 1992, Allen et al., 2003). Energy is concentrated to these frequencies.	Position with highest contrast. If the abrupt onset of energy is not clearly apparent, the F0, F1, and F2 onset is sought.
Occlusion			
	Slow voice weakening, and therefore slow weakening of F0, F1, F2, and F3. Fuzzy due to weak time and energy contrast.	Energy of F0, F1, F2, and F3 slowly weakens. The F2-vowel offset is considered the best indicator of occlusion onset (Duez, 2007).	Used especially to boost the robustness of labeling. Needed especially due to slow energy weakening.

4.3.3 Graphical User Interface

For the purposes of manual labeling and the visual control of automatically detected labels, a basic graphic user interface (GUI) was developed. The GUI consists of a command window used for assignment of the directory containing files and a simple GUI screen which includes a list of utterances, the graph illustrating the entire utterance and the graph describing 0.3s of utterance as a detail and the control panel, Figure 4.2. Correction and obtaining is performed by dragging the lines in the screen or using mouse double clicks to delete an old or create a new label position. To be able to navigate easily through single utterances, the three buttons are implemented as: “load previous”, “load next” and ”load”, where the load is used to load the signal highlighted in the signal list. The fourth button is used to save new position labels. For the needs of automatic detection correction, the displaying of detected positions is enabled through the check box option, which enables the loading of the mat-file containing detected positions.

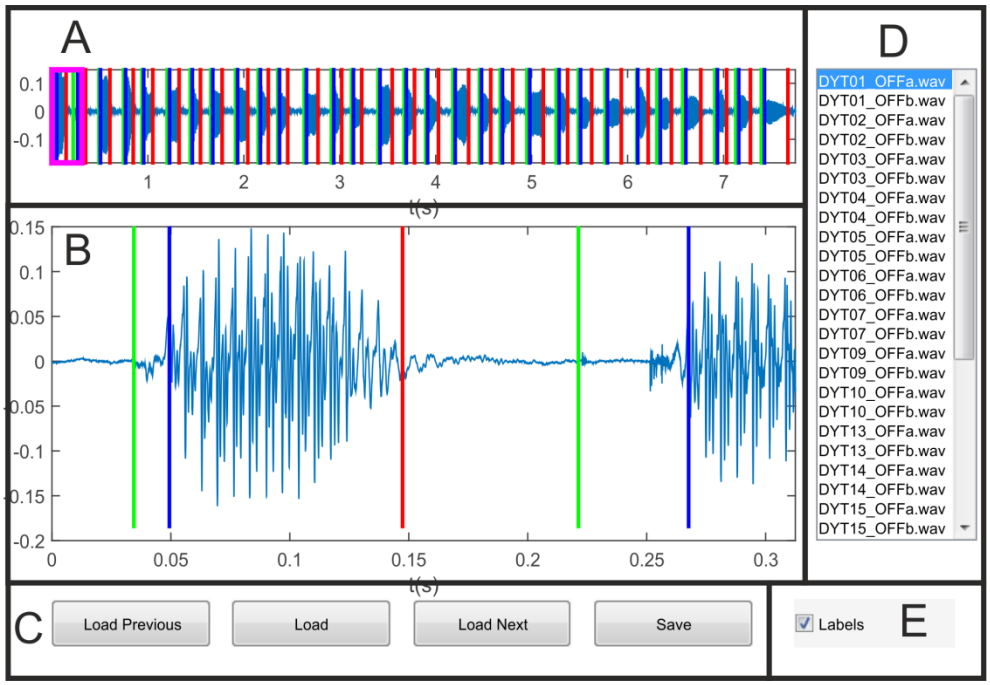

Figure 4.2: The graphic user interface divided into five functional areas: A) section illustrating the entire utterance, with marked positions of initial burst (green lines), vowel onset (blue lines) and occlusion (red lines), and the magenta window marking the position of the detail window B) the detail window depicting 0.3 seconds of the signal. Similarly as in A), the single event positions are marked by lines with different color. C) Panel of control buttons enabling loading of antecedent or consecutive signals or load signal chosen in the signal list shown in D). The fourth button is used to save reference or corrected labels. The check box in E) is used to distinguish between obtaining of reference labels and correction of automatically detected.

4.3.4 Algorithm of Automatic Segmentation

The manual labelling is a time consuming process, and may be biased by subjective evaluation. To lower time demands and provide objective results, the deterministic detector of stop release, vowel onset, and occlusion was designed. The algorithm is presented in several subsections describing pre-processing, a rough segmentation, a detection of initial burst, a detection of vowel onset, and a detection of occlusion.

4.3.4.1 Pre-processing

The pre-processing step comprises re-sampling of the signal to 20 kHz which lowers the computational complexity and maintains useful speech information (Titze, 1994). The pre-processing step also includes DC offset removal and normalization of the signal to the interval [-1, 1].

4.3.4.2 Rough segmentation

The first problem of automatic processing is the unknown number of syllables. This problem is solved by rough segmentation which splits utterance to single syllables (see Figure 4.3). These syllables are than processed separately. To split the signal into single syllables, the approximate position of each syllabic nucleus has to be estimated. We may assume that in the DDK task, each syllable consists of one low-energy consonant and one high-energy vowel. Therefore, positions of syllabic nuclei may be identified by high-energy vowel peaks. However, the presence of a higher noise component in PD utterances may bias the nuclei search, and therefore filtering must be performed. Filtering was accomplished by a low-pass FIR filter with a linear phase and order of 500 with a 300 Hz cut-off frequency. The filtered signal is squared and smoothed by the moving average filter of order 800 and local energy maxima are detected. We noted that when one syllable has considerably lower energy than its neighboring syllables, detection based on 300 Hz filtering tends to omit the syllable. Hence, the same detection based on a low-pass filter with a 1000 Hz cut-off frequency was used. The detector based only on 1000 Hz filtering was more vulnerable to the higher noise component included in PD utterances, and therefore it was used only as a complement to more robust, 300 Hz filter-based detector. The maximal distance between two consecutive nuclei was estimated and enlarged 1.1 times, providing the length of a single syllable segment. This length was distributed before and behind the energy peaks, providing the approximate borders for each syllable.

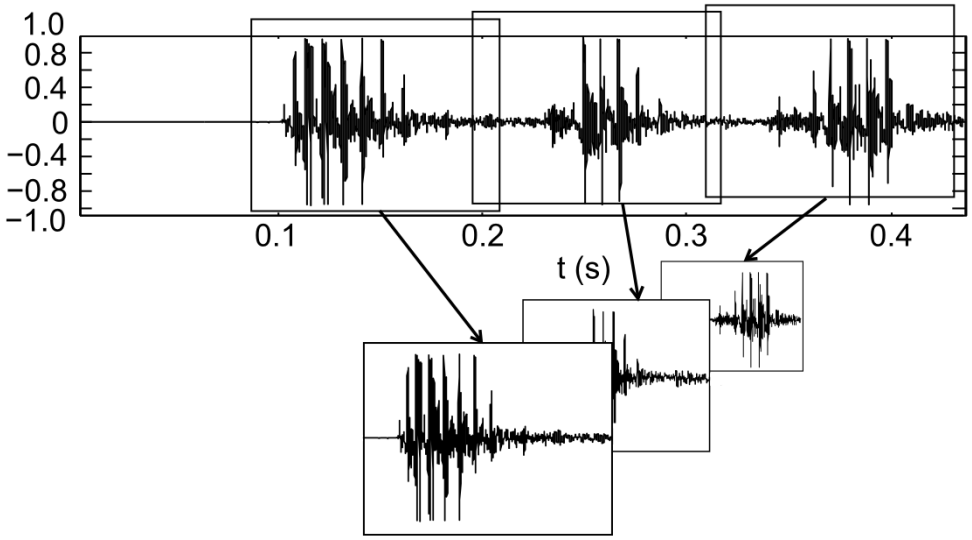

Figure 4.3: Detail of an utterance divided by rough segmentation into single syllabic segments.

To avoid false detections due to the high sensitivity of the detector, the elimination of false positions must be implemented as the second step of rough segmentation. The elimination was based on the comparison of high and low energy centroid positions obtained from the filtered spectrogram around the vowel onset. Due to higher computational complexity, the spectrogram was also utilized during the detection of the initial burst, which was also spectrogram-based.

The spectrogram window length was defined as the length of the processed signal divided by 120, and the overlap was equal to one half of a window. To increase efficiency, the unnecessary rear part was omitted and the spectrogram was computed only from the front part of the syllable (gray part of the syllable highlighted in Figure 4.3).

Spectrogram processing consisted of the elimination of negligible values and computation of energy envelopes. To determine which value was negligible, the spectrogram was treated as a matrix $\vec{P}$ with *m* rows for frequency bins and *n* columns for time bins. The threshold matrix $\vec{T}$ was an *n* by *m* matrix, where the *i*-th row was computed from the *i*-th frequency bin of the spectrogram according to equation (1).

$$\vec{T}(i,1\ldots n) = 0.8 * \frac{1}{n}\sum^{n}\vec{P}(i,1\ldots n). \tag{1}$$

This equation sets each row of the threshold matrix as the weighted mean value of energy contained in the equivalent frequency row of matrix. Filtering was then performed as shown in equation (2).

$$P_{filtered}(i,j) = \begin{cases} P(i,j), & P(i,j) \geq T(i,j) \\ 0 & P(i,j) < T)(i,j) \end{cases}, \tag{2}$$

where $P_{filtered}$ denotes element contained in the *i*–th frequency bin and the *j*-th time bin of the filtered matrix. An example of a filtered spectrogram can be seen in Figure 4.4.

The next processing step was the computation of two energy envelopes. The first was calculated by summing the values in each column (Figure 4.4 (C)), while the second was determined by summing values only in the upper half of each column (Figure 4.4 (D)). The first envelope considers the high energy of vowels contained mostly in low frequencies; the second emphasizes high frequencies generated during the initial burst. Centroids were computed from these envelopes and their absolute and mutual positions were used for the elimination of false detections. The centroid positions are marked as black arrows in Figure 4.4 (C) and Figure 4.4 (D). The energy

envelope comprising the entire frequency bandwidth provides facilities for rough vowel onset estimation. The position of vowel onset was set as the first peak of the voicing periodic sequence. This approach was based on the assumption that, during the voicing, the vocal tract is excited by quasi-periodically repeating glottal pulses (Harris and Nelson, 1993). In processing the front part of the syllable (see Figure 4.3), peaks may be traced from the end of the envelope (see Figure 4.4 (C)). However, this estimation sometimes marks the accentuated initial burst instead of vowel onset, and therefore, it is sufficient only for the correction of syllable position.

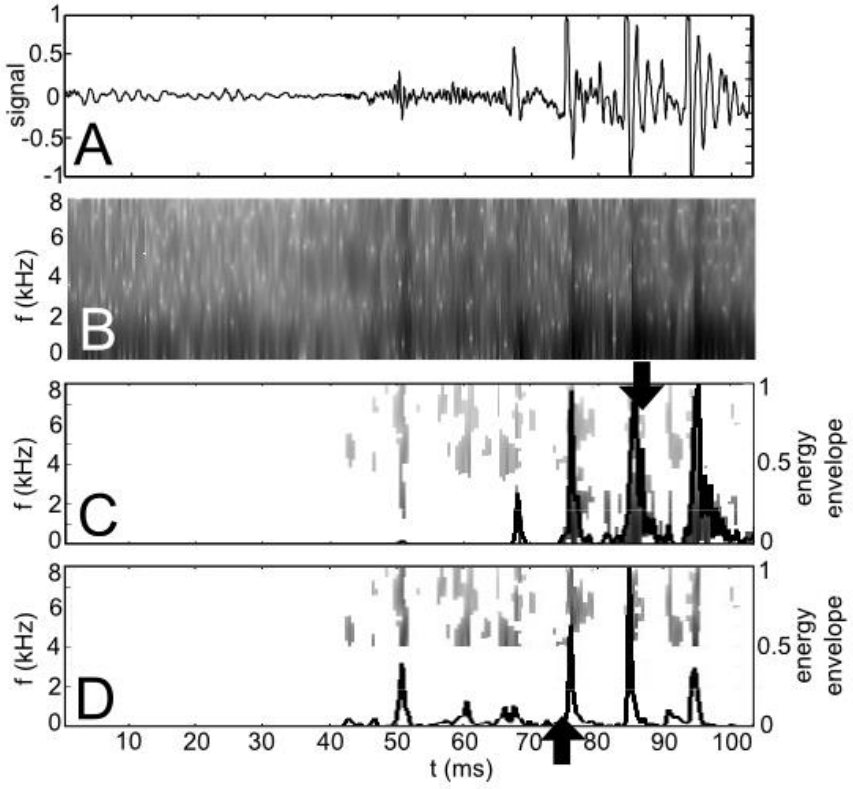

Figure 4.4: Signal in the time domain (A), signal spectrogram (B), filtered spectrograms with marked energy envelopes and arrows pointing to spectral centroid positions in the entire frequency range (C), and the upper half of the frequency range (D).

4.3.4.3 Detection of the Initial Burst

After the elimination of false detections and correction of the segment borders, the noise burst connected with the initial stop release was sought. For the purposes of burst detection, the previously computed spectrogram was processed according to a modification of eq. (2) (see eq. (3)),

$$P_{stop\_release}(i,j)=\begin{cases}1 & P(i,j)\geq T(i,j)\\ 0 & P(i,j)<T)(i,j)\end{cases}, \qquad (3)$$

where the $\vec{T}$ matrix is given by eq. (1). The result of this filtering can be seen in Figure 4.4.

The envelope, given by summing all values in each time window of the matrix $P_{stop\_release}$, emphasizes information about frequency bandwidth at the expense of information about energy distribution. This method emphasizes the noise burst, which has lower energy uniformly distributed through the entire spectrum. Furthermore, due to abrupt onset, the difference of the envelope highlights and specifies the stop release position as shown in Figure 4.5.

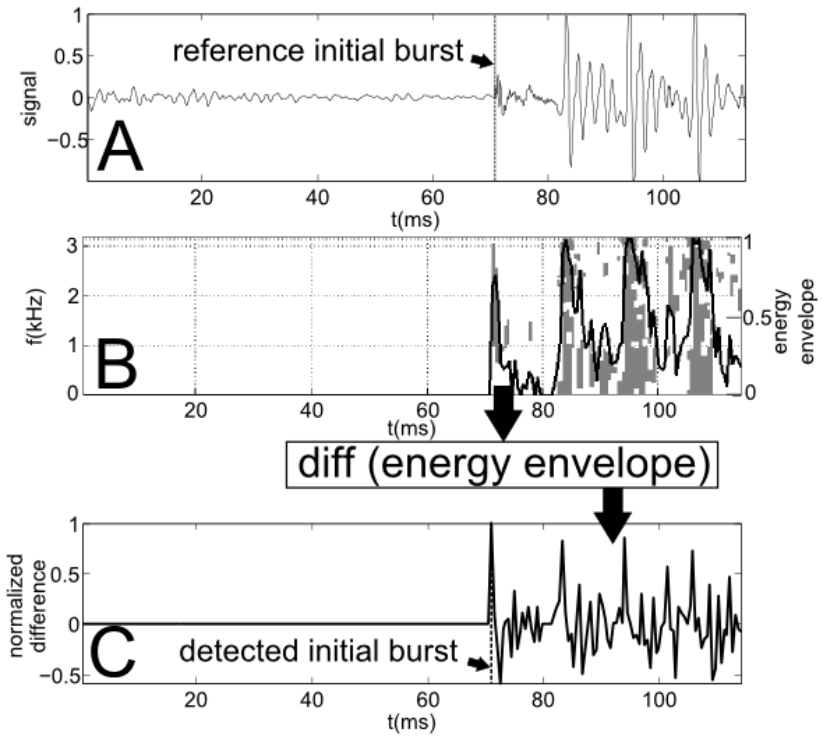

Figure 4.5: Front part of a syllable in the time domain (A), filtered spectrogram with the gray color denoting 1 and the white color denoting 0 and its marked energy envelope (B), and the normalized difference of the energy envelope used for the final initial burst detection (C).

4.3.4.4 Detection of Vowel Onset

The quasi-periodic character of a vowel with an abrupt onset of energy was detected using the Bayesian Step Changepoint Detector (BSCD) (Ruanaidh and Fitzgerald, 1996, Cmejla et al., 2013). In general, the BSCD assumes that (i) the signal is composed of two different constant values (e.g., 0.05 and 0.3 marked as lines in Figure 4.6 (B)), and (ii) that it is possible to calculate the posteriori probability of

changes in the signal through Bayesian marginalization. Whereas the approach with the $P_{stop\_release}$ matrix emphasizes the abrupt noise burst, the assumption of signal being composed of two different constant steps emphasizes a boundary between two different signals. The input of the detector represents the first part of the syllable from the initial burst to the end of the front part of the signal (see Figure 4.4); this can be seen in Figure 4.6 (A), where the reference position of the vowel onset is highlighted. Subsequently, due to the differing character of consonants and vowels, we may assume that the position of vowel onset is located in one of several local maxima of the BSCD output. This output is depicted in Figure 4.6(C), where the detected position is marked. To detect the local maximum corresponding to vowel onset, we may assume that the entire consonant is longer than the distance between single glottal pulses. This presumption allows delineation of the local maximum, following the largest gap between two consecutive maxima, as the position of vowel onset.

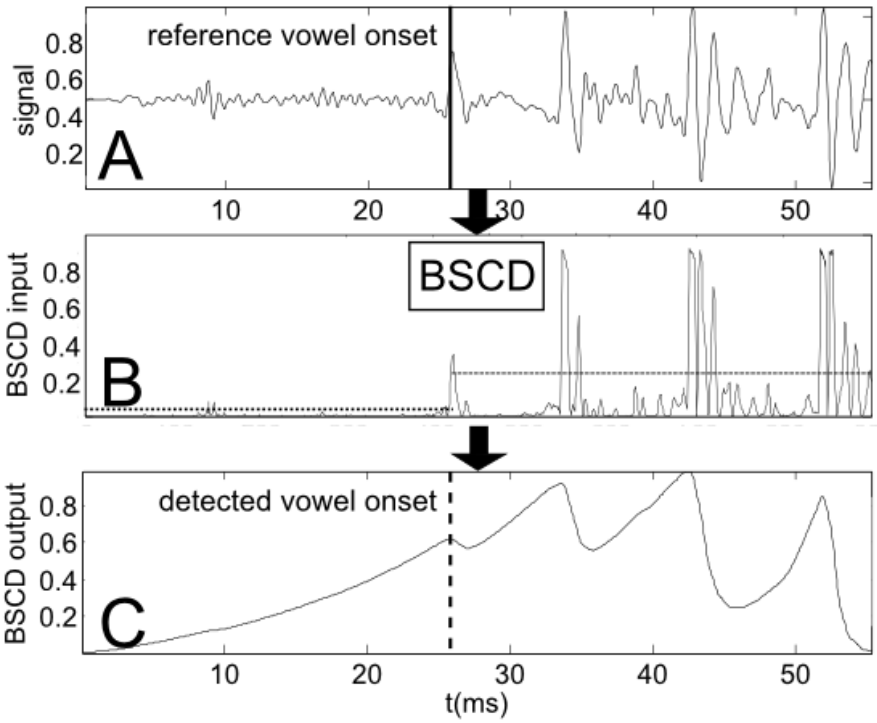

Figure 4.6: Original signal with marked reference position (A), input of the BSCD detector represented by the squared original signal and marked BSCD steps (B), and output of the BSCD detector with marked positions of the detected vowel onset position (C).

4.3.4.5 Detection of Occlusion

The position of occlusion is the most difficult to detect due to its slow subsidence and fuzzy borders. Due to decreased voice quality of PD speakers, a low-pass FIR filter with an order of one quarter of the signal length including the 1.5 kHz cut-off frequency was used. Contrary to the noise component, the F0, F1 and F2 provide a major contribution to signal energy in this frequency band. Signal energy was estimated from the filtered rear-part of the signal (see Figure 4.7 (A)) as the squared signal (see Figure 4.7 (B)). Subsequently, the flexible threshold was adjusted for occlusion detection. The threshold was given as an inverted polynomial energy approximation, and therefore the threshold was lowered with an increase in energy and vice versa, as illustrated by Figure 4.7(B). The definition of the threshold may be written as

$$T_o = \prod_{j=0}^{k} c_k x^k + 2\overline{E}, \tag{4}$$

where c_k denotes the k-th coefficient, $\overline{E}$ gives the mean value of energy, and k is the order of polynomial approximation. The order was experimentally set at nine, providing a good compromise between threshold elasticity and boundary fuzziness. The exact occlusion position was then marked as the place of the last intersection of energy and the threshold, which is no further than 20 ms from the preceding intersection. The 20 ms rule eliminates false detections connected with abrupt noises in distant parts of the signal.

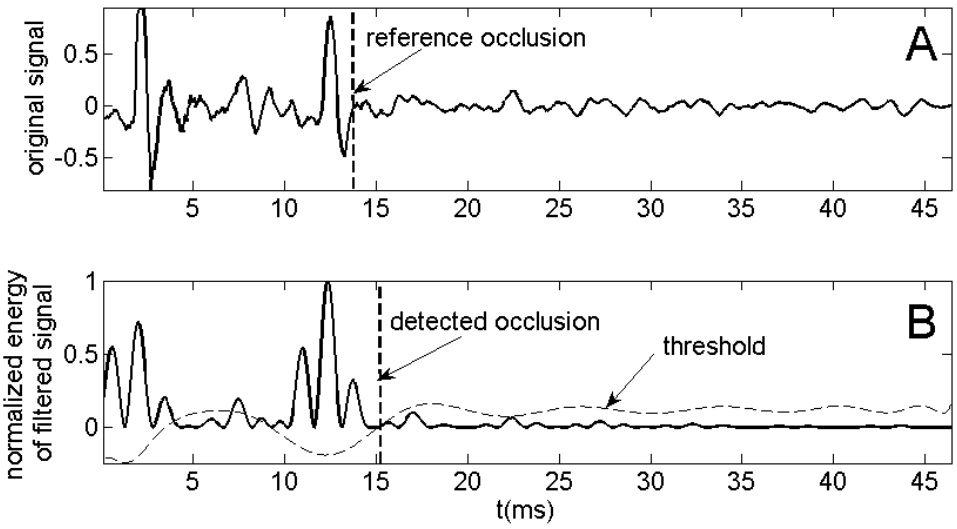

Figure 4.7: Rear part of a syllable with marked reference occlusion position (A), and energy of the filtered signal with polynomial threshold and marked occlusion position (B).

4.3.4.6 Algorithm enhancement

Even though the previously described algorithm showed sufficient robustness in hypokinetic utterances it is prone to high variability rate caused by forced inspirations and expiration present in HD utterances. Therefore to address this aspect of hyperkinetic speech an algorithm based on analysis of the linear prediction (LP) residual was used (Prasanna et al., 2009). This approach uses LP residual for the detection of voiced parts of the utterance. The LP residual was estimated from the signal which was down sampled to 8 kHz and filtered by 500 Hz fir filter with order of 100. Subsequently, the Hilbert envelope of LP residual was estimated and smoothened by moving average filter with order of 500 (Prasanna et al., 2009).

Positions of peaks in the smoothened envelope were set as the positions of the vowel nuclei. To detect these peaks, an envelope slope was computed using first-order difference and every positive to negative zero crossing was marked as the vowel nuclei. To eliminate false detections caused by low signal-to-noise ratio or by intensity fluctuations along single vowels, the minimal distance between two vowel nuclei was set as 10ms. When two peaks were found to be within 10ms distance a higher peak was selected as vowel nuclei see (Figure 4.8).

To detect the borders of syllables, the local minima of smoothened Hilbert envelope were detected. Subsequently, the each syllabic nucleus was associated with the nearest local minimum. According to the relative position of local minima to the nucleus, it was decided whether it is position of the syllable beginning or syllable end. Then the second border was chosen according to this decision.

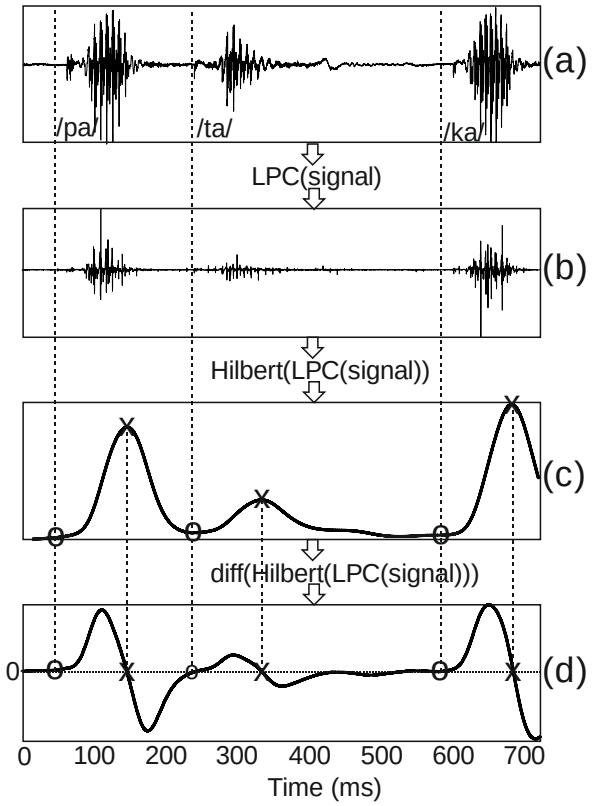

Figure 4.8: Detection of single syllable nuclei based on (Mahadeva Prasanna et al., 2009). Section (a) shows utterance in the time domain. Section (b) represents LP analysis residua and section (c) smoothened Hilbert envelope and section (d) illustrates slope of this envelope. With final positions of signal nuclei (x) and syllable borders (o) marked in Hilbert envelope (c) and its slope (d).

Signal within these borders was resampled to 20 kHz and analyzed to determine the positions of the initial burst and the vowel onset values for each syllable. Detection of initial burst was based on the approach described in the section 4.3.4.3 and the vowel onset detection was depicted in the section 4.3.4.4.

4.3.5 Articulatory Features

To evaluate the impact of PD related hypokinetic dysarthria on speaker performance, we propose 13 features representing six aspects of speech. The features describing *Voice Quality*, *Coordination of Laryngeal and Supralaryngeal Activity*, *Precision of Consonant Articulation*, *Tongue Movement*, *Occlusive Weakening*, and *Speech Timing* are listed in Table 2.1. Furthermore to provide quantitative acoustic vocal assessment of articulatory deficits related with, spastic or ataxic dysarthria (Duffy, 2013, Darley et al., 1969), which correspond to previous descriptions of speech and neuropathological findings in patients with PSP and MSA (Kluin et al., 1993, Kluin et al., 1996). We have also encompassed vowel length to extend speech timing dimension.

Due to the differing spectral characteristics of /p/, /t/, and /k/ consonants and their following vowels, features describing the precision of consonant articulation and tongue movement were performed on different types of syllables (bilabial /pa/, alveolar /ta/, and velar /ka/), separately. Moreover, the measurements connected with the coordination of laryngeal and supralaryngeal activity were performed for separate and mixed syllables. Therefore, the final number of measurements performed was 28. All of the measurements ranked each utterance with an average feature value computed from the first 5 syllabic trains (15 syllables overall). This approach helps to separate the involvement a of single speech feature from the impact of varying speech length.

4.3.5.1 Voice Quality

One of the muscle groups affected by PD is the group of laryngeal muscles. Distortion of this muscle group may lead to decreased vocal fold adduction and decreased ability to keep laryngeal muscles in a fixed position, which may result in increased jitter, shimmer, noise, distortion of F0 in general, and voice tremor (Goberman and Blomgren, 2008, Lindblom and Sundberg, 1971). To obtain general information about voice quality, two vowel similarity quotients and one vowel variability quotient were utilized. The vowel similarity quotient of the entire voicing (VSQ) and the vowel similarity quotient of the first 30 ms of voicing (VSQ_{30}) are defined as the first autocorrelation coefficients, and estimate the ability to produce a steady vocal tone. The motivation behind a 30 ms window in VSQ_{30} was based on a previous study on vowel articulation in PD (Sapir et al., 2010); in the present study, the 30 ms window represented the midpoint of the vowel that should manifest the greatest periodicity through the entire vowel duration. The vowel variability quotient (VVQ) is given as the standard deviation of vowel duration, which reflects the stability of the timing of vocal fold abduction and adduction.

4.3.5.2 Coordination of Laryngeal and Supralaryngeal Activity

The PD-induced disruption of movement patterns may lead to disturbances in muscle group coordination. To evaluate the impact of PD on the coordination of laryngeal and supralaryngeal muscle groups, the VOT and the VOT ratio were used. The VOT parameter, defined as the duration between stop release and the onset of voicing (Hansen et al., 2010), was motivated by the assumption that acoustic events, including the initial burst and vowel onset, are associated with articulatory gestures (i.e., the release of consonant constriction, the onset of vocal fold vibration; Hansen et al.,

2010). In addition, the VOT ratio, defined as VOT divided by the length of entire syllable, was estimated as the parameter suppressing the effect of speech rate (Fischer and Goberman, 2010).

4.3.5.3 Precision of Consonant Articulation

Effort to achieve a normal repetition rate may lead to reduced articulatory displacement. This reduced movement may manifest as airflow leaking around insufficiently closed articulators as well as decreased energy during the initial burst. To assess the impact of imprecise articulator setup, spectral characteristics describing a consonant spectral trend (CST) and a consonant spectral moment (CSM) were employed. The consonant spectral trend is computed as the slope of the line obtained using Fourier spectrum regression in a certain frequency interval. To emphasize the different spectral characteristics of /p/, /t/, and /k/ consonants, three different frequency bands were selected as: /p/ [2500, 3500] Hz; /t/ [2000, 3000] Hz; /k/ [1500, 2500] Hz (Hansen et al., 2010). The CSM represents the first spectral moment describing a centroid of energy contained in the entire Fourier spectrum of the consonant.

4.3.5.4 Tongue Movement

As one of the major articulators, the tongue has a crucial influence on the shape of the oral cavity and formant frequencies, and therefore, change of formant frequency behavior may reveal PD-induced disruption of tongue movement. In general, the acoustic-articulatory relationship can be easily understood, as the F1 frequency varies inversely with tongue height and the F2 frequency varies directly with tongue advancement (Kent et al., 2000b, Lindblom and Sundberg, 1971, Sapir et al., 2010). To assess tongue movement during vocalization, the first formant trend (1FT) and the second formant trend (2FT) were computed as the angle of the linear regression line of F1 or F2 tracked in the vowel.

4.3.5.5 Occlusion Weakening

Reduced articulatory movements may also be present during the silent gap between two syllables. Reduced movements may lead to the leakage of turbulent airflow, which results in increased noise during the silent gap (Duez, 2007). To describe the noise contained in the silent gap, the signal-to-noise ratio (SNR) defined a according to equation (5)

$$SNR = 10\log_{10}\frac{P_S}{P_N} \qquad (5)$$

where P_S represents power contained in voicing and P_N represents power obtained in the signal during the silent gap.

4.3.5.6 Speech Timing

Disrupted movement patterns do not only influence two particular muscle groups separately (e.g., coordination of laryngeal and supralaryngeal activity), but may also affect all aspects of speech timing. Therefore, five parameters were proposed to evaluate the impact of PD on speech timing. The first designed parameter investigates the overall DDK speech rate (DDK rate). The DDK rate is defined as the number of syllables per second and is computed as the number of initial bursts across the entire utterance. The second parameter estimates the ratio of silent gaps during the DDK task (DDK pace), and it is defined as the average value of silent gaps obtained in each utterance. The DDK pace, in connection with the DDK rate, provides information about the speech-silence duration ratio. The third parameter reflects the subject's ability to maintain a steady rhythm during the DDK speech task (DDK fluctuation), and is computed as the standard deviation of the duration of silent gaps in an utterance. The fourth the vowel duration defines ratio of voiced segment length related towards the length of entire syllable. And the fifth parameter describes steadiness of speech rhythm defined as standard deviation of distances between single syllables.

Table 4.3: Definitions of Articulatory Features

Name	Defined in interval	Definition
Voice Quality		
VSQ	vowel onset to occlusion	Vowel similarity quotient, the autocorrelation of the entire vowel duration, representing the rate of regularity of the vowel
VSQ_{30}	first 30 ms after vowel onset	VSQ of the first 30 ms of the vowel, representing the rate of the regularity of the vowel beginning
VVQ	vowel onset to occlusion	Vowel variability quotient, the level of variability in vowel length
Coordination of Laryngeal and Supralayngeal Activity		
VOT	initial burst to vowel onset	Voice onset time defining the length of the entire consonant
VOT ratio	initial burst to vowel onset	The voice onset time ratio defining the length of the entire consonant relative to syllable length
Precision of Consonant Articulation		
CST	initial burst to vowel onset	Consonant spectral trend, the regression of consonant spectrum computed in defined intervals
CSM	initial burst to vowel onset	Consonant spectral moment, the first spectral moment of the consonant
Tongue Movement		
1FT	vowel onset to occlusion	First formant trend, regression of the first formant frequency
2FT	vowel onset to occlusion	Second formant trend, regression of the second formant frequency
Occlusion Weakening		
SNR	vowel onset to occlusion (harmonic signal) as compared to occlusion to subsequent initial burst (noise signal)	Signal-to-noise ratio, representing the amplitude of total noise component
Speech Timing		
DDK Rate	entire utterance	Diadochokinetic rate, the number of syllables per second
DDK pace	occlusion to subsequent initial burst	Diadochokinetic pace, the mean length of silent gaps between syllables
DDK fluctuation	occlusion to subsequent initial burst	Diadochokinetic instability, the level of instability in silent gaps between syllables
Vowel duration	Vowel onset to occlusion	The ratio of vowel length to length of entire syllable
DDK regularity	entire utterance	Variability in rhythm of speech production

4.3.6 Statistics

Statistical analyses were performed in three separate parts: algorithm performance evaluation for automatic segmentation of an utterance, the evaluation of group differences across articulatory features estimated from segmented utterances, evaluation of the classification experiment based on previously computed articulatory features. Although these three parts are interconnected, the evaluation of each was performed separately, i.e., single syllables were used in the evaluation of algorithm performance, average performances of each participant for group difference estimation, and single utterances for the classification task (two per subject).

4.3.6.1 Algorithm Performance

Algorithm performance is illustrated by the cumulative distributions of absolute differences between reference-manual labels and automatically detected positions. For each syllable's event (i.e., initial burst, vowel onset, occlusion), three cumulative distributions were computed. The first was based on all 1644 tokens (across both PD_U and HC groups). Two other distributions were based on PD_U or HC tokens separately (753 tokens for PD_U and 891 for HC). Furthermore, to compare the performance of our algorithm with previous results, a method based on the teager energy operator (TEO) published by Hansen et al. was implemented (Hansen et al., 2010). This approach uses the amplitude modulation component (AMC), which is derived from the TEO, to detect the initial burst and vowel onset in single words. The TEO-based algorithm is not designed for the detection of occlusion. The AMC was applied on the filtered signal, whereas the parameters of the filter were set according to the event (i.e., initial burst or voice onset), and also according to the type of consonant (i.e., /p/, /t/, /k/) when considering burst. The TEO-based algorithm was used to detect the initial burst and voice onset in our data and the cumulative distributions of absolute differences for all PD_U and HC syllables.

To compare the performance of the original and the extended algorithm, the cumulative distributions were computed using 4211 tokens including 753 tokens from the PD_U group and 3458 tokens from the HD group. Additionally, the 10 ms threshold was chosen as the representative threshold value with respect to previous section and study (Stouten and Van Hamme, 2009). The cumulative distributions were computed using all syllables contained in the PD_U or HD groups and falsely detected or missed

syllables were always set as erroneous. Cumulative distributions were estimated separately for PD_U and HD participants.

4.3.6.2 Group Differences and Relationships between Metrics

For assessment of group differences, the average feature values were calculated for each participant prior to analyses. As the one-sample Kolmogorov-Smirnov test (D = 0.08 to 0.20, $p > 0.05$) showed that articulatory features were normally distributed, the two-sample *t*-test was used to assess group differences between PD_U and HC groups and HD and HC groups. The group differences between PD_{SDD}, PSP and MSA groups were computed using the Kruskal-Wallis test. Cohen's effect size (ES) was additionally calculated to assess the strength of differences for PD_U vs. HC, HD vs. HC, PD_{SDD} vs. PSP, PD_{SDD} vs. MSA and MSA vs. PSP groups, with Cohen's d, with $d > 0.5$ indicating a medium effect and $d > 0.8$ indicating a large effect.. Finally, the Pearson correlation coefficient was used to evaluate the correlation between results obtained by automatic detection and reference values, as well as the extent to which single measurements were correlated.

4.3.6.3 Classification Experiment

The experiment based on the support vector machine (SVM) classifier was performed using all PD_U and HC utterances (two per subject) in order to obtain more robust classifier estimates, i.e., the utterances provided by the same participant were not averaged as in the evaluation of group differences. The aim of the experiment was to separate two classes of PD_U and HC participants, based on automatically extracted articulatory features, which were pre-selected using Pearson's correlation and distance correlation. Being linearly inseparable, the features had to be mapped to the space with higher dimensionality, where the linear separability was achieved. For this purpose a Gaussian radial basis function (RBF) kernel was used. The RBF is defined as

$$\vec{K}(z, z') = \exp(-\gamma \|z - z'\|^2), \tag{6}$$

where $\|z - z'\|$ is Euclidean distance of the input vectors and the kernel parameter γ is used to set width of Gaussians approximating the decision boundary. The SVM model may be then written as

$$sign\left(\sum_{\alpha_n > 0} \alpha_n y_n \vec{K}(z, z') + \beta\right) \tag{7}$$

where z and z' are vectors of input features, y_n are labels of data used for training and α_n are Lagrange multipliers based on Lagrange formulation of the optimization task. To prevent overfitting the penalty coefficient C was used to constrain the maximal value of Lagrange multipliers.

The determination of the optimal parameter C and was performed using a grid search over the sets $C = [2^{-15}, 2^{-13}, \ldots, 2^{15}]$ and $\gamma = [2^{-15}, 2^{-13}, \ldots, 2^{3}]$ (Tsanas et al., 2012). Once the optimal parameters C and were found the classifier was trained and tested using these values.

To validate the generalization, empirical findings of previous studies suggest cross-validation or bootstrap methods as the most reliable (Tsanas et al., 2012, Hastie et al., 2009). For the purposes of the generalization estimation the standard cross-validation splitting entire dataset (80 utterances) to the training set containing only 60% of the data (48 utterances) and the testing set containing 40% of all recordings (32 utterances) was employed. For the purposes of the cross-validation a total number of 20 repetitions were performed, with random permutation of the data prior to splitting into training and test subsets. Furthermore, leave-one-subject- out (LOSO) cross-validation, excluding all utterances of the subject used for testing, was utilized and run throughout the entire data.

The testing error was estimated during each iteration of both cross-validations (Hsu et al., 2010). Subsequently, the errors were averaged over all repetitions and the overall performance was determined as the average percentage of correctly classified utterances. Furthermore, the true positive (number of correctly classified PD participants) and true negative (number of correctly classified HC participants) classification performances were assessed.

4.4 Velopharyngeal Control

4.4.1 Subjects

A total of 136 Czech native speakers, including 37 HD patients, 37 PD patients, 12 PSP patients, 13 MSA patients and and 37 healthy participants were recorded.

The HD group consisted of 19 men and 18 women with genetically confirmed HD with mean age 49.1 ± SD 12.7 years, mean disease duration 6.1 ± SD 3.4 years, mean number of CAG triplets 44.7 ± SD 3.3. Most of the patients (32/37) were treated with monotherapy or a combination of benzodiazepines, antipsychotics, amantadine and

antidepressants. All HD patients underwent extensive examination by an experienced neurologist and were rated according to the Unified Huntington's Disease Rating Scale. In addition, several UHDRS subscores including rigidity, bradykinesia, dystonia and chorea were assessed separately for subsequent analysis (Walker, 2007, Rusz et al., 2013c). The UHDRS motor score was 25.7 ± SD 12.2 and the UHDRS speech item was 0.8 ± SD 0.5.

The PD group consisted of 23 men and 14 women, mean age 63.1 ± SD 14.0 years, mean disease duration 8.0 ± SD 4.8 years. All PD patients fulfilled the diagnostic criteria for PD (Hughes et al., 1992). All participants were on stable dopaminergic medication for at least 4 weeks before the examinations, which were conducted in the on-medication state. All PD patients underwent neurological examinations by an experienced neurologist and were rated according to the Hoehn & Yahr staging scale and motor Unified Parkinson's Disease Rating Scale. In addition, several UPDRS composite subscores including rigidity, bradykinesia, resting tremor and axial score were estimated for subsequent analysis (Hughes et al., 1992, Jankovic, 2008). The H&Y score was 2.1 ± SD 0.4, UPDRS III score 17.5 ± SD 8.2, and the UPDRS III speech item 18 score was 0.8 ± SD 0.6.

The PSP participants (10 men 2 women) comprised 9 participants diagnosed with PSP-RS, 2 participants were diagnosed with PSP-P and one with PSP-PAGF. The age of PSP group was 65.8 ± SD 5.4 years with the symptom duration 3.8 ± SD 1.4 years. The H&Y score 3.3 ± SD 0.8 and the NNIPS score of PSP group was 66.3 ± SD 28.7 and the UPDRS III item 18 was 2.0 ± SD 1.0.

The MSA group (6 men and 7 women) included 10 participants diagnosed with MSA-P and 3 participants diagnosed with MSA-C. The age of MSA group was 60.8 ± SD 4.9 years with symptom duration 3.68 ± SD 1.3 years. The H&Y score 3.6 ± SD 0.7 and the NNIPS score of MSA group was 78.5 ± SD 19.9 and the UPDRS III item 18 was 2.0 ± SD 0.7.

The healthy control group consisted of 23 men and 14 women, mean age of 63.1 ± SD 8.7 years. None of the HC participants had a history of neurological or speech disorder.

None of the HD, PD, MSA, PSP or HC subjects suffered from chronic obstructive pulmonary disease, respiratory tract infection, allergy, asthma, facial paresis, or other malady that could negatively influence participant speech performance. All the patient characteristics are summarized in Table 4.4

.

Table 4.4: Clinical characteristics of participants

	HC mean/SD	PD mean/SD	HD mean/SD	PSP mean/SD	MSA mean/SD
Age	63.1/8.7	63.1/14.0	49.1/12.7	60.8/6.5	65.8/5.4
Symptoms duration		8.0/4.8	6.1/3.4	3.6/1.3	3.8/1.4
CAG triplets			44.7/3.3		
L-dopa equivalent		0.0/0.0		899/394	800/373
Amantadine				300/89	200/107
NNIPPS				78.5/19.9	66.3/28.7
UPDRS III		17.5/8.2			
UPDRS III speech item 18		0.8/0.6		2.0/0.7	2.0/0.1
H&Y		2.1/0.4		3.6/0.7	3.3/0.8
UHDRS			25.7/12.2		
UHDRS speech item			0.8/0.5		

4.4.2 Perceptual Analysis

As connected speech is more demanding for velopharyngeal control, it is considered the most valid task for perceptual nasality estimation (Kuehn and Moller, 2000, Peterson-Falzone et al., 2001). The nasality rating was therefore based upon HC, PD and HD group monologue perception and performed by 10 raters including one speech-language pathologist, three clinicians and six acoustic speech specialists using a graded scale (0 = normal nasality, 1 = mild hypernasality, 2 = moderate hypernasality, 3 = severe hypernasality), based on The Great Ormond Street Speech Assessment ’98 (GOS.SP.ASS.’98; Sell et al., 1999). The perceptual assessment was performed blindly on randomized data consisting of HC, PD and HD participant groups. The presentation of samples was self-paced and performed by each rater separately, and each speech sample could be repeated at the discretion of the listener. The final score was obtained by the median value computed from all perceptual assessments. Intra-rater reliability was based upon two perceptual assessments performed with three months delay by one rater.

4.4.3 Acoustic Analysis

For the purposes of instrumental analysis, two recording parts equal to 10% of signal length were cut off from both the beginning and end of the vowel /i/ to avoid distortion by initial vocal fold adjustment and fatigue at the end of the utterance. The remaining signal was then resampled to 20 kHz, which lowered the computational complexity and preserved all useful information. The preprocessed signal was divided

using a hamming 60 ms window with 55 ms overlap. Subsequently, each window was analyzed using a 1/3-octave spectra method.

The process of 1/3-octave spectra analysis based on the multirate filter bank is illustrated in Figure 4.9. The three highest 1/3-octave frequency band filters were designed according to this method. For our purposes, the 3rd order IIR butterworth filters were used and centered on octave frequencies of 2500 Hz [passband from 2244.9 Hz to 2828.4 Hz], 3150 Hz [passband from 2828.4 Hz to 3563.6 Hz], and 4000 Hz [passband from 3563.6 Hz to 4489.8 Hz]. After filtering, the highest components were removed from recording and the signal was then down-sampled by a factor of 2, i.e., sampling frequency (fs) to fs/2. Being defined in relation to the fs, the filter characteristics related to fs/2 yielded one octave lower for each down sampling. Based on this approach, the entire filter bank was achieved by the iterative use of signal down sampling. In each 1/3-octave frequency band, the root-mean-square (RMS) energy was estimated and achieved energy was transformed into decibels. A sum of energy contained in the entire 1/3-octave spectra was used as a reference value for the transformation into decibels, as described by equation 8.

$$E(i) = 10\log_{10}\left(\frac{E_{filtered}(i)}{\sum_{k=1}^{18} E_{filtered}(k)}\right), \tag{8}$$

where $E_{filtered}$ is energy contained in the single band of 1/3-octave and E(i) is the decibel value of energy contained in the i-th band.

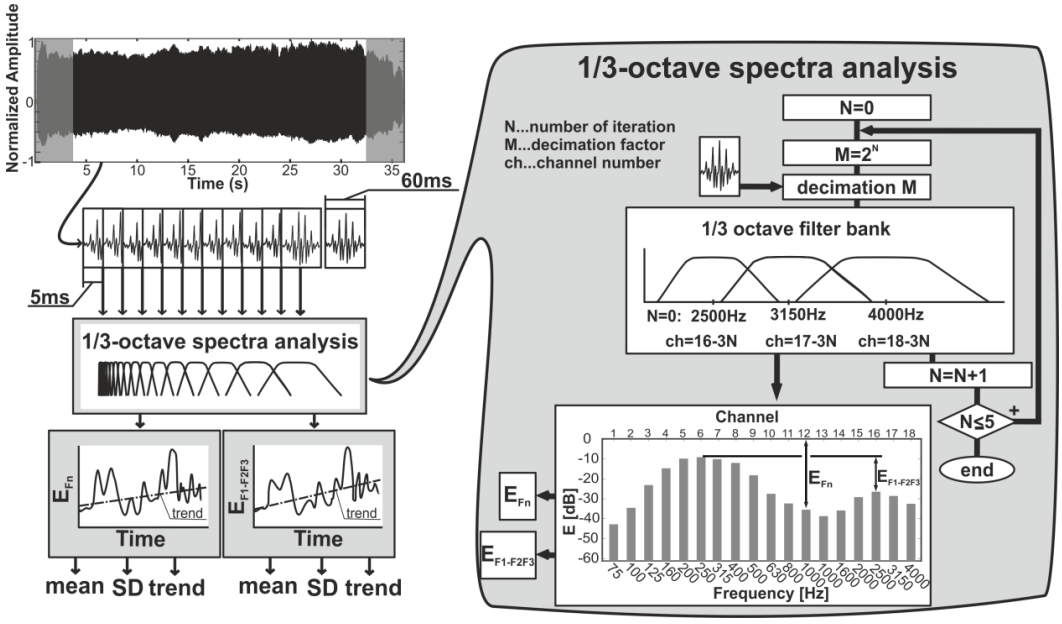

Figure 4.9: Principle of acoustic analysis based on 1/3-octave spectra assessment.

Considering the effect of spectral flattening, nasality in sustained phonation of the vowel /i/ was evaluated using the E_{Fn} parameter, which represented energy in a 1/3-octave band centered around 1 kHz [passband from 890.9 Hz to 1122.5 Hz]. This parameter reflected the addition of nasal resonance and additive nasal pole to the transfer function at 1 kHz. The overall level of hypernasality was estimated by the mean value of E_{Fn} parameter (E_{Fn} mean) across all windows in the entire utterance. The variability of nasality (E_{Fn} SD) in speech was evaluated as the standard deviation of each parameter across the entire utterance. Finally, the evolution of hypernasality in the course of the utterance (E_{Fn} trend) was described using a linear regression tangent for each parameter.

4.4.4 Statistics

As the vowel /i/ was recorded twice for all speakers, average values of both estimated acoustic parameters were used for all consecutive analysis.

The Kolmogorov-Smirnov test for independent samples was used to evaluate normality. Analysis of variance (ANOVA) with post-hoc Bonferroni adjustment was used for the estimation of group differences between HC, HD and PD groups across acoustic variables, the assessment of differences between PD, MSA and PSP groups was performed using Kruskal-Wallis with post-hoc Bonferroni adjustment. Furthermore, effect sizes for proposed parameters were evaluated using Cohen's *d* for differences between the HD and HC, PD and HC, PD and MSA, PD and PSP groups as well as between the PSP and MSA groups, with $d > 0.5$ indicating a medium a $d > 0.8$ indicating a large effect.

Relationships between variables were evaluated using Pearson's correlation, Spearman's correlation and the intraclass correlation coefficient (ICC). Pearson's correlation was applied to normally distributed data (speech metrics and disease severity scores), whereas Spearman's correlation was used for non-normally distributed data (perceptual assessment). The ICC was employed for the evaluation of inter-rater and intra-rater reliability of perceptual scores. Due to the explorative nature of the current study, adjustment for multiple comparisons with regard to correlations was not performed and the level of significance was set to $p < 0.05$.

Assessment of the percentage of affected participants from acoustic data was based on the Wald task, which is a non-Bayesian statistical decision-making method (Schlesinger and Hlavac, 2002). This method allows predefined false positives and

false negatives by extending two basic classes (i.e., healthy and hypernasal), by an indecisive class. Use of the indecisive class enables set boundaries where the possibility of a false positive or false negative result reaches an acceptable value. Therefore the indecisive class is used in cases where measured data do not provide sufficient information for safe classification. As a result, the method provides optimal cut-off values indicating if the subject already reached hypernasal speech performance or manifest normal nasality of wider norm of healthy speakers. Comprehensive details on the Wald task have been published previously (Rusz et al., 2011b).

5 RESULTS

5.1 Rapid articulatory moves

5.1.1 Algorithm performance

Figure 5.1 shows the cumulative distribution representing the absolute difference between reference manual labels and automatically detected positions for the initial burst, vowel onset and occlusion, where the left column represents the performance results of our algorithm and the right column the performance of the TEO-based algorithm proposed by Hansen et al. (2010). Considering a 5ms threshold of absolute difference, performance of our algorithm for all syllables was 79.2% for the initial burst, 81.7% for vowel onset and 59.2% for occlusion. The detection performance for occlusion increased to 77.3% at a 10ms threshold.

Considering a 5ms threshold for initial burst of HC group, our approach achieved 85.4% in comparison to TEO-based algorithm with 65.1% accuracy (Hansen et al., 2010). In the case of vowel onset, our approach reached 86.7% compared to 76.2% by the TEO-based algorithm. In PD group, our approach achieved a score of 71.9% in comparison to 45.2% by the TEO-based algorithm for the initial burst and 5ms threshold. Similarly, we reached a performance of 75.8% in comparison to 45.6% by the TEO-based algorithm in the detection of vowel onset. Although the TEO-based approach achieved high and even comparable performances in HC group, its accuracy was relative low in PD group due to overall decreased speech quality.

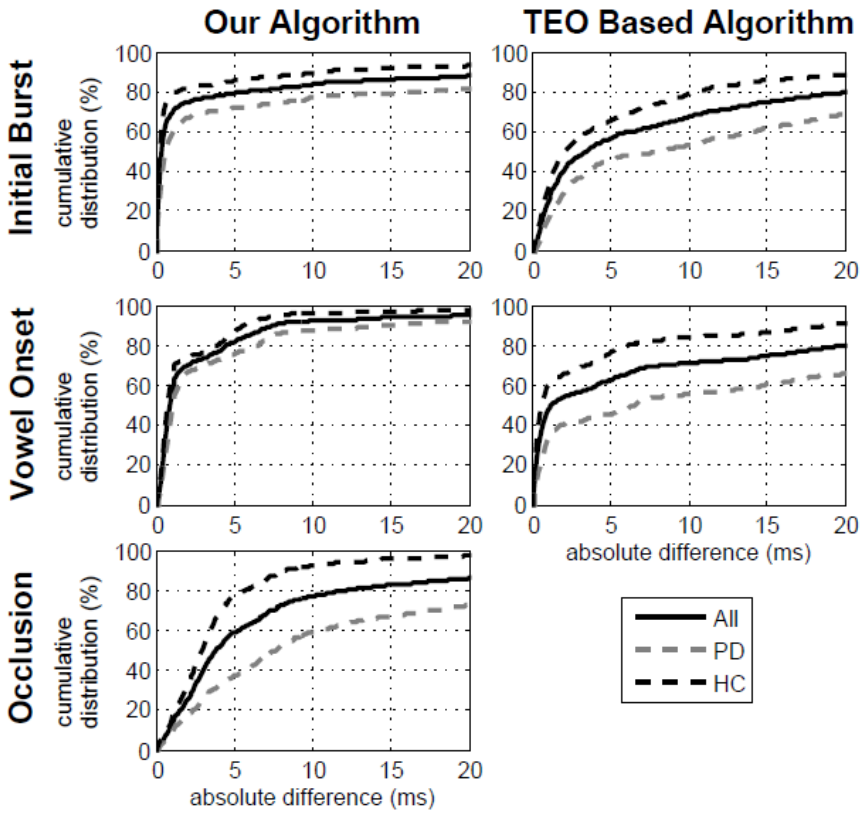

Figure 5.1: Cumulative distributions of algorithm performance based on the absolute difference between automatic detection and reference labels. Performances are estimated separately for syllables in Parkinson's disease subjects (PD) and healthy controls (HC), as well as for all syllables together (All). The first column shows the performance of the algorithm presented in this study and the second column illustrates the performance achieved by the TEO-based algorithm.

Figure 5.2 illustrates performance of updated algorithm in a comparison to the performance of original algorithm designed for assessment of PD. The VOT boundaries detection is illustrated by solid lines in the Fig. 19, which shows the algorithm performance using cumulative distributions of absolute differences between detected and reference positions. For the purposes of comparison, the dashed lines in Fig. 19 represent performance of previously designed PD-aimed algorithm. Considering 10ms threshold and PD sufferers, the presented algorithm achieved slightly improved score of 81.5 % for the initial burst and 89.5 % for the vowel onset, when compared to score of 78.2 % for the initial burst and 88.6 % for the vowel onset achieved by previously PD-aimed algorithm. Nevertheless, considering 10ms threshold and HD sufferers, the presented algorithm score of 77.8% for the initial burst and 80.1% for the vowel onset significantly outperforms score of 45.8% for the initial burst and 55.1% for the vowel onset achieved by the previous PD-aimed algorithm.

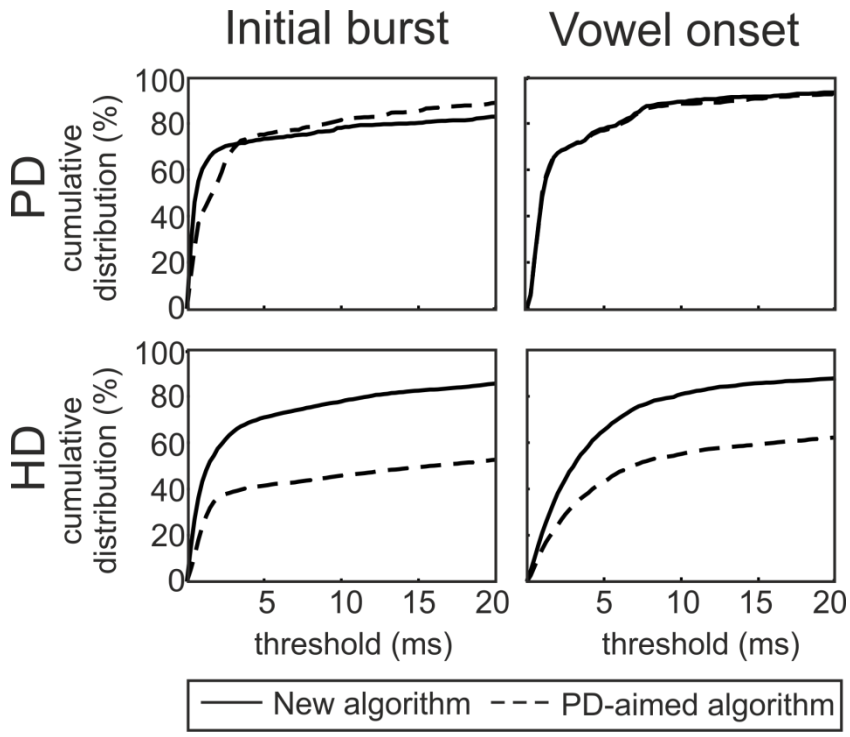

Figure 5.2: Cumulative distributions of absolute difference between detected and reference values using the enhanced algorithm. The results obtained for the Initial burst (1st column) and Vowel onset (2nd column) and for PD (1st row) and HD (2nd row) speakers. For the purposes of comparison the dashed line represent results of PD-aimed algorithm.

5.1.2 Group differences and relationships between metrics

The characteristics of each measurement, including the mean and standard deviation of values and effect sizes are listed in Table 5.1 for HC, PD_U and HD groups and Table 5.2 for PD_{SDD}, MSA and PSP groups. Significant differences between PD_U and HC, and HD and HC performances were found in each feature group. Differences between APS and PD_{SSD} were found in features describing voice quality, tongue movements and speech timing. The correlations between features based on automatic detection and manual reference labels showed high reliability ($r = 0.70$ to 0.99, $p < 0.001$) for all features except for those based upon precision of consonant articulation which showed moderate reliability ($r = 0.40$ to 0.69, $p < 0.001$).

5.1.2.1 Group differences between PD, HD and HC metrics

In the voice quality dimension, the VVQ was significantly increased PD_U patients when compared to controls ($t\,(44) = -3.13$, $p = 0.003$). Similarly, the VSQ_{30} was decreased in PD_U patients ($t\,(44) = 2.42$, $p = 0.02$). As for the HD vs. HC comparison both VVQ ($t\,(57) = -2.75$, $p = 0.008$) and VSQ_{30} ($t\,(57) = -4.59$,

$p < 0.0001$) were significantly decreased. In the dimension considering the coordination of laryngeal and supralaryngeal articulators, both VOT (e.g. VOT:all $t\,(44) = -7.54$, $p = 0.003$) and VOT ratio (e.g. VOT ratio: all $t(44) = -3.57$, $p = 0.003$) features reflected a considerable increase for PD_U participants, with VOT generally providing superior results to VOT ratio as demonstrated by effect sizes. In the HD group, however, only VOT (e.g. VOT:all $t\,(57) = -4.70$, $p < 0.0001$)showed significant increase. Considering the disrupted precision of consonant articulation, a significant difference in CST between the HC and PD_U groups for /pa/ ($t\,(44) = -2.48$, $p < 0.02$) and /ka/ ($t\,(44) = -4.54$, $p < 0.0001$) syllables was observed, whereas only a trend was detected for /ta/ ($t\,(44) = -1.7899$, $p = 0.08$). However, we found no significant group differences for CSM extracted through various consonants. Contrary to $PD_{U,}$ the HD group differed significantly for all CSM (e.g. CSM:/pa/ $t\,(57) = -3.25$, $p = 0.002$) and two CST features (CST:/pa/ $t\,(57) = -2.41$, $p = 0.02$; CST:/ka/ $t\,(57) = -4.27$, $p < 0.0001$). In the tongue movement dimension, all the 1FTs for /pa/ ($t\,(44) = 3.88$, $p = 0.0004$), /ta/ ($t\,(44) = 3.61$, $p = 0.0008$) and /ka/ ($t\,(44) = 3.75$, $p = 0.0006$) syllables were significantly different between the PD_U and HC groups as well as between HD and HC groups (1FT:/pa/ $t\,(57) = 2.87$, $p = 0.006$; 1FT:/ta/ $t\,(57) = 3.63$, $p = 0.0006$; 1FT:/ka/ $t\,(57) = 2.93$, $p = 0.005$). In contrast, only 2FT for the /ta/ syllable ($t\,(44) = 3.72$, $p = 0.0006$) was found to be impaired in PD_U patients and the /ta/ ($t\,(57) = 3.01$, $p = 0.004$) and /ka/ ($t\,(57) = -5.28$, $p < 0.0001$) syllables were impaired in HD group. Lower SNR compared to HC in the PD_U group and even more in HD group provided significant distinction (PD_U: $t\,(44) = 2.05$, $p = 0.047$; HD: $t\,(57) = 3.92$, $p = 0.0002$) in the occlusive weakening dimension. Finally, the speech timing dimension exhibited a considerable decrease in the DDK rate ($t\,(44) = 4.45$, $p < 0.0001$), and increase in DDK fluctuation ($t\,(44) = -2.78$, $p = 0.0082$) in the PD_U group. Furthermore, t HD showed strong distortion in all speech timing features excluding the vowel duration (DDK Rate $t\,(57) = 8.04$, $p < 0.0001$); DDK pace: $t\,(57) = -3.95$, $p = 0.0002$); DDK fluctuation $t\,(57) = -3.43$, $p = 0.001$); DDK regularity $t\,(57) = -3.95$, $p = 0.0002$)).

Table 5.1: Overview of Results of HC, PDU and HD Groups

#	Feature	HC	PD	HD	Cohen's d†	
		Mean ± SD	Mean ± SD	Mean ± SD	PD vs. HC	HD vs. HC
	Voice Quality					
1	VSQ (-)	0.45 ± 0.10	0.41 ± 0.33	0.41 ± 0.13	0.33	0.33
2	VSQ_{30} (-)	0.45 ± 0.11	0.37 ± 0.11	0.55 ±0.17	0.74*	0.79**
3	VVQ (ms)	14.51 ± 4.59	11.49 ± 4.62	32.49 ± 20.77	0.96**	1.41***
	Coordination of Laryngeal and Supralayngeal Activity					
4	VOT:all (ms)	20.33 ± 6.14	34.50 ± 6.17	25.58 ± 5.92	-2.30***	-1.29***
5	VOT:/pa/ (ms)	14.08 ± 4.66	26.57 ± 6.15	20.39 ± 6.63	-2.30***	-1.32***
6	VOT:/ta/ (ms)	22.21 ± 7.91	36.42 ± 10.33	25.87 ± 7.84	-1.54***	-1.17***
7	VOT:/ka/ (ms)	24.73 ± 8.39	40.49 ± 7.05	30.52 ± 8.75	-2.03***	-0.78**
8	VOT ratio:all (%)	28.32 ± 6.51	35.43 ± 6.57	25.50 ±7.14	-1.08***	0.15
9	VOT ratio:/pa/ (%)	22.40 ± 5.77	30.84 ± 8.20	22.96 ±7.03	-1.19***	-0.22
10	VOT ratio:/ta/ (%)	29.90 ± 7.98	36.41 ± 7.81	26.01 ±8.12	-0.83**	0.06
11	VOT ratio:/ka/ (%)	32.65 ± 8.08	39.04 ± 6.97	27.53 ± 10.22	-0.85**	0.43
	Precision of Consonant Articulation					
12	CST:/pa/ (rad×10^{-9})	-3.23 ± 1.30	-2.25 ± 1.31	-1.12 ± 1.68	-0.76*	-0.61*
13	CST:/ta/ (rad×10^{-9})	-2.89 ± 2.02	-2.00 ± 1.15	-1.61 ±2.70	-0.54	-0.35
14	CST:/ka/ (rad×10^{-9})	-4.29 ± 1.91	-2.02 ± 1.34	-1.35 ±8.78	-1.38***	-1.12***
15	CSM:/pa/ (kHz)	4.93 ± 0.38	4.98 ± 0.47	6.80 ±1.11	-0.11	-0.86**
16	CSM:/ta/ (kHz)	5.00 ± 0.61	5.42 ± 1.01	8.41 ±1.59	-0.50	-1.29***
17	CSM:/ka/ (kHz)	4.81 ± 0.41	4.87 ± 0.49	7.53 ±1.41	-0.13	-1.26***
	Tongue Movement					
18	1FT:/pa/ (rad)	0.02 ± 0.11	-0.13 ± 0.13	-0.14 ±0.17	1.19***	0.76**
19	1FT:/ta/ (rad)	0.03 ± 0.14	-0.11 ± 0.13	-0.02 ±0.19	1.10**	0.97***
20	1FT:/ka/ (rad)	0.14 ± 0.14	-0.02 ± 0.13	-0.04 ±0.29	1.14**	0.81**
21	2FT:/pa/ (rad)	-0.09 ± 0.26	-0.06 ± 0.25	0.01 ±0.36	-0.12	-0.31
22	2FT:/ta/ (rad)	0.55 ± 0.22	0.28 ± 0.26	0.01 ±0.45	1.14***	0.84**
23	2FT:/ka/ (rad)	-0.53 ± 0.21	-0.43 ± 0.21	0.61 ±0.40	-0.46	-1.47***
	Occlusion Weakening					
24	SNR (dB)	28.02 ± 4.16	25.13 ± 5.03	22.41 ±5.78	0.63*	1.12***
	Speech Timing					
25	DDK Rate (syll/s)	7.74 ± 0.65	6.69 ± 0.88	4.97 ±1.39	1.36***	2.38***
26	DDK pace (ms)	64.34 ± 11.26	58.29 ± 16.66	119.49 ± 58.56	0.42	-1.21***
27	DDK fluctuation (ms)	22.91 ± 14.67	14.28 ± 15.02	56.62 ±56.15	0.85**	-1.06**
28	Vowel duration (%)	33.37 ± 6.43	36.71 ± 6.32	37.14 ±6.43	-0.55	-0.07
29	DDK regularity(ms)	26.91 ±7.98	17.02 ±16.76	70.45 ±61.48	-0.75	1.21***

† Measurements reaching significance are denoted by asterisks: *) $p < 0.05$, **) $p < 0.01$, and ***) $p < 0.001$.

5.1.2.2 Group differences between PD, PSP and MSA metrics

The Kruskal-Wallis analysis revealed significant differences between PD_{SDD} and both MSA and PSP only in VSQ_{30} (H (2, 38) = 9.90, p = 0.007) and VVQ (H (2, 38) = 9.84, p = 0.01) features representing voice quality dimension and in DDK regularity (H (2, 38) = 6.47, p = 0.04) feature included in speech timing dimension. In

addition the PSP compared to PD_{SDD} differed significantly in the 2FT:/ta/ (H (2, 38) = 9.7, p = 0.008) and the Vowel duration features (H (2, 38) = 7.07, p = 0.03).

Table 5.2: Overview of Results of PD_{SDD}, MSA and PSP Groups

#	Feature	PD_{SDD} Mean ± SD	MSA Mean ± SD	PSP Mean ± SD	Cohen's d†		
					MSA vs. PD_{SDD}	PSP vs. PD_{SDD}	MSA vs. PSP
	Voice Quality						
1	VSQ (-)	0.52 ± 0.21	0.71 ± 0.21	0.60 ± 0.19	-0.89	-0.42	0.53
2	VSQ_{30} (-)	0.40 ± 0.11	0.56 ± 0.16	0.55 ±0.13	-1.13**	-1.25**	0.05
3	VVQ (ms)	17.03 ± 12.58	35.47 ± 24.68	71.08 ± 113.25	-0.94*	-0.67*	-0.43
	Coordination of Laryngeal and Supralayngeal Activity						
4	VOT:all (ms)	20.33 ± 7.64	33.90 ± 12.86	36.68 ± 13.36	-0.66	-0.91	-0.22
5	VOT:/pa/ (ms)	20.40 ± 6.76	28.39 ± 16.28	29.29 ± 13.57	-0.65	-0.84	-0.06
6	VOT:/ta/ (ms)	25.86 ± 10.00	35.49 ± 16.44	37.98 ± 16.05	-0.71	-0.91	-0.15
7	VOT:/ka/ (ms)	34.56 ± 10.15	37.84 ± 10.96	43.74 ± 14.45	-0.31	-0.74	-0.46
8	VOT ratio:all (%)	31.37 ± 7.16	27.95 ± 5.92	31.53 ± 10.53	0.52	-0.02	0.42
9	VOT ratio:/pa/ (%)	27.38 ± 8.91	25.81 ± 8.04	26.60 ± 9.45	0.19	0.08	0.09
10	VOT ratio:/ta/ (%)	30.53 ± 8.83	27.96 ± 7.33	33.82 ± 12.76	0.32	0.30	-0.56
11	VOT ratio:/ka/ (%)	36.19 ± 6.83	30.10 ± 5.85	33.89 ± 12.03	0.96	0.24	0.40
	Precision of Consonant Articulation						
12	CST:/pa/ (rad×10^{-9})	-0.75 ± 0.69	-0.50 ± 0.50	-0.70 ± 0.55	-0.41	-0.07	0.38
13	CST:/ta/ (rad×10^{-9})	-1.90 ± 1.27	-1.00 ± 0.93	-0.54 ± 1.52	-0.82	-0.98	0.37
14	CST:/ka/ (rad×10^{-9})	-9.71 ± 6.53	-7.88 ± 7.62	-5.19 ± 5.65	-0.25	-0.74	-0.40
15	CSM:/pa/ (kHz)	7.53 ± 0.79	6.95 ± 0.98	6.97 ± 0.97	0.65	0.64	-0.01
16	CSM:/ta/ (kHz)	8.49 ± 0.74	8.52 ± 1.22	8.01 ± 1.37	-0.03	0.43	0.39
17	CSM:/ka/ (kHz)	7.69 ± 0.96	7.38 ± 1.11	7.19 ± 5.56	0.30	0.64	0.22
	Tongue Movement						
18	1FT:/pa/ (rad)	-0.03 ± 0.02	-0.01 ± 0.02	-0.03 ± 0.04	-0.78	0.00	0.57
19	1FT:/ta/ (rad)	-0.01 ± 0.02	-0.01 ± 0.03	-0.03 ± 0.03	-0.31	0.39	0.64
20	1FT:/ka/ (rad)	-0.01 ± 0.3	0.00 ± 0.2	-0.02 ± 0.04	-0.39	0.44	0.75
21	2FT:/pa/ (rad)	0.01 ± 0.04	0.03 ± 0.04	0.00 ± 0.05	-0.46	0.08	0.50
22	2FT:/ta/ (rad)	0.07 ± 0.04	0.03 ± 0.03	0.01 ± 0.06	0.95	1.15**	0.55
23	2FT:/ka/ (rad)	-0.07 ± 0.04	-0.04 ± 0.04	-0.05 ± 0.05	-0.59	-0.39	0.14
	Occlusion Weakening						
24	SNR (dB)	25.67 ± 6.13	25.02 ± 5.90	23.61 ± 5.47	0.11	0.36	-0.25
	Speech Timing						
25	DDK Rate (syll/s)	6.82 ± 1.12	5.45 ± 1.32	5.72 ± 1.32	1.12	0.90	-0.20
26	DDK pace (ms)	78.85 ± 43.78	79.41 ± 27.38	86.59 ± 32.04	-0.02	-0.19	-0.22
27	DDK fluctuation (ms)	19.07 ± 16.38	36.05 ± 32.04	35.77 ± 34.45	-1.12	-0.90	0.20
28	Vowel duration (%)	37.90 ± 5.81	49.48 ± 10.8	46.41 ± 18.69	-0.61	-1.34*	0.20
29	DDK regularity(ms)	18.57 ± 8.82	43.73 ± 27.11	51.09 ± 40.30	-1.25*	-1.12*	-0.22

† Measurements reaching significance are denoted by asterisks: *) $p < 0.05$, **) $p < 0.01$, and ***) $p < 0.001$.

5.1.3 Classification experiment

The classification experiment was based on PD_U and HC groups. Considering relations between speech features, the Pearson's correlation revealed correlation higher than 0.9 between the VOT and VOT ratio measurements. Accordingly, distance correlation reached value higher than 0.8 only between VOT and VOT ratio measurements. Therefore, all 29 features were retained for the classification experiment. The most representative classification results are presented in Table 5.3, where the correct overall, true positive and true negative performance rates are listed. Interestingly, the best correct overall classification score of 87.1 ± 5.4 % obtained by standard cross-validation and 88.4 ± 26.4 % obtained by LOSO cross-validation was achieved for the combination of six parameters (VSQ_{30}, VOT:/pa/, CST:/ka/, 2FT:/ta/, SNR, DDK rate), each representing one different speech dimension. Figure 5.3 shows probability distributions for six representative features with the best classification accuracy estimated using the Gaussian kernel density method.

Table 5.3: Representative Classification Results

Feature set (number of measerements)	Correct overall (%)	True positive (%)	True negative (%)
Cross-validation based on 60% training set and 40% testing set			
VSQ30, VOT:/pa/, CST:/ka/, 2FT:/ta/, SNR, DDK rate (6)	87.1 ± 5.4	86.2 ± 9.6	88.0 ± 7.5
VOT:all, 2FT:/ta/, DDK rate (3)	85.2 ± 4.5	84.5 ± 9.3	86.3 ± 8.1
All easurements (27)	82.4 ± 7.0	91.4 ± 9.9	74.8 ± 10.3
VOT:all (1)	83.3 ± 5.4	87.8 ± 7.3	78.1 ± 11.5
Leave-one-subject-out cross-validation			
VSQ30, VOT:/pa/, CST:/ka/, 2FT:/ta/, SNR, DDK rate (6)	88.4 ± 26.4	86.4 ± 37.6	90.5 ± 20.1
VOT:all, 2FT:/ta/, DDK rate (3)	83.7 ± 28.3	81.8 ± 29.1	85.7 ± 28.0
All easurements (27)	82.6 ± 32.5	88.6 ± 21.4	76.2 ± 40.7
VOT:all (1)	79.1 ± 34.9	90.9 ± 25.1	66.7 ± 66.7

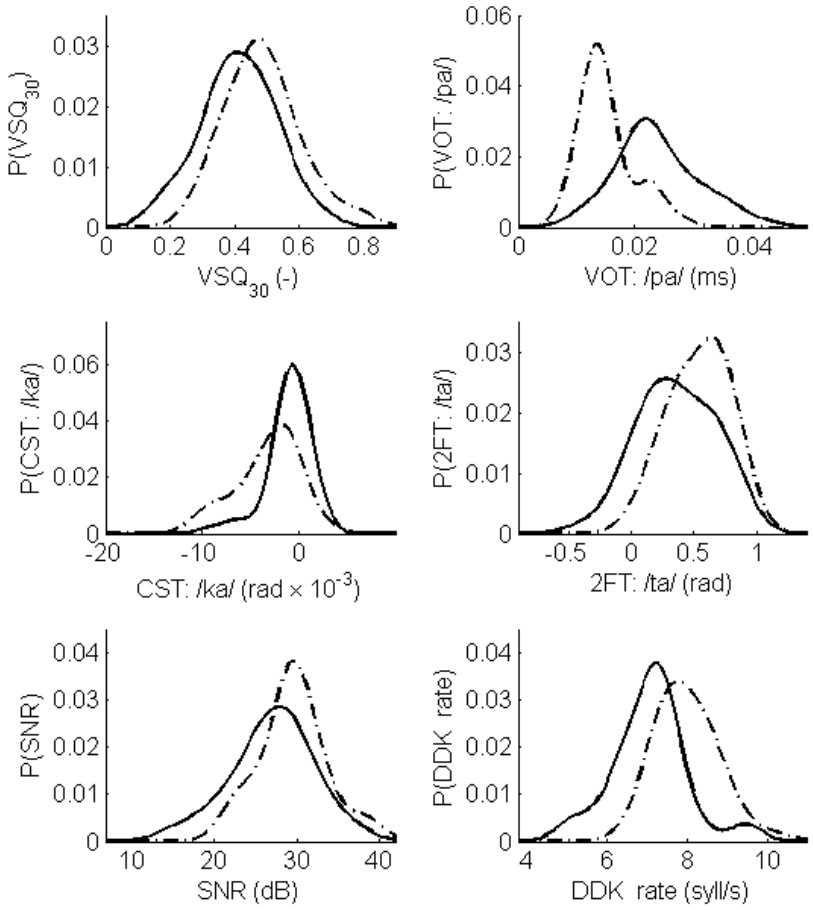

Figure 5.3: Probability densities of six representative features with the best SVM classification performance. The vertical axes are the probability densities P(measure) of feature values estimated using the Gaussian kernel density method. The dashdot lines represent the HC group and solid lines the PD group.

5.2 Velopharyngeal control

5.2.1 Perceptual analysis

The estimated ICC was 0.85 ($p < 0.001$) for inter-rater reliability and 0.86 ($p < 0.001$) for intra-rater reliability. According to perceptual tests, 89% of HD and 65% of PD patients showed mild or moderate hypernasal speech performance whereas mild hypernasality was observed in 27% of healthy speakers. The distribution of participants across four perceptual rating grades (no, mild, moderate, severe) are presented Figure 5.4.

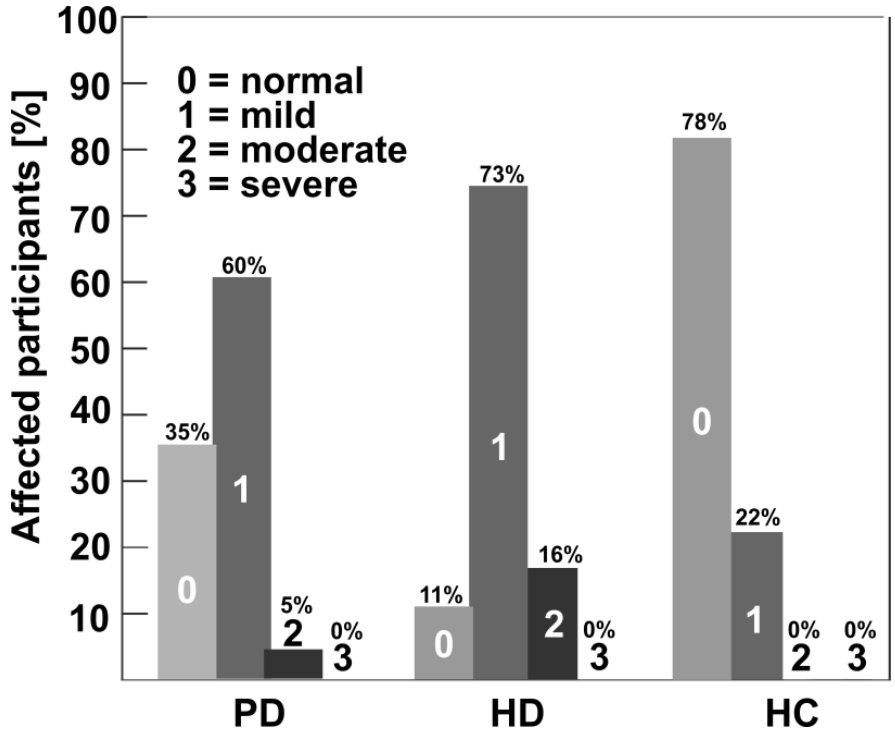

Figure 5.4: Percentage of participants perceptually scored according to GOS.SP.ASS.'98 (Sell et al., 1999) into four grades (0 = no, 1 = mild, 2 = moderate, 3 = severe).

5.2.2 Acoustic analysis

Figure 5.5 illustrates the average energy distributions in PD, HD and HC groups across 18 frequency bands. As can be seen, the HD group demonstrates spectral flattening in the area between the F1 and F2 formant frequencies.

Analysis of test-retest reliability of the proposed parameter E_{Fn} showed strong positive correlation for mean ($r = 0.87$, $p < 0.001$) and SD ($r = 0.79$, $p < 0.001$) parameters, whereas trend analysis showed only moderate positive correlation ($r = 0.47$, $p < 0.001$). Table 5.4 lists the mean, SD, and trend values for acoustic parameters E_{Fn} as well as ANOVA results with post-hoc comparison between HD vs. HC and PD vs. HC groups. Statistically significant differences between all groups were observed for E_{Fn} mean and E_{Fn} SD ($p < 0.001$), particularly due to differences between HD and HC groups ($p < 0.001$).

Figure 5.5 A-C shows the percentage of affected participants according to Wald analysis. Using cutoff values -33dB for E_{Fn} mean and 3 dB for E_{Fn} SD, we found increased nasality in 54% and abnormal nasality variability in 78% of all HD participants. In PD, increased incidence of hypernasality was observed in 27% of all patients. In HC, 19% of speakers showed hypernasality.

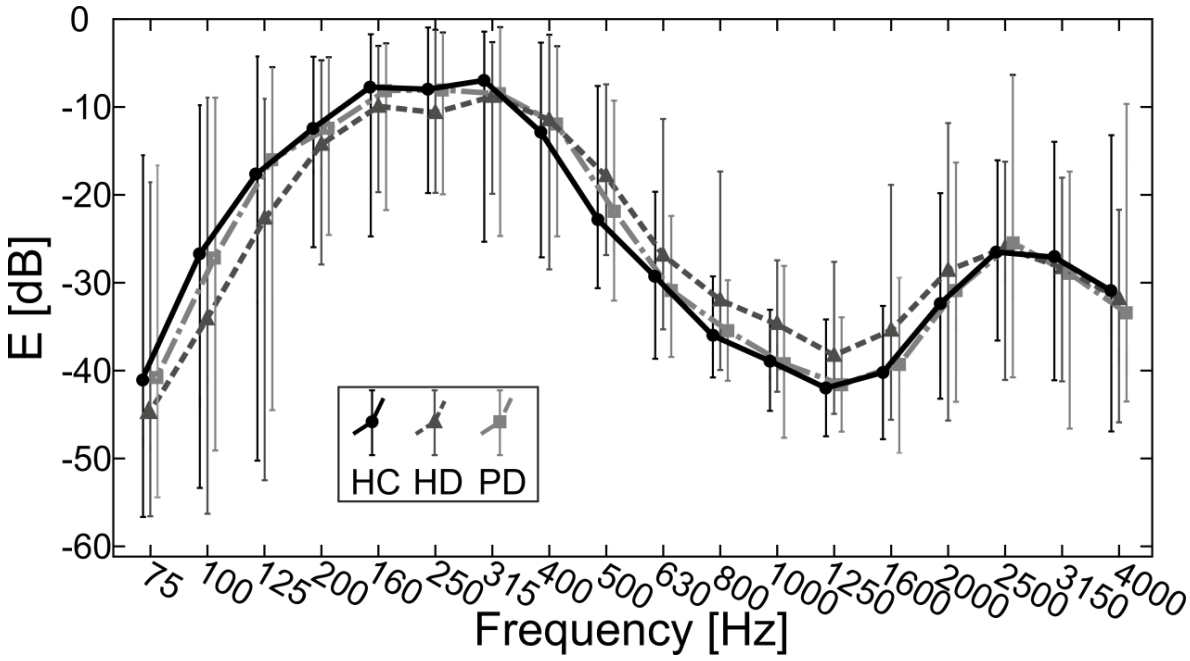

Figure 5.5: Measured values of 1/3octave spectra for 75 – 4000 Hz bands for HC, HD and PD groups.

Table 5.4: Overview of Results for HC, PD and HD Groups

Parameter	HC	HD	PD	ANOVA		Cohen's *d*†	
	Mean ± SD	Mean ± SD	Mean ± SD	F(2,108)	p	HD vs. HC	PD vs. HC
E_{Fn} mean [dB]	-39.10 ± 3.06	-34.85 ± 4.59	-38.93 ± 4.37	11.82	$p < 0.001$	1.09*	0.16
SD [dB]	2.03 ± 0.44	4.29 ± 2.17	2.17 ± 0.64	59.08	$p < 0.001$	1.44*	0.26
trend [dB/s]	-3.68 ± 17.76	-2.22 ± 82.32	-4.78 ± 18.58	0.21	$p = 0.81$	-0.10	0.06

†Mean, SD, and trend values of E_{Fn}, ANOVA F and *p* results and Cohen's *d* for post-hoc Bonferroni comparisons between HC and HD, and HC and PD groups, with the level of statistical significance marked as *) $p < 0.001$.

Table 5.5 lists the mean and SD values for acoustic parameters EFn as well as Kruskal-Wallis results with post-hoc comparison between PSP vs. PD, MSA vs. PD and PSP vs. MSA groups. Statistically significant differences between MSA and PD were observed for EFn SD ($p < 0.001$).

Table 5.5: Overview of Results for PD, PSP and MSA Groups

Parameter	PD	PSP	MSA	Kruskal-Wallis		Cohen's *d*†		
						PSP vs. PD	MSA vs. PD	PSP vs. MSA
	Mean ± SD	Mean ± SD	Mean ± SD	F(2,57)	p			
E_{Fn}								
mean [dB]	-39.10 ± 3.06	-38.99 ± 4.06	-37.95 ± 4.79	0.23	$p = 0.89$	0.12	0.12	0.23
SD [dB]	2.03 ± 0.44	2.66 ± 0.58	3.35 ± 0.68	21.12	$p < 0.001$	0.81	1.78*	1.09

†Mean and SD values of E_{Fn}, ANOVA F and *p* results and Cohen's *d* for post-hoc Bonferroni comparisons between PD, PSP, and PD and MSA, MSA and PSP groups, with the level of statistical significance marked as *) $p < 0.001$.

Figure 5.6 A and B shows the percentage of affected participants according to Wald analysis. Using cutoff values -33dB for E_{Fn} mean and 3 dB for E_{Fn} SD, we did not observe significantly different level of nasality in PSP or MSA participants compared to those diagnosed with PD; however we did observe abnormal nasality variability in 67% of all MSA participants compared to 27 % of PD participants and 25 % of PSP participants.

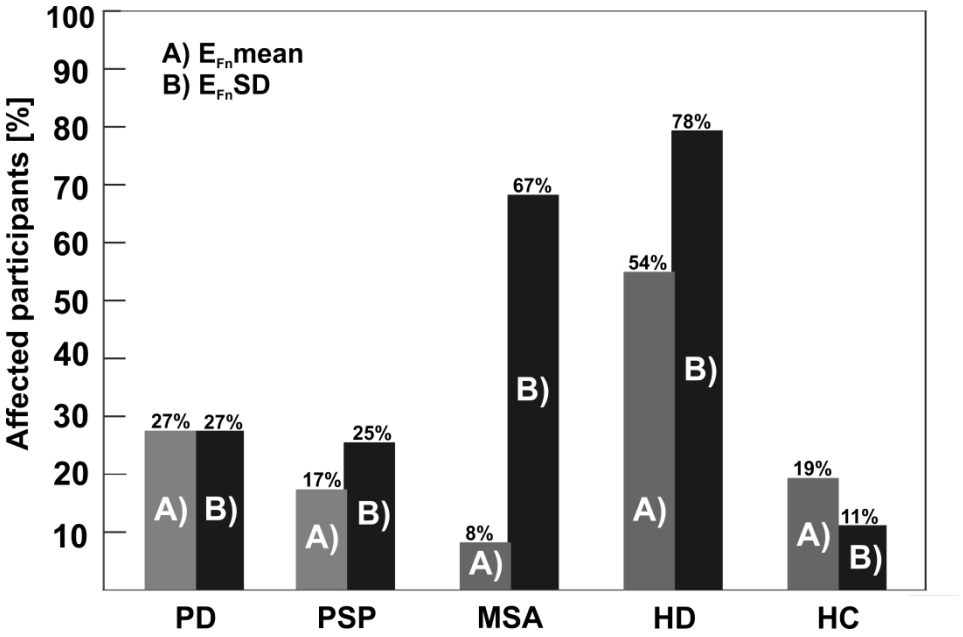

Figure 5.6: Percentage of participants marked as hypernasal using A) E_{Fn} mean and B) E_{Fn} SD.

5.2.3 Relationships between perceptual and acoustic analysis

Figure 5.7 shows comparisons related to the percentage of participants rated as hypernasal by acoustic and perceptual tests across HD, PD, and HC groups. We

observed significant correlation between perceptual test and the acoustic E_{Fn} SD parameter ($r = 0.42$, $p < 0.001$). No other correlations between perceptual and acoustic assessment were detected.

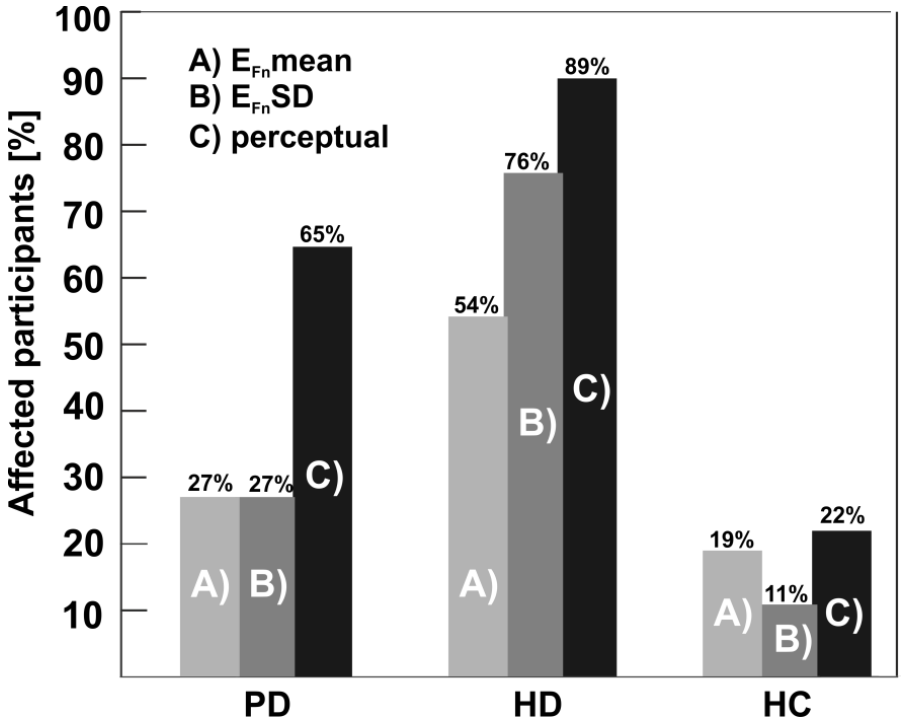

Figure 5.7: Percentage of participants marked as hypernasal using A) E_{Fn} mean, B) E_{Fn} SD, and C) perceptual rating.

5.2.4 Relationships between hypernasality and clinical manifestations

In the HD group, we observed significant relationship between the UHDRS chorea subscore and E_{Fn} SD *(*$r = 0.42$, $p = 0.01$). No other acoustic parameters showed significant correlation to disease severity scales. In the PD group, we did not detect any relationship between acoustic assessment and UPDRS score or subscores. We did not detect correlation between perceptual assessment and clinical manifestations in either PD or HD groups. No relationship between hypernasality scores and disease duration was found.

6 DISCUSSION

6.1 Rapid articulatory moves

In the thesis we present a fully automatic approach to assess articulatory and disorders in PD, which is in contrast to previous research that primarily focused on the assessment of dysphonic patterns of hypokinetic dysarthria. For the purposes of other dysarthrias estimation we have extended the original algorithm which did not reflect some other aspects including audible inspirations and inappropriate silences. Our designed speech features proved capable of describing parkinsonian hypokinetic dysarthria and huntingtonian hyperkinetic dysarthria and even described some articulatory differences between hypokinetic dysarthria and mixed dysarthrias connected with APS. Based on the speech features we were able to differentiate between speech in de novo PD patients and controls with a high classification accuracy of 88%. Interestingly, the strongest classification accuracy for a single articulatory feature was obtained through the VOT, suggesting consonant articulation is a very powerful PD indicator.

6.1.1 Algorithm performance

Automatic segmentation represented by cumulative distributions showed rapid growth of the performance in the first 5 ms of absolute difference between the detected and reference positions. Considering the 5 ms threshold for initial burst and vowel onset, our algorithm performance exceeded 85% accuracy for HC speakers and 70% accuracy for PD patients, illustrating adequate precision of the designed algorithm in the evaluation of both healthy and dysarthric speech. Since the occlusion does not provide

such abrupt change in signal energy as the initial burst or vowel onset, our algorithm reached the lowest performance of 59% within 5 ms threshold for occlusion detection but its accuracy was substantially increased to 77% when considering 10 ms threshold. Moreover, the results of the majority of our features exhibit strong or even very strong correlation to the results obtained using precise manual labels, while none fell below moderate correlation. This is crucial from the clinical point of view, as it is more important to achieve a correct estimation of the patient's speech performance than to obtain the precise position of individual boundaries.

Comparing our results with those obtained by the TEO-based algorithm using a 5 ms threshold (Hansen et al., 2010), both algorithms showed relatively high performances, exceeding 65% for utterances in healthy speakers. However, taking into account the performance of the PD group separately, the performance of the TEO-based algorithm declined under 50% accuracy, while our algorithm still maintained sufficient accuracy, exceeding 70%. Thus, the presented comparison shows that results provided by our algorithm are less vulnerable to PD-induced signal aggravation than those obtained by the TEO-based approach. Nevertheless, it is important to note that the TEO-based algorithm was primarily designed for real-time accent analysis, whereas our algorithm is focused on reliable dysarthric speech assessment, which does not require real-time processing.

The HD extension of original algorithm for the estimation of VOT in dysarthria achieved a high performance score of up to 90% in PD speakers and up to 80% in HD speakers for a 10 ms threshold. In the case of hyperkinetic dysarthria in HD, the current approach was superior to a previous algorithm designed particularly for hypokinetic dysarthria in PD, with increased performance by over 25%. Indeed, the previous approach was not sufficient in the evaluation of HD speech as hyperkinetic dysarthria may be particularly associated with audible inspirations and inappropriate voice breaks, which could affect the detection of syllable nuclei and their borders, and therefore leads to an increased occurrence of false detections. The robustness of the present algorithm to uncontrollable confounding effects in HD speech seems very promising for the automatic detection of VOT in different types of dysarthria due to various neurological conditions.

6.1.2 Articulatory deficits

Due to pathological changes in the basal ganglia, PD, APS and HD disrupt the effective execution of articulatory movements leading to various phonatory, articulatory, and prosodic disturbances. Accordingly, the analysis of freely connected speech seems to be the best way to assess the impact neurodegenerative pathology on speech (Kent et al., 1999, Rusz et al., 2013b). However, the fully automatic estimation of relevant articulatory features such as VOT from free running speech is a very difficult task and to the best of our knowledge, no such algorithm has been presented to date. To provide a robust, fully automatic classifier, previous studies have primarily used speech tests with a fixed frame such as sustained phonation (Little et al., 2009, Tsanas et al., 2012), which significantly lowers the complexity of analysis and preserves as much useful information as possible. Moreover, the advantage of analyzing sustained phonation resides in fact that the speaker's native language has no or only a small effect on dysphonia parameters. Although sustained phonation measurements provide a precise estimation of dysphonic features, Parkinsonian dysphonia is only a subset of dysarthric aspects of speech, whereas dysarthria is primarily a distinctive disorder of articulation (Duffy, 2013). Contrary to sustained phonation measurements, our approach based upon DDK task assessment provides a wide range of articulatory aspects related to dysarthria that may be subjected to evaluation, and allows their automatic assessment; however, possible language dependency cannot be excluded.

6.1.2.1 Articulatory deficits in PD

Voice quality is represented by decreased, and by increased VVQ. Decreased in both PD participants reflects increased noise caused by insufficient vocal fold adduction and phonatory instability caused by a decreased ability to keep laryngeal muscles in a fixed position (Kent et al., 2000b, Lindblom and Sundberg, 1971). Increased VVQ illustrates disrupted timing of vowel gestures (Goberman and Blomgren, 2008).

Voice onset time as the most powerful PD predictor suggests the imprecise coordination of laryngeal and supralaryngeal articulation as an early, prominent sign of PD. Each VOT measurement showed considerable prolongation of consonant duration, which may indicate disrupted coordination between the laryngeal muscle group and supralaryngeal articulators (tongue, jaws, and lips). However, previous studies focused on VOT in PD have provided inconsistent results. While some researchers reported increased or unchanged VOT in PD patients (Forrest et al., 1989, Bunton and Weismer,

2002), other studies suggested decrease in VOT due to parkinsonian articulatory disorders (Flint et al., 1992, Weismer, 1984). A study by Fischer and Goberman (2010) suggested that this inconsistency may be related to different analysis methods used and the fact that measurements were not performed rate-independently. As PD patients may be able to willingly compensate decreased speech rates, Fischer and Goberman (2010) identified the VOT ratio as an appropriate rate-independent measurement.

The willing compensation of speech rate is at the cost of reduced range of motion of the supralaryngeal articulators. The range of motion may also be reduced due to hypokinesia. Incomplete articulatory movements may be manifested as increased turbulent airflow leakage around the insufficiently closed obstacle, causing increased noise and alterations of the frequency spectrum. The significant difference between PD and HC groups, as captured by the CST of /pa/ and /ka/ syllables, illustrates the impact of insufficient articulatory movements during consonant enunciation

The effect of dysarthria on vowels may be also described by increased noise and spectral alterations. The increased noise component in consonants is probably a result of insufficient closure of the supralaryngeal articulators, whereas the vowel noise component may be the result of insufficient vocal fold adduction (Kent et al., 2000b). On the other hand, the distorted setup of supralaryngeal articulators may evoke notable changes in formant frequencies. Therefore, the 1FT and 2FT are used to indicate disruptions of articulatory movements during voicing (Kent et al., 2000b, Lindblom and Sundberg, 1971). The 1FT, which is connected with movement of the tongue in the vertical direction, illustrates impairment in all /pa/, /ta/, /ka/ syllables in both PD and HD participants. The 2FT, describing advance of the tongue, in PD shows disruption only during the /ta/ syllable, which is articulated by the tip of the tongue.

Disruption of articulatory movements leading to occlusive weakening during silent gaps between single words can be captured by decreased SNR in PD. Similar to the case of consonant articulation; this is likely caused by insufficient articulatory closure resulting in leakage of turbulent airflow (Kent et al., 2000b, Duez, 2007).

The general effect of dysarthria is well described by a considerable decrease of the DDK rate in PD speakers. Although the DDK pace measurement did not prove significant alterations in silent gap lengths, the DDK fluctuation revealed considerable instability of silent gaps in PD. The silent gap instability and non-significant DDK pace may suggest the effect of short rushes of speech, which can be caused by a combination of akinesia and speech hastening (Darley et al., 1975).

6.1.2.2 Articulatory deficits in HD

Similarly to PD, the HD voice quality dimension showed significant differences for the VSQ_{30} and VVQ features. However, the VVQ parameter was, in contrast to PD, significantly increased, suggesting severe distortion in vowel length variability. This finding may be the result of the myoclonic vowel prolongation and irregular alternating motion rates (AMR) which are prominent hyperkinetic dysarthria distortions (Duffy, 2013).

The coordination of laryngeal and supralaryngeal articulators showed less severe VOT prolongation than PD; nevertheless all VOT parameters were significantly increased. This fact may be connected with the slow and irregular AMRs also reported as dominant hyperkinetic feature (Duffy, 2013). Unlike in the case of VOT, no VOT ratio parameter revealed significant difference. This is probably due to the lower effect of HD and also due to the presence of high syllable length variability, particularly because of high VVQ parameter, which was used to normalize the VOT length.

Distorted consonant articulation showed increased CST /pa/ and /ka/ in similar manner as PD. Conversely the CSM in HD showed significant increase in all /pa/, /ta/ and /ka/ syllables suggesting inappropriate articulation during consonant articulation. This finding may be possibly caused by the presence of excessive movements which make precise articulation more difficult. This inference supports the presence of imprecise consonant articulation listed by Duffy (2013) as one of the most deviant speech dimensions.

The tongue movements were significantly distorted no matter whether they involved protrusion or elevation, which is in line with studies reporting HD related dysphagia and distorted vowels (Duffy, 2013, Leopold and Kagel, 1985). Another possible effect could have been the presence of intermittent hypernasality, which as a resonatory disruption also affects formant frequencies.

The SNR feature showed a significantly decreased distance between signal and noise. This result is probably due to the harshness and strained-strangled quality of the voice in the signal area and audible inspirations in the noise area, which are also the most prominent features of HD-related hyperkinetic dysarthria (Duffy, 2013).

The speech timing of HD utterances showed a significant decrease of speech rate, possibly due to the presence of prolonged phonemes, and slow AMRs. More interestingly, the features describing speech rhythm variability revealed a significant decrease in the ability to keep a steady rhythm. One of the most affected parameters is

DDK regularity, being more than three times longer than the one measured in HC or even in PD. Nonetheless the DDK pace and DDK fluctuation also showed highly significant results confirming the presence of distorted rhythm.

6.1.2.3 Articulatory deficits in APS

Regarding the PSP and MSA, our results show that the characteristics of speech disorder may reflect the underlying neuropathology of PD and APS. Nevertheless, basin diagnosis only on articulatory features is unreliable. Of all the twenty-nine features describing articulation, only five features from voice quality, tongue movements and speech timing dimensions differed significantly between PD and APS.

In voice quality, the VSQ_{30} feature was significantly decreased in both PSP and MSA probably due to the presence of ataxic (including harsh voice, voice tremor) and spastic (including harshness, hyprnasality, breathy voice, distorted vowels) features. Furthermore the VVQ features showed a significant increase, suggesting distorted vowels. Interestingly, though not too significantly, PSP speakers showed a noticeable increase in VVQ compared to MSA, suggesting that with more samples the difference may emerge as significant.

Considering the speech timing, the vowel length and the DDK regularity measurements were significantly increased, suggesting the more severe impact of APS on rhythm than exhibited PD participants. In addition to the aspects that were significantly distorted for both PSP and MSA, the PSP utterances showed also significantly distorted 2FT/ta/ connected with tongue tip protrusion and the vowel duration feature, suggesting phonemes prolongation in PSP.

6.1.3 Classification experiment

The presented classification experiment shows that a complex view on various aspects of Parkinsonian speech impairment using simple the task of fast syllable repetition provides great potential for fully automatic assessment of the severity of hypokinetic dysarthria in PD speakers. Using our novel DDK-based approach, we were able to predict PD_U group membership with a very high performance of approximately 87.1% using standard cross-validation and 88.4% using LOSO cross-validation. Since our database consists only of 80 speech samples from 46 participants, the advantage of standard cross-validation is that it provides lower variance in results due to possibility to set up larger test group. Yet, training and testing subsets may contain different utterances from the same individuals. This problem is treated by using of LOSO cross-

validation, however, the result variance is increased because only 2 utterances were available per subject.

Notably, the best SVM feature subset comprises six measurements where each one represents a different aspect of speech, confirming the importance of complex speech assessment in PD. It has already been shown that the complex assessment of speech profile in PD may be essential in providing information about the effect of therapy in the course of disease progression on a particular speech apparatus (Rusz et al., 2013a).

Recent studies focused on the differentiation between PD and healthy speakers presented very high classification performances of 89% (Little et al., 2009) and 98% (Tsanas et al., 2012) using a single sustained phonation task for the evaluation of dysphonia. However, considering that speech severity may be influenced by the severity of motor manifestations, disease duration, and specific effects of dopaminergic treatment ((Rusz et al., 2013a, Skodda et al., 2010, Schulz and Grant, 2000), an exact comparison with previous results is not possible. Our PD patients were investigated immediately after the diagnosis was established and before symptomatic treatment was initiated, whereas previous datasets consisted of treated Parkinsonian patients with various disease durations after diagnosis (years in (Little et al., 2009)). In our preliminary findings (Rusz et al., 2011b), we achieved 85% performance in the differentiation between PD_U and HC participants. However, this classification score was obtained using various features estimating prosody, phonation and articulation aspects together. The classification based upon single aspects achieved classification score of 81% for prosody using monologues, 76% for phonation using sustained vowels, and only 71% for articulation using fast syllable repetitions. Therefore, in comparison to these previous results, the current approach provides a performance improvement.

6.1.4 Limitations

Certain limitations of the present study must be considered. Due to the problematic recruitment of de novo PD patients, the current dataset consisted of only 24 Parkinsonian native Czech speakers. The small sample size of the present study may bias the performance of the classifier to a certain extent. Although newly diagnosed, the majority of our patients were already in the middle H&Y stages 2 or 2.5. However, to consider speech tests as diagnostic decision support tool for an early diagnosis of PD,

we would need to differentiate between controls and untreated PD speakers in their very early disease stages.

As our PD_{SDD} patients were investigated in their ON condition, we cannot exclude that some differences between PD_{SDD} and APS were more pronounced due to the beneficial effect of dopaminergic therapy. However, it is assumed that short-term dopaminergic therapy has no or very little effect on speech in PD (Ho et al., 2008). We did not differentiate between speech in the various subtypes of PSP and MSA due to the limited opportunity in recruiting a larger number of participants. Nevertheless, at least in PSP patients, different subtypes of disease seem to have no substantial effect on global speech performance (Skodda et al., 2011).

Furthermore, the language dependency of features extracted from the DDK task cannot be excluded as such patterns have never been investigated. Another limitation of the current dataset is gender imbalance, related to the greater incidence of PD in males (Van Den Eeden et al., 2003, Baldereschi et al., 2000). Previous studies have documented a confounding effect of sexual dimorphism on particular speech impairments (Hertrich and Ackermann, 1995), and we therefore cannot exclude the possibility that articulatory impairment is influenced by gender-specific aspects of speech. Finally, our algorithm was primarily designed for parkinsonian patients with mild to moderate stages of disease and thus does not need to be sufficiently sensitive to evaluation of articulatory disorders in PD patients with advanced motor stages and severe dysarthria.

6.2 Velopharyngeal control

Based upon the 1/3-octave spectra analysis presented by Kataoka et al. (1996) and the acoustic model of the vocal tract published by (Stevens, 2000), we designed the parameter E_{Fn} to evaluate the presence and character of hypernasality in prolonged vowels. Using acoustic analysis, we revealed an occurrence of hypernasality in 54% of HD, 27% of PD, 17% PSP, 8% MSA and 19% of HC speakers. In addition, our results showed a high occurrence of intermittent hypernasality in 78% of HD patients and 67% of MSA patients. Perceptual analysis showed the occurrence of mild to moderate hypernasality in 89% HD, 65% PD and 22% HC speakers. Significant correlation between the acoustic parameter representing nasality fluctuation and perceptual assessment was observed. Furthermore, we revealed significant correlation between acoustic metric representing nasality fluctuation and chorea in HD patients.

6.2.1 Nasality in PD

Although using acoustic analysis we detected hypernasality in 27% of PD speakers, the non-significant difference between PD and HC groups suggests that hypernasality is a non-prominent speech manifestation. Previous studies focused on hypernasality in PD have provided rather inconsistent conclusions. Based on perceptual evaluation, Ludlow and Basich (1983) included hypernasality among the 10 most salient features connected with dysarthria, whereas Darley et al. (1975) and Duffy (2013) found hypernasality to be non-prominent manifestation of PD. In particular, based on a large sample of PD patients, Logemann et al. (1978) observed hypernasality in only 10% of participants. Considering instrumental analyses, only Mueller (1971) failed to detect hypernasality in PD speakers, contrary to the majority of studies reporting an increased occurrence of hypernasality in PD participants (Hoodin and Gilbert, 1989, Netsell et al., 1975, Theodoros et al., 1995). While the differences in perceptual assessments could be explained by the fact that listeners from various cultures may have a different level of tolerance for perceived hypernasality, inconsistencies in the instrumental assessment are likely due to the differing sensitivity of particular methods. Moreover, both perceptual and instrumental assessment could be biased by differences in the sample data, as the majority of previous studies have reported hypernasality in a minority of PD speakers. One further explanation for these discrepancies may be that the severity of hypernasality parallels overall disease progression to some extent (Hoodin and Gilbert, 1989). However, we did not observe any relation between hypernasality metrics and disease duration or motor severity scales in PD.

6.2.2 Nasality in HD

The presence of hypernasality was observed in the majority of our HD speakers, which was mainly associated with the occurrence of abnormal nasality variability. Although this is the first known study to objectively examine hypernasality in HD, our findings are in accordance with Duffy (2013), perceptually indicating intermittent hypernasality as a salient feature of HD. Indeed, we observed correlation between acoustic nasality variability and the chorea UHDRS subscore, demonstrating the significant impact of chorea on velopharyngeal mechanism. This is also in accordance

with Duffy (2013), who proposed the effect of chorea on articulatory-resonatory incompetence.

6.2.3 Nasality in APS

The observed results suggest that MSA patients do not exhibit increased nasality, even though 67% of them exhibited significantly increased nasality variability. Notably, very little was found in the literature on the topic of nasality in MSA. One possible explanation is that MSA patients experiencing largely ataxic dysarthria disorders (Rusz et al., 2015, Duffy, 2013), are affected by intermittent hyponasality connected with ataxic dysarthria (Duffy, 2013).. This assumption is also supported by the very low presence of increased nasality occurring only in 8% of MSA speakers compared to 19% of HC speakers.

In the PSP group, the 1/3-octave band spectra method has not detect any significant difference between PSP and PD or even PSP or HC. This fact suggests that hypernasality is not a prominent feature describing PSP-induced dysarthria. The result is in congruence with the study (Rusz et al., 2015), which found the hypokinetic dysarthria and spastic dysarthria aspects to be the most defining.

6.2.4 Acoustic assessment of hypernasality

In the present study we applied an acoustic method designed for the objective evaluation of velopharyngeal insufficiency, to determine the presence and nature of velopharyngeal incompetency in PD, MSA, PSP and HD. This methodology has been previously verified by Vogel et al. (2009) and successfully applied to patients with Friedriech ataxia resulting in velopharyngeal incompetency (Poole et al., 2015). Based upon an acoustic model of the vowel /i/ published by Stevens (2000) and recommendations presented by Kent et al. (1999), we designed the E_{Fn} parameter to describe the presence of nasal resonance in speech due to properties of the nasal cavity present in the 1 kHz 1/3-octave band (Stevens, 2000, Kataoka et al., 1996). This assumption is valid for all vowels; nevertheless the wide plateau between F1 and F2 frequencies in the vowel /i/ makes the presence of nasal resonance more pronounced (Stevens, 2000, Kataoka et al., 1996). Compared to controls, the parameter E_{Fn} mean showed significantly increased energy in HD patients with a large effect size, suggesting an abnormal presence of hypernasality in HD patients. Furthermore, using the parameter E_{Fn} SD, we revealed significant differences in fluctuations of nasality

between HD and control, and MSA and control speakers, suggesting intermittent hypernasality in HD and MSA patients. The parameter E_{Fn} trend was found to be unreliable, as it demonstrated no significant differences between groups and low test-retest reliability.

6.2.5 Perceptual assessment of hypernasality

Previous studies have reported perceptual assessment of hypernasality in dysarthria as rather unreliable as hypernasality is masked by more dominant dysarthria signs, which confounds the perceived nasalance (Brancewicz and Reich, 1989). Accordingly, our results indicate more HD and PD participants systematically rated as hypernasal by perceptual assessment than by an instrumental approach, likely due to difficulty in achieving accurate perception of hypernasality when other abnormal dysarthria characteristics are present. Furthermore, the difference between speech tasks used during perceptual and instrumental evaluation could be a source of discrepancy between acoustic and perceptual assessments.

There is a little evidence for correlation between perceptual and instrumental measurements of hypernasality in dysarthrias (Poole et al., 2015, Theodoros et al., 1995). In our HD sample, acoustic analyses identified only 50% of all HD speakers as hypernasal in comparison to the perceptual rating of nearly 90%. Yet, the abnormally intermittent character of nasality was also acoustically observed in nearly 80% of all HD participants. As we observed significant correlation between acoustic parameters measuring intermittent hypernasality and perceptual ranking, we may hypothesize that fluctuation in the level of nasality makes resonatory disruptions more obvious to perceptual raters. Interestingly, these correlations were evident even if perceptual and acoustic assessment were performed using different speech material.

In agreement with our findings, previous studies have perceptually rated the majority of PD participants as mildly hypernasal (Theodoros et al., 1995, Hoodin and Gilbert, 1989). However, our raters tended to score PD speakers more strictly, and in ambiguous cases decided in favor of mild hypernasality. Indeed, some mild hypernasality is not rare even in healthy subjects and was observed in up to 22% of our control speakers, which is in accordance with previous research (Poole et al., 2015). Given this evidence, we may suppose that the perceptual decision between normal and mildly hypernasal speech can be misleading, particularly in dysarthrias with other perceptually dominant speech deviations.

6.2.6 Limitations

We did not perform aerodynamic measurements, which would provide direct information about nasal airflow. Nevertheless, a previous study by Vogel et al. (2009) provided exhaustive evaluation of the 1/3-octave method and other studies have successfully applied this method to hypernasality assessment (Poole et al., 2015, Lee et al., 2003, Yoshida et al., 2000, Kataoka et al., 1996). The advantage of the current approach is that it provides an easy-to-administer acoustic assessment, which would be possible to integrate into a larger battery of acoustic tests.

One limitation is that we used different speech tasks for the perceptual and acoustic evaluation of hypernasality, as accurate perceptual evaluation of hypernasality from sustained vowel phonation is not feasible. Indeed, the different speech tasks used likely make correlation analyses between perceptual and acoustic variables problematic. In future studies, it may therefore be beneficial to include rating for consistency, as with the Consensus Auditory Perceptual Evaluation of Voice (Kempster et al., 2009).

We did not test the consistency and reliability of UPDRS and UHDRS metrics. Nevertheless, relationships between nasality and motor abnormalities were found only for the UHDRS chorea subscore, which showed high inter-rater reliability with an ICC of 0.82 (Huntington-Study-Group, 1996).

As HD generally has an earlier onset than PD, the PD and HD participant groups could not be age-matched. Therefore, we matched the age of the control group to the age of generally older PD group, as nasality is expected to remain stable throughout life or may slightly deteriorate as a consequence of aging (Ramig and Ringel, 1983, Hoit et al., 1994). This approach ensures that the results of the PD group were not favored in comparison with the HC group. Moreover, we did not match our groups according to gender. Nevertheless, previous studies did not find differences in nasality between male and female speakers (Litzaw and Dalston, 1992, Joos et al., 2006).

7 CONCLUSION

The main purpose of the thesis was to provide approaches towards an objective assessment of articulatory and resonatory deficits resulting from dysarthria. Results obtained for both of these approaches show the algorithm to be feasible and applicable. Moreover, the acoustic approaches show greater sensitivity in comparison with commonly used perceptual assessment. Therefore, the acoustic analyses are presented as a useful tool providing good conditions for the early diagnosis, monitoring of treatment efficacy, speech therapy monitoring and adjustment, and even the differential diagnosis. Analysis of a number of acoustic parameters enabled the examination of a wide variety of articulatory deficits. Statistical analysis revealed distortion in every examined articulatory dimension of speech. The classification experiments showed the applicability of the designed parameters for the classification of healthy and pathological utterances. Moreover, the distinctive patterns of dysarthria in different neurodegenerative diseases showed a possible application of differential diagnosis cues.

7.1 Summary

In the summary section we provide a brief list of the thesis goals and achieved results. For more detailed discussion of results obtained in each study, see the section Discussion.

- **To develop an automatic segmentation algorithm allowing for the accurate detection of the initial burst, vowel onset, and occlusion in patients diagnosed with PD.**
  - The presented algorithm shows sufficient sensitivity and robustness, as was documented by the cumulative distributions measured for each detected event i.e. initial burst, vowel onset and occlusion. The score obtained for the 5ms error range in the initial burst and the vowel onset detection reached nearly 80% in the joint PD and HC group and for the 10ms error range in the occlusion detection reached 77% in the joint PD and HC group. Furthermore, the parameters computed from the detected positions show a high to very correlation (r = 0.70 to 0.99) with those assessed using the reference labels.

- **To extend the detection algorithm to be applicable to different dysarthria subtypes, and to demonstrate its applicability by evaluating PD and HD speakers.**
  - Comparison of the algorithm designed for the PD utterance assessment and the updated algorithm shows similar results in the score obtained assessing PD participants. Nevertheless in HD participants, the 10ms performance score of original algorithm reached only 45%, which is significantly lower than the 10ms performance score of over 77% reached by the updated algorithm.

- **Using the proposed segmentation algorithm to introduce several acoustic features sensitive to possible articulatory deficits caused by hypokinetic dysarthria.**
  - The section Articulatory Features presented thirteen acoustic parameters (i.e. VSQ, VSQ_{30}, VVQ, VOT, VOT ratio, 1FT, 2FT, CSM, CST, SNR, DDK rate, DDK fluctuations and DDK pace) which were used to assess six different speech dimensions (i.e. speech quality, laryngeal and supralaryngeal coordination,

precision of consonant articulation, tongue movement, occlusion weakening, and speech timing). Statistical analysis revealed significantly affected articulatory aspects in every examined neurodegenerative disease. Furthermore, each speech dimension contained at least one significantly distorted parameter.

- **To explore the suitability of the designed acoustic features in capturing parkinsonian articulatory disorder, and to perform a classification experiment in order to differentiate PD subjects from controls.**
  - Using the SVM classifier the performed classification experiment illustrates the ability of the automatic measurement to distinguish between PD and HC groups. Interestingly the best classification score ($87 \pm 5\%$ using 60/40 cross-validation and $88 \pm 26\%$ using LOSO cross-validation) was obtained using six parameters (i.e. VSQ_{30},VOT:/pa/,CST:/ka/, 2FT:/ta/, SNR, DDK rate) representing every assessed speech dimension. Moreover the score of classification based only on one parameter (VOT:all) reached $83 \pm 5\%$ using 60/40 cross-validation and $79 \pm 34\%$ using LOSO cross-validation.

- **To determine specific patterns of articulatory disruptions and estimate their reliability in differentiating between PD, PSP and MSA**
  - We have detected distinctive patterns in PD, PSP and MSA dysarthria types. The only distinctive manifestation of PSP was dysfluency, which was only rarely detected in MSA. Regarding MSA, participants showed poorer voice control described by VVQ and VSQ_{30} features and distinctive patterns in DDK regularity. Furthermore, PSP showed significantly changed 2FT/ta/ and vowel duration features.

- **To employ methods of objective hypernasality assessment, which may be easily grasped by the tools of the differential diagnosis.**
  - We have implemented the 1/3-octave method, which was originally designed for the purposes of assessment of hypernasality resulting from VIS and based on this approach we have designed objective parameters of dysarthria assessment. The presented parameters enabled the objective evaluation of nasality in the presence and behavior of hypernasality in the dysarthric utterance, which is connected with VIC hypernasality.

- **To evaluate the presence and character of hypernasality in PD and HD speakers.**
  - The perceptual analysis has revealed the occurrence of hypernasality in 89% HD, 65% PD and 22% HC speakers. The acoustic analysis detected hypernasality in 54% of HD, 27% of PD and 19% HC speakers. In addition our results showed a high occurrence of intermittent hypernasality in 78% of HD patients.

- **To examine possible relationships between the severity of hypernasality and disease-specific motor manifestations, to provide more insight into the pathophysiology responsible for development of hypernasality in basal ganglia disorders.**
  - The relationship analysis revealed a moderate correlation ($r = 0.42$) between the chorea UHDRS subscore and the E_{Fn} suggesting the impact of chorea on the mechanism of velopharyngeal control. The correlation analysis has not showed any correlation between hypernasality and clinical manifestations in PD.

- **To determine specific resonatory disruption patterns characteristic for PSP, MSA and PD.**
  - Using the Wald task 67 % of MSA compared to 27% of PD participants have been detected with increased nasality variability. This result has been confirmed as statistically significant using the Kruskal-Wallis test with post-hoc Bonfferoni adjustment. With regard to PSP no statistically significant increase compared to PD has been detected.

7.2 Further work

Dysarthrias, as a very complex speech disorder, requires more research to be better understood. Further research provides two possible directions of interest. The first one is to extend the description of previously investigated dysarthrias, whereas the second is to extend the database of different dysarthrias themselves.

The improvement of the dysarthria description should incorporate assessments of different speech dimensions to provide better insight into the speech pathologies and

the underlying neuropathology. For this reason, one of the future aims is to provide dysarthria assessment describing respiration, phonation, articulation, resonance and prosody. A combination of parameters from different speech dimensions has already been described to some extent in the study (Rusz et al., 2015), showing that the combination of features describing different speech dimensions enables differentiation even between PSP and MSA.

Furthermore, additional non-speech information may be included, for instance in mimics. Studies by Crosiers et al. (2011) and Zingler et al. (2005) suggest the presence of hemihypomimia in PD, though both studies are more or less exploratory and further research needs to be conducted. The other area of distorted facial movements is the disruption of saccades in PSP, which is used as a distinctive pattern for differential diagnosis (Rehman, 2000, Litvan et al., 1996). In this case automatic assessment may provide objective and more sensitive method of saccadic movement estimation.

The extension of the dysarthrias database should provide several different points of view. First view may be the duration of the disease. For instance, this could involve cases when, the PD_U group has been recorded right after diagnosis and before administration of pharmacological treatment, therefore providing a sample with earliest diagnosed PD available. The fact that PD is diagnosed several years after onset and that during the time of diagnosis the 60 % - 70 % of dopaminergic neurons is lost and the dopamine concentration is reduce by 80 %, even the PD_U group could be hardly described as patients in the early stage (Bernheimer et al., 1973, Postuma et al., 2012). A solution of this issue may be provided through longitudinal recording of participants experiencing rapid eye movement sleep behavior disorder, which has been shown to often develop in PD (Postuma et al., 2012), or longitudinal assessment of genetically diagnosed HD participants.

Furthermore, other samples of different motor neurodegenerative disorders should be examined to provide better understanding of different neuronal pathology effects on speech performance. For this reason, patients diagnosed with multiple sclerosis or amyotrophic lateral sclerosis should be recorded.

8 REFERENCES

ABBS, J. H., HUNKER, C. J. & BARLOW, S. M. 1983. Differential speech motor subsystem impairments with suprabulbar leisons: neurophysiological framework and supporting data. *In:* BERRY, W. R. (ed.) *Clinical dysarthria.* San Diego: College-Hill Press.

ACKERMANN, H. & HERTRICH, I. 1997. Voice onset time in ataxic dysarthria. *Brain and Language,* 56**,** 321-333.

ACKERMANN, H., HERTRICH, I. & HEHR, T. 1995a. Oral Diadochokinesis in - Neurological Dysarthrias. *Folia Phoniatrica Et Logopaedica,* 47**,** 15-23.

ACKERMANN, H., HERTRICH, I. & SCHARF, G. 1995b. Kinematic Analysis of Lower Lip Movements in Ataxic Dysarthria. *Journal of Speech and Hearing Research,* 38**,** 1252-1259.

ACKERMANN, H., KONCZAK, J. & HERTRICH, I. 1997. The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and Language,* 56**,** 312-319.

ACKERMANN, H. & ZIEGLER, W. 1995. Articulatory deficits in Parkinson's and Huntington's disease: An acoustic analysis. *Alzheimer's and Parkinson's Diseases,* 44**,** 29-32.

ADAMS, S. G. 1997. Hypokinetic dysarthria in Parkinson's disease. *In:* MCNEIL, M. R. (ed.) *Clinical management of sensorimotor speech disorders.* New York: Thieme.

AIZMAN, O., BRISMAR, H., UHLEN, P., ZETTERGREN, E., LEVEY, A. I., FORSSBERG, H., GREENGARD, P. & APERIA, A. 2000. Anatomical and physiological evidence for D-1 and D-2 dopamine receptor colocalization in neostriatal neurons. *Nature Neuroscience,* 3**,** 226-230.

ALBIN, R. L., YOUNG, A. B. & PENNEY, J. B. 1989. The Functional-Anatomy of Basal Ganglia Disorders. *Trends in Neurosciences,* 12**,** 366-375.

ALEXANDER, M. P., NAESER, M. A. & PALUMBO, C. L. 1987. Correlations of subcortical CT lesion sites and aphasia profiles. *Brain,* 110 (Pt 4)**,** 961-91.

ALLEN, J. S., MILLER, J. L. & DESTENO, D. 2003. Individual talker differences in voice-onset-time. *J Acoust Soc Am,* 113**,** 544-52.

ARONSON, A. E. 1990. *Clinical voice disorders,* New York, Thieme.

BALDERESCHI, M., DI CARLO, A., ROCCA, W. A., VANNI, P., MAGGI, S., PERISSINOTTO, E., GRIGOLETTO, F., AMADUCCI, L., INZITARI, D. &

GRP, I. W. 2000. Parkinson's disease and parkinsonism in a longitudinal study - Two-fold higher incidence in men. *Neurology,* 55**,** 1358-1363.

BATES, G. P. & BENN, C. 2002. The polyglutamine diseases. *In:* BATES, G., HARPER, P. & JONES, L. (eds.) *Huntington's disease.* New York: Oxford University Press.

BECKER, G., MULLER, A., BRAUNE, S., BUTTNER, T., BENECKE, R., GREULICH, W., KLEIN, W., MARK, G., RIEKE, J. & THUMLER, R. 2002. Early diagnosis of Parkinson's disease. *J Neurol,* 249 Suppl 3**,** III/40-8.

BERNHEIMER, H., BIRKMAYER, W., HORNYKIEWICZ, O., JELLINGER, K. & SEITELBERGER, F. 1973. Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. *J Neurol Sci,* 20**,** 415-55.

BOESCH, S. M., WENNING, G. K., RANSMAYR, G. & POEWE, W. 2002. Dystonia in multiple system atrophy. *J Neurol Neurosurg Psychiatry,* 72**,** 300-3.

BOEVE, B. F. 2012. Progressive supranuclear palsy. *Parkinsonism Relat Disord,* 18 Suppl 1**,** S192-4.

BONNET, A. M., PICHON, J., VIDAILHET, M., GOUIDER-KHOUJA, N., ROBAIN, G., PERRIGOT, M. & AGID, Y. 1997. Urinary disturbances in striatonigral degeneration and Parkinson's disease: clinical and urodynamic aspects. *Mov Disord,* 12**,** 509-13.

BOWER, J. H., MARAGANORE, D. M., MCDONNELL, S. K. & ROCCA, W. A. 1997. Incidence of progressive supranuclear palsy and multiple system atrophy in Olmsted County, Minnesota, 1976 to 1990. *Neurology,* 49**,** 1284-8.

BRANCEWICZ, T. M. & REICH, A. R. 1989. Speech rate reduction and "nasality" in normal speakers. *J Speech Hear Res,* 32**,** 837-48.

BROWN, S., LAIRD, A. R., PFORDRESHER, P. Q., THELEN, S. M., TURKELTAUB, P. & LIOTTI, M. 2009. The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn,* 70**,** 31-41.

BUNTON, K. 2006. Fundamental frequency as a perceptual cue for vowel identification in speakers with Parkinson's disease. *Folia Phoniatr Logop,* 58**,** 323-39.

BUNTON, K. & WEISMER, G. 2002. Segmental level analysis of laryngeal function in persons with motor speech disorders. *Folia Phoniatr Logop,* 54**,** 223-39.

CALIGIURI, M. P. 1989. The influence of speaking rate on articulatory hypokinesia in parkinsonian dysarthria. *Brain Lang,* 36**,** 493-502.

CANTER, G. J. 1963. Speech Characteristics of Patients with Parkinson's Disease: I. Intensity, Pitch, and Duration. *J Speech Hear Disord,* 28**,** 221-9.

CANTER, G. J. 1965a. Speech Characteristics of Patients with Parkinson's Disease. 3. Articulation, Diadochokinesis, and over-All Speech Adequacy. *J Speech Hear Disord,* 30**,** 217-24.

CANTER, G. J. 1965b. Speech Characteristics of Patients with Parkinson's Disease. Ii. Physiological Support for Speech. *J Speech Hear Disord,* 30**,** 44-9.

CMEJLA, R., RUSZ, J., BERGL, P. & VOKRAL, J. 2013. Bayesian changepoint detection for the automatic assessment of fluency and articulatory disorders. *Speech Communication,* 55**,** 178-189.

CROSIERS, D., MARECHAL, E., VAN AEL, Y. & CRAS, P. 2011. Left-sided hemihypomimia in Parkinson's disease. *Acta Neurol Belg,* 111**,** 225-7.

DANIEL, S. E., DE BRUIN, V. M. & LEES, A. J. 1995. The clinical and pathological spectrum of Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy): a reappraisal. *Brain,* 118 (Pt 3)**,** 759-70.

DARLEY, F. L., ARONSON, A. E. & BROWN, J. R. 1969. Differential diagnostic patterns of dysarthria. *J Speech Hear Res,* 12**,** 246-69.

DARLEY, F. L., ARONSON, A. E. & BROWN, J. R. 1975. *Motor speech disorders,* Philadephia, WB Saunders.

DARLING, M. & HUBER, J. E. 2011. Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. *J Speech Lang Hear Res,* 54**,** 1247-59.

DE LETTER, M., SANTENS, P., DE BODT, M., VAN MAELE, G., VAN BORSEL, J. & BOON, P. 2007. The effect of levodopa on respiration and word intelligibility in people with advanced Parkinson's disease. *Clinical Neurology and Neurosurgery,* 109**,** 495-500.

DE RIJK, M. C., LAUNER, L. J., BERGER, K., BRETELER, M. M., DARTIGUES, J. F., BALDERESCHI, M., FRATIGLIONI, L., LOBO, A., MARTINEZ-LAGE, J., TRENKWALDER, C. & HOFMAN, A. 2000. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology,* 54**,** S21-3.

DE RIJK, M. C., TZOURIO, C., BRETELER, M. M., DARTIGUES, J. F., AMADUCCI, L., LOPEZ-POUSA, S., MANUBENS-BERTRAN, J. M., ALPEROVITCH, A. & ROCCA, W. A. 1997. Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON Collaborative Study. European Community Concerted Action on the Epidemiology of Parkinson's disease. *J Neurol Neurosurg Psychiatry,* 62**,** 10-5.

DELONG, M. R. & WICHMANN, T. 2007. Circuits and circuit disorders of the basal ganglia. *Arch Neurol,* 64**,** 20-4.

DILLENSCHNEIDER, E., ZALESKI, T. & GREINER, G. F. 1973. [Sonagraphic study of nasality in cases of palatovelar insufficiency]. *JFORL J Fr Otorhinolaryngol Audiophonol Chir Maxillofac,* 22**,** 201-2.

DUEZ, D. 2007. Acoustic analysis of occlusive weakening in parkinsonian French speech. *International Congress of Phonetic Sciences.* Saarbrücken.

DUFFY, J. R. 2013. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management,* St. Louis, Elsevier Health Sciences.

FANT, G. 1960. *Acoustic theory of speech production,* Mouton, The Hague.

FEDERICO, F., SIMONE, I. L., LUCIVERO, V., DE MARI, M., GIANNINI, P., ILICETO, G., MEZZAPESA, D. M. & LAMBERTI, P. 1997. Proton magnetic resonance spectroscopy in Parkinson's disease and progressive supranuclear palsy. *J Neurol Neurosurg Psychiatry,* 62**,** 239-42.

FISCHER, E. & GOBERMAN, A. M. 2010. Voice onset time in Parkinson disease. *Journal of Communication Disorders,* 43**,** 21-34.

FLAHERTY, A. W. & GRAYBIEL, A. M. 1993. Two input systems for body representations in the primate striatal matrix: experimental evidence in the squirrel monkey. *J Neurosci,* 13**,** 1120-37.

FLETCHER, S. G. 1972. Time-by-count measurement of diadochokinetic syllable rate. *J Speech Hear Res,* 15**,** 763-70.

FLINT, A. J., BLACK, S. E., CAMPBELLTAYLOR, I., GAILEY, G. F. & LEVINTON, C. 1992. Acoustic Analysis in the Differentiation of

Parkinsons-Disease and Major Depression. *Journal of Psycholinguistic Research,* 21**,** 383-399.

FOLSTEIN, S. 1989. *Huntington's disease: a disorder of families,* Maryland, The Johns Hopkins University Press.

FORREST, K. & WEISMER, G. 1995. Dynamic aspects of lower lip movement in parkinsonian and neurologically normal geriatric speakers' production of stress. *J Speech Hear Res,* 38**,** 260-72.

FORREST, K., WEISMER, G. & TURNER, G. S. 1989. Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *J Acoust Soc Am,* 85**,** 2608-22.

FRANCOIS, C., GRABLI, D., MCCAIRN, K., JAN, C., KARACHI, C., HIRSCH, E. C., FEGER, J. & TREMBLAY, L. 2004. Behavioural disorders induced by external globus pallidus dysfunction in primates II. Anatomical study. *Brain,* 127**,** 2055-70.

FRANCOIS, C., SAVY, C., JAN, C., TANDE, D., HIRSCH, E. C. & YELNIK, J. 2000. Dopaminergic innervation of the subthalamic nucleus in the normal state, in MPTP-treated monkeys, and in Parkinson's disease patients. *J Comp Neurol,* 425**,** 121-9.

GALLENA, S., SMITH, P. J., ZEFFIRO, T. & LUDLOW, C. L. 2001. Effects of levodopa on laryngeal muscle activity for voice onset and offset in Parkinson disease. *Journal of Speech Language and Hearing Research,* 44**,** 1284-1299.

GAMBOA, J., JIMENEZJIMENEZ, F., NIETO, A., MONTOJO, J., ORTIPAREJA, M., MOLINA, J. A., GARCIAALBEA, E. & COBETA, I. 1997. Acoustic voice analysis in patients with Parkinson's disease treated with dopaminergic drugs. *Journal of Voice,* 11**,** 314-320.

GENTIL, M. 1990. Acoustic Characteristics of Speech in Friedreichs Disease. *Folia Phoniatrica,* 42**,** 125-134.

GENTIL, M., PINTO, S., POLLAK, P. & BENABID, A. L. 2003. Effect of bilateral stimulation of the subthalamic nucleus on parkinsonian dysarthria. *Brain Lang,* 85**,** 190-6.

GERFEN, C. R., ENGBER, T. M., MAHAN, L. C., SUSEL, Z., CHASE, T. N., MONSMA, F. J. & SIBLEY, D. R. 1990. D1 and D2 Dopamine Receptor Regulated Gene-Expression of Striatonigral and Striatopallidal Neurons. *Science,* 250**,** 1429-1432.

GILMAN, S., WENNING, G. K., LOW, P. A., BROOKS, D. J., MATHIAS, C. J., TROJANOWSKI, J. Q., WOOD, N. W., COLOSIMO, C., DURR, A., FOWLER, C. J., KAUFMANN, H., KLOCKGETHER, T., LEES, A., POEWE, W., QUINN, N., REVESZ, T., ROBERTSON, D., SANDRONI, P., SEPPI, K. & VIDAILHET, M. 2008. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology,* 71**,** 670-6.

GOBERMAN, A. M. & BLOMGREN, M. 2008. Fundamental frequency change during offset and onset of voicing in individuals with Parkinson disease. *Journal of Voice,* 22**,** 178-191.

GOBERMAN, A. M. & COELHO, C. 2002. Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy. *Neurorehabilitation,* 17**,** 237-246.

GOUVEIA, R. G., PINTO, A., EVANGELISTA, T., ATALAIA, A., CONCEICAO, I. & DE CARVALHO, M. 2006. Evidence for central abnormality in respiratory control in primary lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders,* 7**,** 57-60.

GUTEKUNST, C., NORFLUS, F. & HERSCH, S. 2002. The neuropathology of Huntington's disease. *In:* BATES, G., HARPER, P. & JONES, L. (eds.) *Huntington's disease.* New York: Oxford University Press.

HALLETT, M. 1993. Physiology of basal ganglia disorders: an overview. *Can J Neurol Sci,* 20**,** 177-83.

HANSEN, J. H. L., GRAY, S. S. & KIM, W. 2010. Automatic voice onset time detection for unvoiced stops (/p/,/t/,/k/) with application to accent classification. *Speech Communication,* 52**,** 777-789.

HANSON, D. G., GERRATT, B. R. & WARD, P. H. 1984. Cinegraphic observations of laryngeal function in Parkinson's disease. *Laryngoscope,* 94**,** 348-53.

HARDIN, M. A., VANDEMARK, D. R., MORRIS, H. L. & PAYNE, M. M. 1992. Correspondence between Nasalance Scores and Listener Judgments of Hypernasality and Hyponasality. *Cleft Palate-Craniofacial Journal,* 29**,** 346-351.

HARPER, P. & JONES, L. 2002. Huntington's disease: genetic and molecular studies. *In:* BATES, G., HARPER, P. & JONES, L. (eds.) *Huntington's disease.* New York: Oxford University Press.

HARRIS, J. D. & NELSON, D. 1993. Glottal Pulse Alignment in Voiced Speech for Pitch Determination. *Icassp-93 : 1993 Ieee International Conference on Acoustics, Speech, and Signal Processing, Vols 1-5***,** B519-B522.

HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. 2009. *The elements of statistical learning: Datamining, inference, and predicition* New York, Springer.

HAYDEN, M. R. 1981. *Huntington's chorea,* New York, Springer.

HERTRICH, I. & ACKERMANN, H. 1994. Acoustic analysis of speech timing in Huntington's disease. *Brain Lang,* 47**,** 182-96.

HERTRICH, I. & ACKERMANN, H. 1995. Gender-Specific Vocal Dysfunctions in Parkinsons-Disease - Electroglottographic and Acoustic Analyses. *Annals of Otology Rhinology and Laryngology,* 104**,** 197-202.

HIROSE, H. 1986. Pathophysiology of Motor Speech Disorders (Dysarthria). *Folia Phoniatrica,* 38**,** 61-88.

HIROSE, H., KIRITANI, S., USHIJIMA, T., YOSHIOKA, H. & SAWASHIMA, M. 1981. Patterns of dysarthric movements in patients with parkinsonism. *Folia Phoniatr (Basel),* 33**,** 204-15.

HO, A. K., BRADSHAW, J. L. & IANSEK, R. 2008. For better or worse: The effect of levodopa on speech in Parkinson's disease. *Movement Disorders,* 23**,** 574-580.

HOEHN, M. M. 1992. The natural history of Parkinson's disease in the pre-levodopa and post-levodopa eras. *Neurol Clin,* 10**,** 331-9.

HOEHN, M. M. & YAHR, M. D. 1967. Parkinsonism: onset, progression and mortality. *Neurology,* 17**,** 427-42.

HOIT, J. D., WATSON, P. J., HIXON, K. E., MCMAHON, P. & JOHNSON, C. L. 1994. Age and Velopharyngeal Function during Speech Production. *Journal of Speech and Hearing Research,* 37**,** 295-302.

HOODIN, R. B. & GILBERT, H. R. 1989. Nasal airflows in parkinsonian speakers. *J Commun Disord,* 22**,** 169-80.

HORII, Y. 1980. An Accelerometric Approach to Nasality Measurement - a Preliminary-Report. *Cleft Palate Journal,* 17**,** 254-261.

HORNYKIEWICZ, O. 2008. Basic research on dopamine in Parkinson's disease and the discovery of the nigrostriatal dopamine pathway: the view of an eyewitness. *Neurodegener Dis,* 5**,** 114-7.

HSU, C. W., CHANG, C. C. & LIN, C. J. 2010. The elements of statistical learning: Datamining, inference, and prediction. Taipei: National Taiwan University.

HUBER, J. E. & DARLING, M. 2011. Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: respiratory physiologic and linguistic considerations. *J Speech Lang Hear Res,* 54**,** 33-46.

HUBER, J. E., STATHOPOULOS, E. T., RAMIG, L. O. & LANCASTER, S. L. 2003. Respiratory function and variability in individuals with Parkinson disease: Pre- and post-Lee Silverman voice treatment. *Journal of Medical Speech-Language Pathology,* 11**,** 185-201.

HUGHES, A. J., BEN-SHLOMO, Y., DANIEL, S. E. & LEES, A. J. 1992. What features improve the accuracy of clinical diagnosis in Parkinson's disease: a clinicopathologic study. *Neurology,* 42**,** 1142-6.

HUNKER, C. J. & ABBS, J. H. 1984. Physiological analyses of parkinsonian tremors in the orofacial system. *In:* MCNEIL, M. R., ROSENBEK, J. C. & ARONSON, A. E. (eds.) *The dysarthrias.* Austin: Pro-Ed.

HUNTINGTON-STUDY-GROUP 1996. Unified Huntington's Disease Rating Scale: reliability and consistency. . *Mov Disord,* 11**,** 136-42.

HUSSAIN, I. F., BRADY, C. M., SWINN, M. J., MATHIAS, C. J. & FOWLER, C. J. 2001. Treatment of erectile dysfunction with sildenafil citrate (Viagra) in parkinsonism due to Parkinson's disease or multiple system atrophy with observations on orthostatic hypotension. *J Neurol Neurosurg Psychiatry,* 71**,** 371-4.

CHANG, S. E., KENNEY, M. K., LOUCKS, T. M., POLETTO, C. J. & LUDLOW, C. L. 2009. Common neural substrates support speech and non-speech vocal tract gestures. *Neuroimage,* 47**,** 314-25.

CHENAUSKY, K., MACAUSLAN, J. & GOLDHOR, R. 2011. Acoustic Analysis of PD Speech. *Parkinsons Disease.*

INDEFREY, P. & LEVELT, W. J. 2004. The spatial and temporal signatures of word production components. *Cognition,* 92**,** 101-44.

JAN, C., FRANCOIS, C., TANDE, D., YELNIK, J., TREMBLAY, L., AGID, Y. & HIRSCH, E. 2000. Dopaminergic innervation of the pallidum in the normal state, in MPTP-treated monkeys and in parkinsonian patients. *Eur J Neurosci,* 12**,** 4525-35.

JANKOVIC, J. 2008. Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry,* 79**,** 368-76.

JANKOVIC, J., GILDEN, J. L., HINER, B. C., KAUFMANN, H., BROWN, D. C., COGHLAN, C. H., RUBIN, M. & FOUAD-TARAZI, F. M. 1993. Neurogenic orthostatic hypotension: a double-blind, placebo-controlled study with midodrine. *Am J Med,* 95**,** 38-48.

JOOS, U., WERMKER, K., KRUSE-LOESLER, B. & KLEINHEINZ, J. 2006. Influence of treatment concept, velopharyngoplasty, gender and age on hypernasality in patients with cleft lip, alveolus and palate. *Journal of Cranio-Maxillofacial Surgery,* 34**,** 472-477.

KATAOKA, R., MICHI, K., OKABE, K., MIURA, T. & YOSHIDA, H. 1996. Spectral properties and quantitative evaluation of hypernasality in vowels. *Cleft Palate-Craniofacial Journal,* 33**,** 43-50.

KEMPSTER, G. B., GERRATT, B. R., ABBOTT, K. V., BARKMEIER-KRAEMER, J. & HILLMAN, R. E. 2009. Consensus Auditory-Perceptual Evaluation of Voice: Development of a Standardized Clinical Protocol. *American Journal of Speech-Language Pathology,* 18**,** 124-132.

KENT, R. D. 2000. Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders,* 33**,** 391-428.

KENT, R. D., KENT, J. F., DUFFY, J. R., THOMAS, J. E., WEISMER, G. & STUNTEBECK, S. 2000a. Ataxic dysarthria. *Journal of Speech Language and Hearing Research,* 43**,** 1275-1289.

KENT, R. D., KENT, J. F., ROSENBEK, J. C., VORPERIAN, H. K. & WEISMER, G. 1997. A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatrica Et Logopaedica,* 49**,** 63-82.

KENT, R. D., KENT, J. F. & WEISMER, G. 2000b. What dysarthrias can tell us about the neural control of speech. *Journal of Phonetics,* 28**,** 273-302.

KENT, R. D. & READ, C. 2002. *The Acoustic Analysis of Speech,* Albany, Thomson Learning.

KENT, R. D. & ROSENBEK, J. C. 1982. Prosodic Disturbance and Neurologic Lesion. *Brain and Language,* 15**,** 259-291.

KENT, R. D., VORPERIAN, H. K., KENT, J. F. & DUFFY, J. R. 2003. Voice dysfunction in dysarthria: application of the Multi-Dimensional Voice Program (TM). *Journal of Communication Disorders,* 36**,** 281-306.

KENT, R. D., WEISMER, G., KENT, J. F., VORPERIAN, H. K. & DUFFY, J. R. 1999. Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders,* 32**,** 141-186.

KLUIN, K. J., FOSTER, N. L., BERENT, S. & GILMAN, S. 1993. Perceptual analysis of speech disorders in progressive supranuclear palsy. *Neurology,* 43**,** 563-6.

KLUIN, K. J., GILMAN, S., LOHMAN, M. & JUNCK, L. 1996. Characteristics of the dysarthria of multiple system atrophy. *Arch Neurol,* 53**,** 545-8.

KUEHN, D. P. & MOLLER, K. T. 2000. The state of the art: Speech and language issues in the cleft palate population. *Cleft Palate-Craniofacial Journal,* 37**,** 348-348.

LEE, A. S., CIOCCA, V. & WHITEHILL, T. L. 2003. Acoustic correlates of hypernasality. *Clin Linguist Phon,* 17**,** 259-64.

LEMOINE, C. & BLOCH, B. 1995. D1 and D2 Dopamine-Receptor Gene-Expression in the Rat Striatum - Sensitive Crna Probes Demonstrate Prominent Segregation of D1 and D2 Messenger-Rnas in Distinct Neuronal Populations of the Dorsal and Ventral Striatum. *Journal of Comparative Neurology,* 355**,** 418-426.

LEOPOLD, N. A. & KAGEL, M. C. 1985. Dysphagia in Huntington's disease. *Arch Neurol,* 42**,** 57-60.

LEVELT, W. J. M. 1989. *Speaking: from intention to articulation,* Cambridge, The MIT Press.

LINDBLOM, B. E. & SUNDBERG, J. E. 1971. Acoustical Consequences of Lip, Tongue, Jaw, and Larynx Movement. *Journal of the Acoustical Society of America,* 50**,** 1166-&.

LITTLE, M. A., MCSHARRY, P. E., HUNTER, E. J., SPIELMAN, J. & RAMIG, L. O. 2009. Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease. *Ieee Transactions on Biomedical Engineering,* 56**,** 1015-1022.

LITVAN, I., AGID, Y., CALNE, D., CAMPBELL, G., DUBOIS, B., DUVOISIN, R. C., GOETZ, C. G., GOLBE, L. I., GRAFMAN, J., GROWDON, J. H., HALLETT, M., JANKOVIC, J., QUINN, N. P., TOLOSA, E. & ZEE, D. S. 1996. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology,* 47**,** 1-9.

LITZAW, L. L. & DALSTON, R. M. 1992. The Effect of Gender Upon Nasalance Scores among Normal Adult Speakers. *Journal of Communication Disorders,* 25**,** 55-64.

LOGEMANN, J. A., FISHER, H. B., BOSHES, B. & BLONSKY, E. R. 1978. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *J Speech Hear Disord,* 43**,** 47-57.

LUDLOW, C. L. & BASICH, C. J. 1983. The results of acoustic and perceptual assessment of two types of dysarthria. *In:* BERRY, W. R. (ed.) *Clinical dysarthria.* San Diego: College-Hill Press.

LUDLOW, C. L. & BASICH, C. J. 1984. Relationship between perceptual ratings and acoustic measures of hypokinetic speech. *In:* MCNEIL, M. R. (ed.) *The dysarthrias: physiologic, acoustics, perception, management.* Austin: Pro-Ed.

MACDONALD, M. E., VONSATTEL, J. P., SHRINIDHI, J., COUROPMITREE, N. N., CUPPLES, L. A., BIRD, E. D., GUSELLA, J. F. & MYERS, R. H. 1999. Evidence for the GluR6 gene associated with younger onset age of Huntington's disease. *Neurology,* 53**,** 1330-2.

MALLET, L., SCHUPBACH, M., N'DIAYE, K., REMY, P., BARDINET, E., CZERNECKI, V., WELTER, M. L., PELISSOLO, A., RUBERG, M., AGID, Y. & YELNIK, J. 2007. Stimulation of subterritories of the subthalamic nucleus reveals its role in the integration of the emotional and motor aspects of behavior. *Proc Natl Acad Sci U S A,* 104**,** 10661-6.

MCCLEAN, M. D., BEUKELMAN, D. R. & YORKSTON, K. M. 1987. Speech-Muscle Visuomotor Tracking in Dysarthric and Nonimpaired Speakers. *Journal of Speech and Hearing Research,* 30**,** 276-282.

MIDI, I., DOGAN, M., KOSEOGLU, M., CAN, G., SEHITOGLU, M. A. & GUNAL, D. I. 2008. Voice abnormalities and their relation with motor dysfunction in Parkinson's disease. *Acta Neurol Scand,* 117**,** 26-34.

MINK, J. W. 1996. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol,* 50**,** 381-425.

MUELLER, P. B. 1971. Parkinson's disease: motor-speech behavior in a selected group of patients. *Folia Phoniatr (Basel),* 23**,** 333-46.

MURDOCH, B. E., CHENERY, H. J., STOKES, P. D. & HARDCASTLE, W. J. 1991. Respiratory kinematics in speakers with cerebellar disease. *J Speech Hear Res,* 34**,** 768-80.

MURRY, T. 1983. The production of stress in three types of dysarthric speech. *In:* BERRY, W. R. (ed.) *Clinical dysarthria.* San Diego: College-Hill Press.

NETSELL, R., DANIEL, B. & CELESIA, G. G. 1975. Acceleration and weakness in parkinsonian dysarthria. *J Speech Hear Disord,* 40**,** 170-8.

ORIEUX, G., FRANCOIS, C., FEGER, J., YELNIK, J., VILA, M., RUBERG, M., AGID, Y. & HIRSCH, E. C. 2000. Metabolic activity of excitatory parafascicular and pedunculopontine inputs to the subthalamic nucleus in a rat model of Parkinson's disease. *Neuroscience,* 97**,** 79-88.

PARENT, A. 1990. Extrinsic connections of the basal ganglia. *Trends Neurosci,* 13**,** 254-8.

PARENT, A. & HAZRATI, L. N. 1995. Functional-Anatomy of the Basal Ganglia .1. The Cortico-Basal Ganglia-Thalamo-Cortical Loop. *Brain Research Reviews,* 20**,** 91-127.

PATEL, R. & CAMPELLONE, P. 2009. Acoustic and Perceptual Cues to Contrastive Stress in Dysarthria. *Journal of Speech Language and Hearing Research,* 52**,** 206-222.

PAULSEN, J. S., HOTH, K. F., NEHL, C. & STIERMAN, L. 2005. Critical periods of suicide risk in Huntington's disease. *Am J Psychiatry,* 162**,** 725-31.

PAYAN, C. A., VIALLET, F., LANDWEHRMEYER, B. G., BONNET, A. M., BORG, M., DURIF, F., LACOMBLEZ, L., BLOCH, F., VERNY, M., FERMANIAN, J., AGID, Y., LUDOLPH, A. C., LEIGH, P. N., BENSIMON, G. & GROUP, N. S. 2011. Disease severity and progression in progressive supranuclear palsy and multiple system atrophy: validation of the NNIPPS--Parkinson Plus Scale. *PLoS One,* 6**,** e22293.

PETERSON-FALZONE, S., HARDIN-JONES, M., KARNELL, M. & MC. WILLIAMS, B. 2001. *Cleft palate speech,* St. Louis.

PINTO, S., GENTIL, M., FRAIX, V., BENABID, A. L. & POLLAK, P. 2003. Bilateral subthalamic stimulation effects on oral force control in Parkinson's disease. *J Neurol,* 250**,** 179-87.

POOLE, M. L., WEE, J. S., FOLKER, J. E., CORBEN, L. A., DELATYCKI, M. B. & VOGEL, A. P. 2015. Nasality in Friedreich ataxia. *Clinical Linguistics & Phonetics,* 29**,** 46-58.

POSTUMA, R. B., LANG, A. E., GAGNON, J. F., PELLETIER, A. & MONTPLAISIR, J. Y. 2012. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. *Brain,* 135**,** 1860-70.

PRASANNA, S. R. M., REDDY, B. V. S. & KRISHNAMOORTHY, P. 2009. Vowel Onset Point Detection Using Source, Spectral Peaks, and Modulation Spectrum Energies. *Ieee Transactions on Audio Speech and Language Processing,* 17**,** 556-565.

PRICE, C. J. 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann N Y Acad Sci,* 1191**,** 62-88.

QUINN, N. 1995. Parkinsonism--recognition and differential diagnosis. *BMJ,* 310**,** 447-52.

RAMIG, L. A. 1986. Acoustic analyses of phonation in patients with Huntington's disease. Preliminary report. *Ann Otol Rhinol Laryngol,* 95**,** 288-93.

RAMIG, L. A. & RINGEL, R. L. 1983. Effects of physiological aging on selected acoustic characteristics of voice. *J Speech Hear Res,* 26**,** 22-30.

RAMIG, L. A., SCHERER, R. C., TITZE, I. R. & RINGEL, S. P. 1988. Acoustic analysis of voices of patients with neurologic disease: rationale and preliminary data. *Ann Otol Rhinol Laryngol,* 97**,** 164-72.

REHMAN, H. U. 2000. Progressive supranuclear palsy. *Postgrad Med J,* 76**,** 333-6.

RIECKER, A., MATHIAK, K., WILDGRUBER, D., ERB, M., HERTRICH, I., GRODD, W. & ACKERMANN, H. 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology,* 64**,** 700-6.

RIECKER, A., WILDGRUBER, D., GRODD, W. & ACKERMANN, H. 2002. Reorganization of speech production at the motor cortex and cerebellum following capsular infarction: a follow-up functional magnetic resonance imaging study. *Neurocase,* 8**,** 417-23.

RODRIGUEZ-OROZ, M. C., JAHANSHAHI, M., KRACK, P., LITVAN, I., MACIAS, R., BEZARD, E. & OBESO, J. A. 2009. Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms. *Lancet Neurol,* 8**,** 1128-39.

ROSEN, K. M., KENT, R. D., DELANEY, A. L. & DUFFY, J. R. 2006. Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers. *J Speech Lang Hear Res,* 49**,** 395-411.

ROSS, C. A. & TABRIZI, S. J. 2011. Huntington's disease: from molecular pathogenesis to clinical treatment. *Lancet Neurol,* 10**,** 83-98.

RUANAIDH, J. J. O. & FITZGERALD, W. J. 1996. *Numerical bayesian methods applied to signal processing,* Berlin, Springer-Verlag.

RUBINSZTEIN, D. C., LEGGO, J., COLES, R., ALMQVIST, E., BIANCALANA, V., CASSIMAN, J. J., CHOTAI, K., CONNARTY, M., CRAUFORD, D., CURTIS, A., CURTIS, D., DAVIDSON, M. J., DIFFER, A. M., DODE, C., DODGE, A., FRONTALI, M., RANEN, N. G., STINE, O. C., SHERR, M., ABBOTT, M. H., FRANZ, M. L., GRAHAM, C. A., HARPER, P. S., HEDREEN, J. C., HAYDEN, M. R. & ET AL. 1996. Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *Am J Hum Genet,* 59**,** 16-22.

RUSZ, J., BONNET, C., KLEMPIR, J., TYKALOVA, T., BABOROVA, E., NOVOTNY, M., RULSEH, A. & RUZICKA, E. 2015. Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Journal of Neurology,* 262**,** 992-1001.

RUSZ, J., CMEJLA, R., RUZICKOVA, H., KLEMPIR, J., MAJEROVA, V., PICMAUSOVA, J., ROTH, J. & RUZICKA, E. 2011a. Acoustic Assessment of Voice and Speech Disorders in Parkinson's Disease Through Quick Vocal Test. *Movement Disorders,* 26**,** 1951-1952.

RUSZ, J., CMEJLA, R., RUZICKOVA, H., KLEMPIR, J., MAJEROVA, V., PICMAUSOVA, J., ROTH, J. & RUZICKA, E. 2013a. Evaluation of speech impairment in early stages of Parkinson's disease: a prospective study with the role of pharmacotherapy. *Journal of Neural Transmission,* 120**,** 319-329.

RUSZ, J., CMEJLA, R., RUZICKOVA, H. & RUZICKA, E. 2011b. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *Journal of the Acoustical Society of America,* 129**,** 350-367.

RUSZ, J., CMEJLA, R., TYKALOVA, T., RUZICKOVA, H., KLEMPIR, J., MAJEROVA, V., PICMAUSOVA, J., ROTH, J. & RUZICKA, E. 2013b. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *Journal of the Acoustical Society of America,* 134**,** 2171-2181.

RUSZ, J., KLEMPIR, J., BABOROVA, E., TYKALOVA, T., MAJEROVA, V., CMEJLA, R., RUZICKA, E. & ROTH, J. 2013c. Objective acoustic quantification of phonatory dysfunction in Huntington's disease. *PLoS One,* 8**,** e65881.

RUSZ, J., KLEMPIR, J., TYKALOVA, T., BABOROVA, E., CMEJLA, R., RUZICKA, E. & ROTH, J. 2014a. Characteristics and occurrence of speech impairment in Huntington's disease: possible influence of antipsychotic medication. *Journal of Neural Transmission,* 121**,** 1529-1539.

RUSZ, J., KLEMPIR, J., TYKALOVA, T., BABOROVA, E., CMEJLA, R., RUZICKA, E. & ROTH, J. 2014b. Characteristics and occurrence of speech impairment in Huntington's disease: possible influence of antipsychotic medication. *J Neural Transm (Vienna),* 121**,** 1529-39.

RUSZ, J., SAFT, C., SCHLEGEL, U., HOFFMAN, R. & SKODDA, S. 2014c. Phonatory dysfunction as a preclinical symptom of Huntington disease. *PLoS One,* 9**,** e113412.

SAPIR, S., RAMIG, L. O., SPIELMAN, J. L. & FOX, C. 2010. Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech. *Journal of Speech Language and Hearing Research,* 53**,** 114-125.

SELL, D., HARDING, A. & GRUNWELL, P. 1999. GOS.SP.ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). *Int J Lang Commun Disord,* 34**,** 17-33.

SHERRARD, K. C., MARQUARDT, T. P. & CANNITO, M. P. 2000. Phonatory and temporal aspects of spasmodic dysphonia and pseudobulbar dysarthria: An acoustic analysis. *Journal of Medical Speech-Language Pathology,* 8**,** 271-277.

SCHLESINGER, M. I. & HLAVAC, V. 2002. *Ten lectures on statistical and structural pattern recognition,* Dordrecht, Kluwer Academic Press.

SCHRAG, A., BEN-SHLOMO, Y. & QUINN, N. P. 1999. Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. *Lancet,* 354**,** 1771-1775.

SCHULZ, G. M. & GRANT, M. K. 2000. Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: a review of the literature. *J Commun Disord,* 33**,** 59-88.

SIMONYAN, K. & HORWITZ, B. 2011. Laryngeal Motor Cortex and Control of Speech in Humans. *Neuroscientist,* 17**,** 197-208.

SKODDA, S. 2011. Aspects of speech rate and regularity in Parkinson's disease. *Journal of the Neurological Sciences,* 310**,** 231-236.

SKODDA, S., RINSCHE, H. & SCHLEGEL, U. 2009. Progression of Dysprosody in Parkinson's Disease Over Time-A Longitudinal Study. *Movement Disorders,* 24**,** 716-722.

SKODDA, S. & SCHLEGEL, U. 2008. Speech rate and rhythm in Parkinson's disease. *Movement Disorders,* 23**,** 985-992.

SKODDA, S., SCHLEGEL, U., HOFFMANN, R. & SAFT, C. 2014a. Impaired motor speech performance in Huntington's disease. *Journal of Neural Transmission,* 121**,** 399-407.

SKODDA, S., SCHLEGEL, U., HOFFMANN, R. & SAFT, C. 2014b. Impaired motor speech performance in premotor stages of Huntington's disease (HD) over time - A longitudinal investigation. *Movement Disorders,* 29**,** S217-S217.

SKODDA, S., VISSER, W. & SCHLEGEL, U. 2010. Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *Journal of Neural Transmission,* 117**,** 197-205.

SKODDA, S., VISSER, W. & SCHLEGEL, U. 2011. Acoustical Analysis of Speech in Progressive Supranuclear Palsy. *Journal of Voice,* 25**,** 725-731.

SMITH, Y., WICHMANN, T. & DELONG, M. R. 1994. Synaptic innervation of neurones in the internal pallidal segment by the subthalamic nucleus and the external pallidum in monkeys. *J Comp Neurol,* 343**,** 297-318.

SOLOMON, N. P. & HIXON, T. J. 1993. Speech Breathing in Parkinsons-Disease. *Journal of Speech and Hearing Research,* 36**,** 294-310.

SOLOMON, N. P., ROBIN, D. A. & LUSCHEI, E. S. 2000. Strength, endurance, and stability of the tongue and hand in Parkinson disease. *J Speech Lang Hear Res,* 43**,** 256-67.

SOROS, P., SOKOLOFF, L. G., BOSE, A., MCINTOSH, A. R., GRAHAM, S. J. & STUSS, D. T. 2006. Clustered functional MRI of overt speech production. *Neuroimage,* 32**,** 376-87.

STEVENS, K. N. 2000. *Acoustic Phonetics,* Cambridge, The MIT Press.

STOUTEN, V. & VAN HAMME, H. 2009. Automatic voice onset time estimation from reassignment spectra. *Speech Communication,* 51**,** 1194-1205.

THEODOROS, D. G., MURDOCH, B. E. & C., T. E. 1995. Hypernasality in Parkinson Disease: A Perceptual and Physiological Analysis *Journal of Medical Speech-Language Pathology,* 3**,** 73-84.

TISON, F., YEKHLEF, F., CHRYSOSTOME, V. & SOURGEN, C. 2000. Prevalence of multiple system atrophy. *Lancet,* 355**,** 495-6.

TITZE, I. R. 1994. Workshop on acoustic voice analysis. *Workshop on acoustic voice analysis.* Denver: National Center for Speech and Voice.

TRIPOLITI, E. 2010. *Effects of deep brain stimulation on speech in patients with Parkinson's disease and dystonia.* PhD, University College London.

TSANAS, A., LITTLE, M. A., MCSHARRY, P. E., SPIELMAN, J. & RAMIG, L. O. 2012. Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease. *Ieee Transactions on Biomedical Engineering,* 59**,** 1264-1271.

TYKALOVA, T., RUSZ, J., CMEJLA, R., RUZICKOVA, H. & RUZICKA, E. 2014. Acoustic Investigation of Stress Patterns in Parkinson's Disease. *Journal of Voice,* 28.

VAN DEN EEDEN, S. K., TANNER, C. M., BERNSTEIN, A. L., FROSS, R. D., LEIMPETER, A., BLOCH, D. A. & NELSON, L. M. 2003. Incidence of Parkinson's disease: Variation by age, gender, and Race/Ethnicity. *American Journal of Epidemiology,* 157**,** 1015-1022.

VANACORE, N., BONIFATI, V., FABBRINI, G., COLOSIMO, C., DE MICHELE, G., MARCONI, R., NICHOLL, D., LOCURATOLO, N., TALARICO, G., ROMANO, S., STOCCHI, F., BONUCCELLI, U., DE MARI, M., VIEREGGE, P., MECO, G. & EUROPEAN STUDY GROUP ON ATYPICAL, P. 2001. Epidemiology of multiple system atrophy. ESGAP Consortium. European Study Group on Atypical Parkinsonisms. *Neurol Sci,* 22**,** 97-9.

VILA, M., LEVY, R., HERRERO, M. T., RUBERG, M., FAUCHEUX, B., OBESO, J. A., AGID, Y. & HIRSCH, E. C. 1997. Consequences of nigrostriatal denervation on the functioning of the basal ganglia in human and nonhuman primates: An in situ hybridization study of cytochrome oxidase subunit I mRNA. *Journal of Neuroscience,* 17**,** 765-773.

VOGEL, A. P., IBRAHIM, H. M., REILLY, S. & KILPATRICK, N. 2009. A Comparative Study of Two Acoustic Measures of Hypernasality. *Journal of Speech Language and Hearing Research,* 52**,** 1640-1651.

VOLAITIS, L. E. & MILLER, J. L. 1992. Phonetic prototypes: influence of place of articulation and speaking rate on the internal structure of voicing categories. *J Acoust Soc Am,* 92**,** 723-35.

VONSATTEL, J. P. & DIFIGLIA, M. 1998. Huntington disease. *J Neuropathol Exp Neurol,* 57**,** 369-84.

WALKER, F. O. 2007. Huntington's disease. *Lancet,* 369**,** 218-28.

WANG, Y. T., KENT, R. D., DUFFY, J. R., THOMAS, J. E. & WEISMER, G. 2004. Alternating motion rate as an index of speech motor disorder in traumatic brain injury. *Clin Linguist Phon,* 18**,** 57-84.

WATKINS, K. E., SMITH, S. M., DAVIS, S. & HOWELL, P. 2008. Structural and functional abnormalities of the motor system in developmental stuttering. *Brain,* 131**,** 50-9.

WEINER, P., INZELBERG, R., DAVIDOVICH, A., NISIPEANU, P., MAGADLE, R., BERAR-YANAY, N. & CARASSO, R. L. 2002. Respiratory muscle performance and the Perception of dyspnea in Parkinson's disease. *Can J Neurol Sci,* 29**,** 68-72.

WEISMER, G. 1984. Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. *In:* MCNEIL, M. R., ROSENBEK, J. C. & ARONSON, A. E. (eds.) *The dysarthrias: physiology, acoustics, perception, management.* San Diego: College-Hill Press.

WENNING, G. K., BEN SHLOMO, Y., MAGALHAES, M., DANIEL, S. E. & QUINN, N. P. 1994. Clinical features and natural history of multiple system atrophy. An analysis of 100 cases. *Brain,* 117 (Pt 4)**,** 835-45.

WENNING, G. K., COLOSIMO, C., GESER, F. & POEWE, W. 2004. Multiple system atrophy. *Lancet Neurol,* 3**,** 93-103.

WEXLER, N. S., LORIMER, J., PORTER, J., GOMEZ, F., MOSKOWITZ, C., SHACKELL, E., MARDER, K., PENCHASZADEH, G., ROBERTS, S. A., GAYAN, J., BROCKLEBANK, D., CHERNY, S. S., CARDON, L. R., GRAY, J., DLOUHY, S. R., WIKTORSKI, S., HODES, M. E., CONNEALLY, P. M., PENNEY, J. B., GUSELLA, J., CHA, J. H., IRIZARRY, M., ROSAS, D., HERSCH, S., HOLLINGSWORTH, Z., MACDONALD, M., YOUNG, A. B., ANDRESEN, J. M., HOUSMAN, D. E., DE YOUNG, M. M., BONILLA, E., STILLINGS, T., NEGRETTE, A., SNODGRASS, S. R., MARTINEZ-JAURRIETA, M. D., RAMOS-ARROYO, M. A., BICKHAM, J., RAMOS, J. S., MARSHALL, F., SHOULSON, I., REY, G. J., FEIGIN, A., ARNHEIM, N., ACEVEDO-CRUZ, A., ACOSTA, L., ALVIR, J., FISCHBECK, K., THOMPSON, L. M., YOUNG, A., DURE, L., O'BRIEN, C. J., PAULSEN, J., BRICKMAN, A., KRCH, D., PEERY, S., HOGARTH, P., HIGGINS, D. S., JR., LANDWEHRMEYER, B. & PROJECT, U. S.-V. C. R. 2004. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A,* 101**,** 3498-503.

WILDGRUBER, D., ACKERMANN, H. & GRODD, W. 2001. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *Neuroimage,* 13**,** 101-9.

WILLIAMS, D. R. & LEES, A. J. 2009. Progressive supranuclear palsy: clinicopathological concepts and diagnostic challenges. *Lancet Neurology,* 8**,** 270-279.

YELNIK, J. 2002. Functional anatomy of the basal ganglia. *Movement Disorders,* 17**,** S15-S21.

YELNIK, J. 2008. Modeling the organization of the basal ganglia. *Revue Neurologique,* 164**,** 969-976.

YOSHIDA, H., FURUYA, Y., SHIMODAIRA, K., KANAZAWA, T., KATAOKA, R. & TAKAHASHI, K. 2000. Spectral characteristics of hypernasality in maxillectomy patients. *Journal of Oral Rehabilitation,* 27**,** 723-730.

ZIEGLER, W. & VON CRAMON, D. 1986. Spastic dysarthria after acquired brain injury: an acoustic study. *Br J Disord Commun,* 21**,** 173-87.

ZIEGLER, W. & WESSEL, K. 1996. Speech timing in ataxic disorders: Sentence production and rapid repetitive articulation. *Neurology,* 47**,** 208-214.

ZINGLER, V. C., STRUPP, M., JAHN, K. & BRANDT, T. 2005. Hemihypomimia in Parkinson's disease. *European Neurology,* 53**,** 92-93.

ZWIRNER, P., MURRY, T. & WOODSON, G. E. 1993. Perceptual-acoustic relationships in spasmodic dysphonia. *J Voice,* 7**,** 165-71.

LIST OF PUBLICATIONS

Journals with impact factor

Rusz, J. - Bonnet, C. - Klempíř, J. - Tykalová, T. - Baborová, E. - et al.

Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy

In: Journal of Neurology. 2015, vol. 262, no. 4, p. 992-1001. ISSN 0340-5354. **[12.5%]**

Novotný, M. - Rusz, J. - Čmejla, R. - Růžička, E.

Automatic Evaluation of Articulatory Disorders in Parkinson's Disease

In: IEEE Transactions on Audio Speech and Language Processing. 2014, vol. 22, no. 9, p. 1366-1378. ISSN 1558-7916. **[25%]**

Journals with peer view

Novotný, M. - Rusz, J. - Čmejla, R.

Automatická segmentace hlásek při rychlém opakování slabik (/pa/ /ta/ /ka/) u hypokinetické dysartrie

In: Akustické listy. 2011, roč. 17, č. 4, s. 10-16. ISSN 1212-4702. **[33%]**

Novotný, M. - Rusz, J. - Čmejla, R.

AUTOMATIC SEGMENTATION OF PHONEMES DURING THE FAST REPETITION OF (/PA/-/TA/-/KA/) SYLLABLES IN A SPEECH AFFECTED BY HYPOKINETIC DYSARTHRIA

In: Lékař a technika. 2012, roč. 42, č. 2, s. 81-84. ISSN 0301-5491. **[33%]**

Publications excerpted in Web of Science or Scopus

Novotný, M. - Pospíšil, J. - Čmejla, R. - Rusz, J.

Automatic detection of voice onset time in dysarthric speech

In: 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing - Proceedings. New Jersey: IEEE Signal Processing Society, 2015, p. 4340-4344. ISBN 978-1-4673-6997-8. **[25%]**

Other publications

Novotný, M. - Rusz, J. - Čmejla, R.

Hypernazalita v dysartrických promluvách

In: V. Letní doktorandské dny 2015. Praha: ČVUT FEL, Katedra teorie obvodů, 2015, s. 73-78. ISBN 978-80-01-05749-0. **[33%]**

Novotný, M.

Analysis of neurodegenerative diseases

In: PROCEEDINGS ABSTRACTS on II. CZECH-ITALY WORKSHOP ON BIOLOGICAL SIGNALS. Prague: CTU, Faculty of Electrical Engineering, Department of Circuit Theory, 2014, ISBN 978-80-01-05671-4. **[100%]**

Novotný, M.

Analysis of Speech Dysarthria

In: PROCEEDINGS ABSTRACTS on II. CZECH-GERMAN WORKSHOP ON SPEECH PATHOLOGY AND BIOLOGICAL SIGNALS. Prague: CTU, Faculty of Electrical Engineering, Department of Circuit Theory, 2014, ISBN 978-80-01-05670-7. **[100%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Hypernazalita v dysartrických promluvách

In: IV. Letní doktorandské dny 2014. Praha: ČVUT FEL, Katedra teorie obvodů, 2014, díl 4, s. 87-91. ISBN 978-80-01-05506-9. **[33%]**

Tykalová, T. - Rusz, J. - Čmejla, R. - Novotný, M.

Využití akustických analýz pro hodnocení hlasu a řeči u Huntingtonovy choroby

In: 22nd Annual Conference Proceedings Technical Computing Bratislava 2014. Praha: Humusoft, 2014, ISSN 2336-1662. ISBN 978-80-7080-898-6. **[25%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Grafické rozhraní pro přípravu a kontrolu dat

In: 21th Annual Conference Proceedings Technical Computing Prague 2013. Praha: Humusoft, 2013, s. 47-50. ISSN 2336-1662. ISBN 978-80-7080-863-4. **[33%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Charakteristiky promluv pacientů s Parkinsonovou nemocí extrahované z řečové diadochokinetické úlohy

In: III. LETNÍ DOKTORANDSKÉ DNY 2013. Praha: ČVUT, Fakulta elektrotechnická, 2013, díl 3, s. 122-127. ISBN 978-80-01-05251-8. **[33%]**

Rusz, J. - Novotný, M. - Tykalová, T.

Akustické analýzy řeči u Parkinsonovy nemoci

In: Sborník 85. akustického semináře. Praha: Nakladatelství ČVUT, 2012, s. 45-50. ISBN 978-80-01-05133-7. **[33%]**

Čmejla, R. - Rusz, J. - Bauer, L. - Lustyk, T. - Nejepsová, M. - et al.

Analýza patologického hlasu a řeči v laboratoři SAMI ČVUT

In: Novinky ve foniatrii. Praha 5, Na bělidle 34, 150 00: Nakladatelství Galén, 2012, s. 28-30. ISBN 978-80-7262-940-4. **[25%]**

Novotný, M.

Automatická segmentace hlásek při rychlém opakovaní slabik (/pa/-/ta/-/ka/) u hypokinetické dysartrie

In: LETNÍ DOKTORANDSKÉ DNY 2012. Praha: ČVUT, 2012, s. 92-97. ISBN 978-80-01-05050-7. **[100%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Automatické hledání významných pozic v Parkinsonických promluvách založených na rychlém opakováni slabik /pa/ - /ta/ - /ka/

In: 20th Annual Conference Proceeding's Technical Computing Bratislava 2012. Praha: Humusoft, 2012, díl 57, s. 1-5. ISBN 978-80-970519-4-5. **[33%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Automatické hodnocení poruch artikulace založené na rychlém opakování slabik u pacientů trpících Parkinsonovou nemocí

In: Sborník 85. akustického semináře. Praha: Nakladatelství ČVUT, 2012, s. 35-40. ISBN 978-80-01-05133-7. **[33%]**

Novotný, M. - Rusz, J. - Čmejla, R.

Evaluation of Parkinsonian speech attributes obtained from utterances based upon the fast /pa/ - /ta/ - /ka/ syllables repetition

In: Czech-German Workshop on Speech Pathology and Biological Signals - Proceedings. Prague: CTU, Faculty of Electrical Engineering, Department of Circuit Theory, 2012, p. 57-58. ISBN 978-80-01-05164-1. **[33%]**

Novotný, M. - Rusz, J. - Čmejla, R. - Růžičková, H.

HODNOCENÍ HYPOKINETICKÝCH PROMLUV U PACIENTŮ S PARKINSONOVOU NEMOCÍ POMOCÍ VOICE ONSET TIME

In: Novinky ve foniatrii. Praha 5, Na bělidle 34, 150 00: Nakladatelství Galén, 2012, s. 118-120. ISBN 978-80-7262-940-4. **[25%]**