

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE

Brno, 2018

Tomáš Koutný



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

VYTVOŘENÍ KNIHOVNY PRO PARAMETRIZACI DYSARTRICKÉ ŘEČI V JAZYCE PYTHON

LIBRARY FOR PYTHON USED FOR DYSARTHIC SPEECH PARAMETERIZATION

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Tomáš Koutný

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Ján Mucha

BRNO 2018

Bakalářská práce

bakalářský studijní obor **Teleinformatika**
Ústav telekomunikací

Student: Tomáš Koutný

Ročník: 3

NÁZEV TÉMATU:

ID: 164748

Akademický rok 2017/18

Vytvoření knihovny pro parametrizaci dysartrické řeči v jazyce Python

POKYNY PRO VYPRACOVÁNÍ:

V rámci bakalářské práce budou nastudovány metody analýzy řečových signálů v časové a frekvenční oblasti jako neinvazivní analýza Parkinsonovy nemoci, která se během posledních let dostává do popředí. Dále budou nastudovány moderní parametrizační techniky, které mají za úkol kvantifikovat poškození motorických aspektů řeči. Vybrané parametry budou implementovány v jazyce Python.

DOPORUČENÁ LITERATURA:

[1] GALÁŽ, Z.; MEKYSKA, J.; MŽOUREK, Z.; SMÉKAL, Z.; REKTOROVÁ, I.; ELIÁŠOVÁ, I.; KOŠTÁLOVÁ, M.; MRAČKOVÁ, M.; BERANKOVA, D. Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE, 2016, č. 1, s. 1-17. ISSN: 0169-2607.

[2] GALÁŽ, Z. Quantification of Prosodic Impairment in Patients with Idiopathic Parkinson's Disease. In Proceedings of the 22nd Conference STUDENT EEICT 2016. Brno: 2016. s. 538-542. ISBN: 978-80-214-5350-0.

Termín zadání: 5.2.2018

Termín odevzdání: 29.5.2018

Vedoucí práce: Ing. Ján Mucha

Konzultant:

prof. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Bakalářská práce je zaměřená na parametrizaci dysartrické řeči. V rámci práce je věnována pozornost metodám analýzy řečového signálu u Parkinsonovy nemoci, moderním parametrizačním technikám, které mají za úkol kvantifikovat poškození motorických aspektů řeči, a implementaci vybraných parametrů v jazyce Python. Hlavním cílem této práce bylo vytvoření knihovny parametrů, která je realizována ve vývojovém prostředí PyCharm.

KLÍČOVÁ SLOVA

Hypokinetická dysartrie, Parkinsonova choroba, Parametrizace, Python, Knihovna

ABSTRACT

Bachelor thesis is focused on parameterization of dysarthric speech. Attention is paid to methods of Speech Signal Analysis for Parkinson's disease, modern parameterization techniques, which are designed to quantify the damage of motoric aspects of speech and implementation of selected parameters in Python. The main goal of this work was to create a parameter library that is realized in the PyCharm development environment.

KEYWORDS

Hyperkinetic dysarthria, Parkinson's disease, Parameterization, Python, Library

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma Vytvoření knihovny pro parametrizaci dysartrické řeči v jazyce Python jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne.....

.....

(podpis autora)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu mé bakalářské práce, panu Ing. Jánmu Muchovi, za odborné vedení, konzultace, trpělivost a poskytnutí cenných rad a materiálů potřebných pro realizaci této bakalářské práce.

V Brně dne.....

.....

(podpis autora)

OBSAH

SEZNAM OBRÁZKŮ.....	9
ÚVOD.....	10
1 ÚVOD DO TVORBY ŘEČI A POPIS ŘEČOVÉHO TRAKTU.....	12
1.1 Respirační ústrojí	12
1.2 Fonační ústrojí.....	13
1.3 Artikulační ústrojí	14
2 DYSARTRIE.....	17
2.1 Charakteristika dysartrie a její klasifikace	17
2.2 Hypokinetická dysartrie	19
2.3 Diagnostické postupy	20
2.4 Terapeutické postupy	21
3 METODY ANALÝZY ŘEČOVÉHO SIGNÁLU.....	23
3.1 Předzpracování řečového signálu.....	23
3.1.1 Ustřednění.....	23
3.1.2 Preemfáze.....	24
3.1.3 Segmentace	25
3.2 Parametrizace	26
3.2.1 Parametry popisující fonaci	27
3.2.2 Parametry popisující intenzitu řečového signálu	28
3.2.3 Parametry popisující tempo řečového signálu	30
3.2.4 Parametry popisující hybnost artikulačních orgánů.....	30
3.2.5 Parametry hodnotící kvalitu řečového signálu.....	31
4 PROGRAMOVACÍ JAZYK PYTHON	33
4.1 Python	33
4.1.1 Dekorátor	33
4.2 PyCharm.....	34
4.3 Tvorba knihovny	35
5 NÁVRH A REALIZACE SYSTÉMU	37
5.1 Popis systému.....	37
5.2 Databáze řečových záznamů	37
5.3 Použité parametry.....	38
5.4 Vyhodnocení	39
ZÁVĚR	40
LITERATURA	42

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK	46
SEZNAM PŘÍLOH.....	48
A OBSAH PŘILOŽENÉHO CD	49

SEZNAM OBRÁZKŮ

Obrázek 1.1: Dýchací ústrojí	12
Obrázek 1.2: Schéma činnosti hlasivek	13
Obrázek 1.3: Dutina ústní	15
Obrázek 4.1: Dekorátor.....	34
Obrázek 4.2: Výpis	34
Obrázek 5.1: Výpis NST.....	39

ÚVOD

Řeč představuje biologickou vlastnost člověka, která mu umožňuje navázat kontakt s okolím, vyjadřovat svá přání, pocity a myšlenky a tím dosahovat vytyčených cílů. Její realizace probíhá na základě kooperace respiračního, fonačního a artikulačního ústrojí. Řeč však není pouze záležitostí mluvních orgánů, ale rovněž mozku, který ji řídí. V případě poškození struktur centrální nervové soustavy, které zajišťují tyto funkce, může dojít ke vzniku narušené komunikační schopnosti, tedy i dysartrie. Dysartrie se projevuje obtížemi v oblasti respirace, fonace, artikulace a prozodie. Na základě symptomů a lokalizace léze lze dysartrii dále dělit na několik typů, z nichž jedním je i hypokinetická dysartrie.

Hypokinetická dysartrie bývá typická pro Parkinsonovu chorobu, která postihuje 100 až 150 pacientů na 100 000 obyvatel a projevuje se, mimo jiné, i v komunikačních schopnostech daného jedince. Pro diagnostiku, zjištění a posouzení pokročilosti onemocnění lze využít subjektivní diagnostické postupy, které však mohou být do jisté míry ovlivněny subjektivním pohledem vyšetřujícího, nebo objektivní způsoby získávání a vyhodnocování parametrů řečového signálu.

Hlavním cílem bakalářská práce je vytvoření knihovny pro parametrizaci dysartrické řeči v jazyce Python. V rámci práce je věnována pozornost metodám analýzy řečového signálu u Parkinsonovy nemoci, moderním parametrizačním technikám, které mají za úkol kvantifikovat poškození motorických aspektů řeči, a implementaci vybraných parametrů v jazyce Python.

Bakalářská práce je rozčleněna do pěti kapitol, první čtyři jsou teoretické, zpracované na základě studia odborné literatury a dalších zdrojů. V první kapitole je zpracována problematika tvorby řeči. Blíže jsou zde popsány jednotlivá ústrojí podílející se na vzniku mluvené řeči, tedy ústrojí respirační, fonační a artikulační.

Druhá kapitola se věnuje dysartrii, její charakteristice a klasifikaci, etiologii a symptomům, které jednotlivé typy doprovází. Blíže je zde specifikovaná hypokinetická dysartrie, která je typická pro Parkinsonovu chorobu. Obsahem druhé kapitoly jsou rovněž diagnostické a terapeutické postupy uplatňované u osob s Parkinsonovou chorobou.

Třetí kapitola je zaměřena na metody analýzy akustického signálu. V této kapitole byla pozornost cílena na předzpracování řečového signálu pomocí ustřednění, preemfáze a segmentace a na popis jednotlivých parametrů. Pro lepší přehlednost jsou parametry členěny do několika skupin: parametry popisující fonaci, parametry popisující intenzitu řečového signálu, parametry popisující tempo řečového signálu, parametry popisující hybnost artikulačních orgánů a parametry popisující kvalitu řečových signálů. Vybrané jsou dále využity v rámci praktické části.

Čtvrtá kapitola charakterizuje nejen programovací jazyk Python, ale i vývojové prostředí PyCharm, které bylo využito v rámci této práce. Součástí je i způsob vytváření knihovny v programovacím jazyce Python.

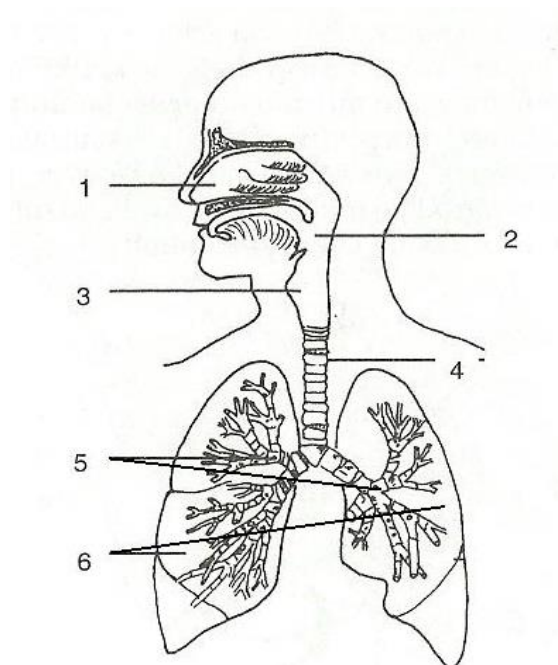
Pátá kapitola je praktická, zaměřuje se na návrh a realizaci systému. V rámci ní byl zpracován popis samotného systému a postupy při výpočtu vybraných parametrů řečového signálu. V závěru páté kapitoly je uveden příklad získaných dat.

1 ÚVOD DO TVORBY ŘEČI A POPIS ŘEČOVÉHO TRAKTU

Řeč je biologickou vlastností člověka, která mu umožňuje vědomě uplatňovat jazyk jako složitý systém znaků a symbolů ve všech jeho formách a pomocí něj sdělovat svá přání, pocity a myšlenky ([1]) a tím dosahovat komunikačního cíle. Řeč se dělí na řeč zevní neboli mluvenou a řeč vnitřní, která úzce souvisí s myšlením. Mluvená řeč je utvářena mluvními orgány. Dle B. Krahulcové ([2]) lze mluvidla v širším pojetí členit na trojí ústrojí rozdílné funkce. Jedná se o ústrojí respirační: sloužící k dýchání, ústrojí fonační: dochází zde k tvorbě hlasu a modulaci lidské řeči, a ústrojí artikulační: vznikají zde základní segmenty lidské řeči neboli hlásky. Kooperace těchto tří ústrojí nastává ve chvíli řečové produkce, v ostatních případech slouží odlišným účelům.

1.1 Respirační ústrojí

Respirační ústrojí, zobrazené na obrázku 1.1, se skládá z horních a dolních cest dýchacích. Horní cesty dýchací jsou tvořeny nosní dutinou, vedlejšími dutinami nosními a nosohltanem. Dolní cesty dýchací se skládají z hrtanu, průdušnice, průdušek a plic.



Obrázek 1.1: Dýchací ústrojí: 1 - dutina ústní, 2 - hltan, 3 - hrtan, 4 - průdušnice, 5 - průdušky, 6 - plice ([3])

Dýchání představuje vrozený reflex, jenž umožňuje výměnu plynů mezi externím prostředím a organismem. Dělí se na dýchání vnější a vnitřní. Při vnějším dýchání dochází k výměně plynů mezi vnějším prostředím a plicemi, při vnitřním dýchání dochází k předání kyslíku z krve do jednotlivých tkání. Proces dýchání probíhá ve dvou fázích: nádech a výdech, které se periodicky opakují. Fyziologický poměr doby nádechu a výdechu je přibližně 1:3 ([4]). Nádech je dějem aktivním, na kterém se podílí především bránice a mezižeberní svaly. Přirozený výdech je pak dějem pasivním, vydechovaný vzduch prochází hrtanem mezi volně rozevřenými hlasivkovými vazy a dále z těla pryč. Při tvorbě řeči však zajišťuje respirační ústrojí výdechový proud potřebný pro tvorbu hlasu, jež je hlavním médiem mluvené řeči. Mluvená řeč se skládá z jednotlivých hlásek, ty se vzájemně liší svou spotřebou výdechového proudu. Nejnižší spotřebu výdechu v českém jazyce má ze souhlásek hláska l, nízkou spotřebu mají hlásky t, ť, d, ď, p, b, k, g, j, n, ň, střední č, š, ž, c, s, z, r, ř, a největší hlásky h, f, ch. Ze samohlásek mají nejnižší spotřebu výdechového proudu hlásky i, u, střední e, o a nejvyšší a ([2]).

1.2 Fonační ústrojí

Základní hlas vzniká tak, že vdechnutý vzduch je rozdílným tlakovým spádem a stahujícími se plicními křídly skrz průdušnici a mezi sevřenými hlasivkami v hrtanu



Obrázek 1.2: Schéma činnosti hlasivek ([5])

opět vytlačován ([6]), výdechový proud pak opakovaně překonává napětí hlasivek a rozráží zavřenou hlasovou štěrbinu, tím dochází ke vzniku primárního hrtanového tónu. Zajímavostí je, že postavení hlasivek se mění v závislosti na artikulaci jednotlivých hlásek. B. Krahulcová ([2]) udává, že při realizaci souhlásek f, s, š, ch jsou hlasivky mírně rozevřeny, při realizaci úžinových souhlásek v, z, ž, j, l, r, ř a závěrových znělých

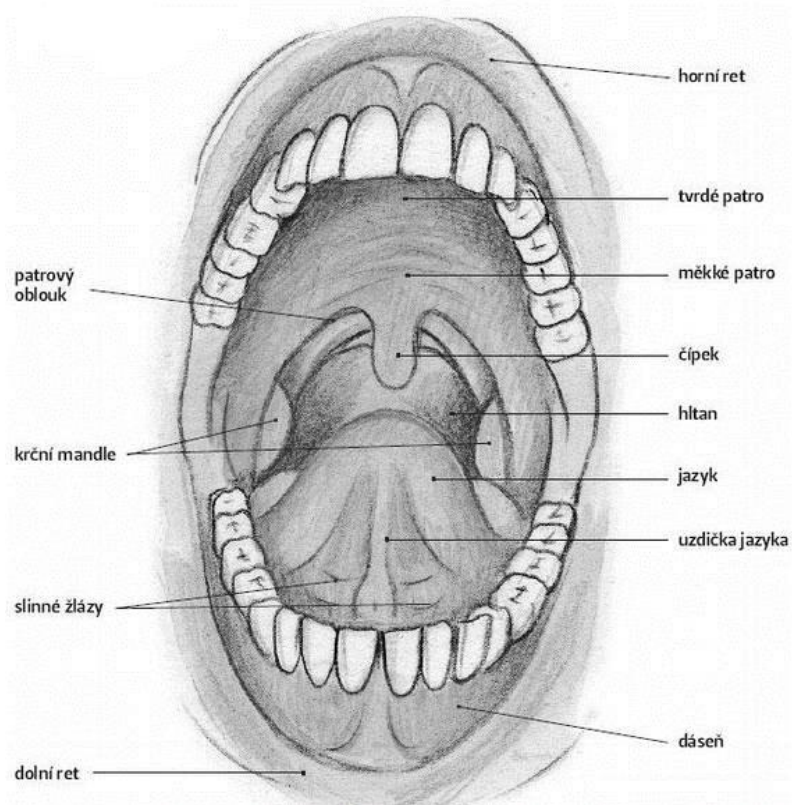
souhlásek b, d, d', g, m, n, ň jsou hlasivky volně přiložené, málo napjaté a pozvolně kmitají a vznikající hlas je hluboký a brumlavý, při realizaci závěrových neznělých souhlásek p, t, t', k, c, č jsou hlasivky přiloženy těsně a k tvorbě hlasu přitom nedochází, při realizaci souhlásky h jsou hlasivky málo napjaté a pomalu kmitají a vzniká hluboký a brumlavý hlas a při realizaci samohlásek a, e, i, o, u vzniká čistý tón, zatímco hlasivky jsou přiloženy k sobě a kmitají.

Výše zmiňovaný primární hrtanový tón vzniklý na hlasivkách má původně charakter tichého, nezvučného a chraptivého zvuku ([2]), který následně prochází rezonančními dutinami vokálního traktu, kde se dále upravuje. Zde dochází k posílení či potlačování určitých frekvencí vyšších harmonických tónů ([7]), které jsou součástí vzniklého hlasu. Mezi rezonanční prostory patří supraglotický prostor hrtanu, hltan, dutina ústní, nosohltan a dutina nosní ([8]). Výsledná rezonance je pak ovlivněna několika kritérii: objem a tvar rezonanční dutiny, rozměr vstupního a výstupního otvoru a hmotnost vzduchu v rezonančním prostoru ([7]). Vzniklý hlas tak dostává typicky lidské zabarvení. Právě barva hlasu je vedle výšky, síly a kvality hlasu, jedním ze znaků, kterým lze lidský hlas charakterizovat. Barva hlasu je dle I. Jedličky ([8]) ovlivněna délkou a hmotou hlasivek a tvarem rezonančních dutin, kvalita hlasu poměrem harmonických tónů a neharmonických zvuků v hlase a výška hlasu koresponduje s frekvencí kmitání hlasivek, čím je hlasivka objemnější a delší, tím je hlas hlubší a opačně. Síla neboli intenzita hlasu je dána rozsahem pohybu kmitajících částic vzduchu, objektivně se měří jako hladina akustického tlaku v určité vzdálenosti od úst – většinou 30 cm ([7]).

1.3 Artikulační ústrojí

Rezonanční prostory dutiny nosní a supraglotického prostoru hrtanu se tvarově nemění na rozdíl od rezonančních dutin hltanu a dutiny ústní, jejichž rozměry jsou ovlivňovány především pohybem jazyka a dolní čelisti ([4]), ale jak udává J. Klenková ([1]), i rty a měkkým patrem. Právě rty, jazyk, měkké patro, dolní čelist společně s hrtanem se řadí mezi pohyblivé artikulační orgány, ke kterým se dále řadí artikulační orgány nepohyblivé. Mezi nepohyblivé artikulační orgány lze řadit tvrdé patro, dásně, zuby a horní čelist. Souhrou nepohyblivých a pohyblivých artikulačních orgánů vznikají určitá artikulační postavení, díky nim dochází k modifikaci výdechového a fonačního proudu a vzniku jednotlivých samohlásek a souhlásek. Samohlásky čili vokály: a, e, i, o, u vznikají volným průchodem výdechového proudu rezonančními dutinami, aniž by musely překonávat nějakou překážku. Souhlásky neboli konsonanty musí vždy nějakou překážku překonávat ať již ve formě závěru, kterou musí výdechový proud rozrazit, nebo úžiny, kudy se výdechový proud tře. Přehledné dělení konsonant uvádí J. Klenková ([1]). Konsonanty lze dělit podle místa tvorby na hlásky bilabiální neboli obouretné: p, b, m, které jsou tvořeny oběma rty vytvářející překážku, labiodentální

neboli retozubné: f, v, tvořené spodním rtem a horními řezáky, alveolární neboli dásňové: t, d, n, c, s, z, l, r, ř, č, š, ž, jež jsou realizovány hrotem jazyka na horní dásni, palatální neboli tvrdopatrové: t', d', ň, j, uskutečňované hřbetem jazyka přibližujícího se k tvrdému patru, velární neboli měkkopatrové: k, g, ch, tvořené zadním hřbetem jazyka přitahujícího se k měkkému patru a laryngeální neboli hrtanové: h, jež je tvořené přímo v hrtanu.



Obrázek 1.3: Dutina ústní ([9])

Kromě dělení konsonant podle místa tvorby se často uplatňuje ještě dělení podle způsobu tvorby na souhlásky závěrové, úžinové a polozávěrové. Hláskami závěrovými neboli okluzivy se označují ty souhlásky, při jejichž tvorbě dochází k úplnému závěru, který je potřeba zrušit: p, b, m, t, d, n, t', d', ň, k, g, dle sluchového dojmu jsou označovány jako výbuchové neboli explozivny. Mezi hlásky úžinové neboli konstriktivny se řadí ty souhlásky, jež jsou tvořeny třením vzduchu úžinou, tzn. f, v, s, z, š, ž, j, h, ch, l, r, ř, dle sluchového dojmu se označují jako hlásky třené, frikativny. Hlásky polozávěrové neboli semiokluzivy mají dvě fáze tvorby, v první fázi dochází k závěru, v druhé k tření vzduchu úžinou, jedná se o hlásky c, č, dle sluchového dojmu jsou označovány jako hlásky polotřené neboli afrikaty. Avšak jak píše J. Klenková ([11]), řeč není pouze záležitostí mluvních orgánů, ale rovněž mozku a jeho hemisfér, které ji řídí. Verbální komunikace je tedy řízena nervovým systémem, ve kterém docházelo

v průběhu tisíciletí k formování nervových struktur a procesů umožňujících člověku dorozumívat se právě mluvenou řečí. Za hlavní centrum tvorby řeči lze považovat především mozkovou kůru, konkrétně se jedná o temporální lalok, parientální lalok a oblast Brocovy arey ve frontálním laloku ([10]), která je však v těsné spolupráci s vývojově staršími podkorovými oblastmi, jež se na řízení rovněž podílí. V případě poškození těchto struktur lze očekávat dopad na komunikační dovednosti jedince v podobě narušené komunikační schopnosti, kam je řazena i dysartrie.

2 DYSARTRIE

2.1 Charakteristika dysartrie a její klasifikace

Řečový projev člověka je závislý na složité souhře dýchacích svalů a svalů v oblasti hrtanu, měkkého patra, jazyka, rtů a čelisti. Aktivita těchto svalů, jak již bylo uvedeno výše, je realizována a kontrolována činností centrální nervové soustavy, především aktivitou motorických okřsků mozkové kůry a nervovými drahami, jmenovitě pyramidovou nervovou dráhou, extrapyramidovým systémem a mozkovými nervy ([11]). Následkem lézí centrálního nebo periferního nervového systému dochází ke vzniku tzv. neurogenních poruch řeči, jež lze blíže dělit na neurogenní motorické poruchy řeči a jazykové neurogenní poruchy ([12]). Za původce neurogenní motorické poruchy řeči lze pokládat poškození motorického systému kdekoli na dráze mezi mozkem a samotným svalem, kdy dochází k ochrnutí, oslabení nebo poruše koordinace svalstva podílejícího se na tvorbě řeči. V důsledku těchto změn mohou v různé míře nastat potíže v oblasti respirace, fonace, rezonance a prozódie, což jsou hlavní modalitky vyskytující se u poruchy motorické realizace řeči zvané dysartrie. Dysartrie se dělí na dysartrii vývojovou a získanou, které se dále člení do několika typů dle lokalizace léze či z hlediska neurologických syndromů ([13]). O vývojové dysartrii lze hovořit tehdy, vyskytují-li se uvedené potíže od počátku vývoje organismu na základě organického poškození centrální nervové soustavy dítěte, nejčastěji je spojována se skupinou dětí narozených s projevy dětské mozkové obrny ([14]). Získanou dysartrii se rozumí náhlé narušení procesu zrání centrální nervové soustavy vzniklé na bázi úrazu mozku, traumatu, infekce či onkologického onemocnění nervové soustavy, přičemž toto narušení vzniká u dospělých osob či dětí starších 1 – 2 let ([12]). Pojmem anartrie je pak označován nejtěžší stupeň poruchy motorických řečových modalit s praktickou ztrátou verbální komunikace s okolním světem ([14]).

Získanou dysartrii lze dělit do následujících typů: dysartrie flacidní, dysartrie spastická, dysartrie při unilaterální lézi centrálního neuronu, dysartrie ataktická, dysartrie hyperkinetická, dysartrie hypokinétická a dysartrie smíšená, která se projevuje kombinací projevů níže uvedených typů získané dysartrie.

Dysartrie flacidní či periferní (chabá) vzniká při lézi hlavových, ale také spinálních nervů ([15]). Vlivem poškození dolního motorického neuronu se stávají svaly účastníci se na mluvení a dýchání hypotonickými nebo chabými, a proto je zasažen každý typ pohybu, ať už volní a automatický, tak pohyb reflexivní ([10]). Dysartrie tohoto typu může vzniknout při chirurgických zákrocích, úrazech hlavy a krku, v důsledku virové infekce, tumoru či cévní mozkové příhody v oblasti mozkového kmene či při onemocnění myasthenia gravis. U flacidní dysartrie bývají přítomny znaky periferní parézy s atrofií postižených svalů a drobné svalové záškuby, tzv. fascikulace ([11]). Typickým příznakem je rovněž patologicky zvýšená nosovost neboli hypernazalita, snížený intraorální tlak, monotónní hlas, chrapot a poruchy polykání a

dýchání ([12]). Artikulace bývá narušena především při poškození lícního nervu (nervus facialis), trojklaného nervu (nervus trigeminus) a podjazykového nervu (nervus hypoglossus).

Dysartrie spastická či centrální vzniká při oboustranné poruše centrálního motoneuronu ([15]). Příčinami mohou být cévní mozkové příhody, traumata hlavy, tumor, infekce degenerativní onemocnění, zánětlivá nebo toxicko-metabolická onemocnění ([10]). Vlivem tohoto poškození dochází k poruchám vykonávání volních pohybů, k vystupňování reflexních pohybů neboli hyperreflexii a ke zvýšenému napětí svalstva neboli spasticitě. Spastická dysartrie se tak projevuje pomalým a pracným řečovým projevem, protahováním slov a nesrozumitelností delšího mluvního projevu, oslabeným dýcháním, hypernazalitou a potíží při polykání ([11]). Hlas pacientů se spastickou dysartií je drsný, tvořený s námahou, může být nepřírozeně hluboký nebo se mohou vyskytovat náhlé zlomy ve výšce hlasu ([10]).

Dysartrie při unilaterální lézi centrálního neuronu se projevují svalovou slabostí dolní částí obličeje, jazyka a rtů. Příčiny a projevy jsou podobné jako u spastické dysartrie, jen jsou tyto příznaky mírnější, neboť jsou kompenzovány nepostiženou stranou ([15]).

Dysartrie ataktická vzniká při lézích mozečku a jeho drah ([14]) a projevuje se špatně cílenými pohyby a špatnou koordinací společně s celkovým snížením svalového napětí neboli hypotonií ([12]). Příčinami bývají záněty a nádory mozečku či degenerativní onemocnění v této oblasti ([11]). Ataktická dysartrie se projevuje nekoordinovaným pohybem dýchacích svalů, ostrým hlasem nebo i hlasovým tremorem, omezeným rozsahem a intenzitou hlasu, zpomaleným tempem řeči, ulpíváním v jednotlivých artikulačních postaveních, důrazem na každé slabice a poruchami artikulace konsonant i vokálů ([15]).

Dysartrie hyperkinetická vzniká v důsledku poškození bazálních ganglií a souvisí se zvýšením míry pohybu ([10]). Nejčastějšími příčinami bývají degenerativní onemocnění nervového systému, cévní mozkové příhody lokalizované v oblasti bazálních ganglií a stavy při podávání léků, nejčastěji neuroleptik ([15], [11]). Hyperkinetická dysartrie bývá doprovázena sníženým svalovým tonem a mimovolní pohyby, které se v různé míře projevují i při mluvené řeči. Řeč je hlasitá, vykřikovaná, přerušovaná mimovolnými pohyby ([14]). Hlas bývá ostrý a tlačný.

Dysartrie hypokinetická vzniká při poškození bazálních ganglií nejčastěji u Parkinsonovy nemoci nebo jednorázovým úrazem hlavy nebo opakovanými údery do hlavy jako tomu je např. u boxerů ([15]).

2.2 Hypokinetická dysartrie

Hypokinetická dysartrie vzniká při lézích bazálních ganglií a jejich okruhů ([10]) a bývá typická pro Parkinsonovu chorobu neboli parkinsonismus. Parkinsonova nemoc je druhým nejčastějším neurodegenerativním onemocněním ([16]) postihující 100 až 150 pacientů na 100 000 obyvatel ([17]) První zmínky o tomto onemocnění lze nalézt v eseji „An Essay on the Shaking Palsy“ Jamese Parkinsona z roku 1817 ([16]) a ([18]). U této nemoci dochází k degenerativním změnám v substantia nigra, což způsobuje deficit neurotransmiteru dopaminu ([10]) a ([16]). Dle odborníků ([17]) se rozlišují dva typy Parkinsonovy nemoci, jedná se o vzácnou juvenilní formu začínající již v adolescenci, která vzniká v důsledku mutace genu Parkin, a častěji se vyskytující idiopatickou Parkinsonovu nemoc vznikající typicky mezi 60. a 70. rokem života člověka. Nemoc postihuje 1-2% osob starších 60 let ([19]). Z toho pak přibližně u 70% pacientů s parkinsonismem se vyskytují problémy s řečí ([18]).

Parkinsonismus se vyznačuje klasickou triádou příznaků, patří sem bradykineze (zpomalený pohyb), rigidita (ztuhlost) a klidový třes ([20]), které v různé míře následně omezují i komunikační schopnosti jedince s Parkinsonovou nemocí. Hypokinetická dysartrie se typicky odráží v oblastech fonace, faciokineze, artikulace, prozodie a plynulosti řeči ([16]). Prvními příznaky, které se u osob s Parkinsonovou nemocí manifestují, jsou odchylky v oblasti fonace, tedy tvorby hlasu, a respirace, která s fonací úzce souvisí ([20]). Poruchy hlasu byly zjištěny u 89% zkoumaných pacientů s hypokinetickou dysatrií ([10]). Příčinou bývá ztuhlé svalstvo hrtanu, kvůli němuž nedochází k dovírání hlasivek a výsledný hlas pak není odpovídající. Rovněž nedostatečná souhra svalstva hrtanu s respiračními svaly ztěžuje komunikaci. Proudění vydechovaného vzduchu z plic pacientů s parkinsonismem je charakteristické výrazným kolísáním, což se následně projevuje hlasovým třesem ([18]). U osob s parkinsonismem se objevuje zvýšená střední hodnota kmitočtu základního tónu F_0 , a to jak při realizaci delších větných celků, tak u prodloužených vokálů, s postupujícím onemocněním tato střední hodnota roste, a tudíž je možné využít tento parametr k posouzení pokročilosti onemocnění ([18]). U těchto osob lze také pozorovat neefektivní hospodaření s dechem, obě fáze dýchání, tedy nádech i výdech, se zkracuje, což má opět dopad na tvorbu hlasu, jeho kvalitu, sílu, i délku promluvy ([20]). Osoba s Parkinsonovou nemocí hovoří v jednoduchých větách, zrychluje tempo řeči, aby dokázala zamýšlenou větu realizovat na jeden výdech, nebo se nadechuje na nevhodných místech, např. uprostřed realizované věty, promluva tak působí rušivě. Při prudkém uvolnění výdechového proudu z plic může dojít k hypernazalitě, kdy velká část výdechu unikne přes nosní dutinu, což se odrazí ve zdůraznění některých formantových kmitočtů ([18]). V běžné komunikaci lze pozorovat monotónní hlas, je přítomná dysfonie - chraptivý, drsný hlas a hypofonie - snížená intenzita hlasu, kdy hlas je tichý a nevýrazný, problémem bývá se zesilováním či zeslabováním hlasu ([20]).

V oblasti faciokineze lze zaznamenat hypokinezi, bradykinezi, akinezi či rigiditu orofaciálního svalstva, která ovlivňuje správnou artikulaci, žvýkání, polykání a mimiku ([20]). Třes v orofaciální oblasti se vyskytuje spíše vzácně, přítomna však bývá hypomimie, která negativně ovlivňuje neverbální komunikaci.

Artikulace souvisí se správnou funkcí svalstva orofaciálního systému. Výslovnost vokálů nebývá obvykle narušena ([21]), poruchy se však objevují při výslovnosti konsonant, obzvláště pak exploziv ([18]), ačkoli změny lze zaznamenat i u ostatních souhlásek. Souhlásky výbuchové a polotřené, které jsou běžně tvořeny nekontinuálně z důvodu překonání překážky (viz Úvod do tvorby řeči a popis řečového traktu), jsou u osob s hypokinetickou dysartrií tvořeny kontinuálně, a hlásky třené, které bývají tvořeny s napětím, jsou tvořeny nenapjatě ([20]). Tyto změny jsou zapříčiněné redukováným rozsahem pohybu artikulačních orgánů ([15]), zejména pak narušenou elevací jazyka a nesprávným uzavřením dýchacích cest ([20]). Řeč pak zní nesrozumitelně a setřele.

Hypokinetická dysartrie bývá doprovázena poruchami prozodie a plynulosti řeči. Prozodickými faktory řeči je myšlen přízvuk, melodie řeči, tempo řeči a zařazování pauz ([1]). U osob s Parkinsonovou nemocí se vyskytují deficity většiny těchto faktorů. Monotónní řeč bývá způsobena absencí větné melodie nebo přízvuku ([20]). Řeč může být zrychlená, hovoří se o tzv. tachyfermii, která má vliv na umístování pauz v řeči, nebo zpomalená, jedná se o bradyfermii. Tempo řeči bývá narušeno ve smyslu palilalie či hezitací. Při palilalii se jedná o zrychlené opakování slova či jednotlivých slabik ([18]) a hezitace představují výskyt neúmyslných pauz v průběhu řečového projevu ([21]).

2.3 Diagnostické postupy

Při diagnostice dysartrie dochází k týmové spolupráci neurologů, klinických logopedů, foniatrů, oftalmologů a případně dalších odborníků. Základ tvoří především závěry z neurologického vyšetření ([13]), neboť charakter neurologického onemocnění výrazně ovlivňuje typ dysartrie, lokalizace léze v rámci nervové soustavy se odrazí v její symptomatologii ([15]). K neurologickým závěrům se dále přidávají informace o stavu smyslových orgánů, které zajišťuje otorinolaryngolog, příp. foniatr a oftalmolog ([13]). Logoped se poté snaží blíže zmapovat úroveň následujících modalit: respirace, fonace, faciokineze, diadochokineze, reflexní činnosti spjaté s polykáním, žvýkáním a kašláním, artikulace, srozumitelnost četby a mluvy, tempo a prozodie řeči ([12]). K těmto účelům byla vytvořena řada diagnostických metodik, v zahraničí např. „*Frenchay Dysarthria Assessment*“ nebo „*The Robertson Dysarthria Profile*“ ([15]). V České republice je pak nejpožávanějším „*Test 3F – dysarthrický profil*“, který je rozdělen na tři základní subtesty: F1 - faciokineze, F2 - fonorespirace a F3 - fonetika. V třetím vydání Test 3F má každý z těchto subtestů tři samostatné části, z nichž každá

obsahuje pět úkolů ([22]). Časová náročnost je 30 až 60 minut a splnění jednotlivých úkonů je hodnoceno pomocí jednoduché stupnice: 0 bodů je přiděleno v případě, kdy pacient úkol neprovedl vůbec nebo pouze v náznaku, 1 bod za potíže středně těžké až lehké a 2 body za správně provedený úkol ([22]). Maximální počet bodů, kterých může hodnocený jedinec dosáhnout, je 90. Na základě získaných bodů je pak možno, z důvodu toho, že test není standardizován, alespoň orientačně určit stupeň postižení: 85 – 90 bodů označuje výkon bez přítomnosti poruchy, tedy bez dysartrie, 74 – 85 bodů naznačuje velmi lehkou dysartrii, 57 – 73 bodů lehkou dysartrii, 36 – 56 bodů středně těžkou dysartrii, 17 – 35 bodů těžkou dysartrii a 0 – 17 bodů velmi těžkou dysartrii – anartrii ([22]). Na základě získaných údajů může logoped dále vyhodnotit, v kterých oblastech má pacient větší či menší obtíže a dle toho volit vhodné postupy při terapii.

Je nutné si uvědomit, že každý z výše uvedených testů však může být do jisté míry ovlivněn subjektivním pohledem vyšetřujícího, proto je potřebné, aby se tyto subjektivní hodnocení doplnila o výsledky objektivní, jež lze získat pomocí speciálních přístrojů a měření ([15]).

2.4 Terapeutické postupy

Základ léčby Parkinsonovy choroby tvoří léky zvané antiparkinsonika, za neúčinnější z těchto léků je v současné době považována levodopa (L-dopa) ([17]). Jedná se o aminokyselinu, která se přes stěnu tenkého střeva dostává do krevního řečiště a odtud je nesena do mozku, kde se přetváří na dopamin. Přibližně u 80% pacientů nastávají po podání L-dopy žádoucí změny k lepšímu, její funkce však není stálá, po určité době začínají účinky L-dopy slábnout. Za zmínku stojí rovněž agonisté dopaminu a inhibitory: iMAO a iCOMT. Agonisté dopaminu fungují v substantia nigra a imitují funkci dopaminu ([23]), v České republice jsou v současné době nejpoužívanější ropinirol a pramipexol ([17]). Inhibitory MAO zabraňují odbourávání dopaminu a tím zvyšují jeho hladinu a inhibitory COMT prodlužují efekt L-dopa a brzdí jednu z cest jejího odbourávání ([17]). Léčba antiparkinsoniky má dle některých studií příznivý vliv nejen na svalovou rigiditu končetin trupu, ale i na řeč, zlepšení bývá patrné v oblasti respirace, srozumitelnosti řečového projevu i ve změnách síly a výšky hlasu, jiné studie však tento vliv nepotvrdily ([20]). Kromě medikamentózní léčby existuje i léčba chirurgická v podobě hloubkové mozkové stimulace, při které dochází pomocí elektrických impulsů ke stimulaci a blokaci tkání v bledém jádru, které bývají při Parkinsonově chorobě nadměrně aktivní ([23]). Rovněž zde se odborníci neshodují na vlivu užití léčby na řeč pacienta, některé studie dokumentují pozitivní změny v řečovém projevu, jiné naopak uvádí negativní dopad na tyto funkce ([20]).

Terapie u osob s Parkinsonovou nemocí by však neměla být založena pouze na medikamentózní a chirurgické léčbě, ale na mezioborové spolupráci lékařů s nelékařskými zdravotními pracovníky, především fyzioterapeuty a klinickými

logopedy. Fyzioterapeut se zaměřuje na pohybový aparát daného pacienta, klinický logoped pak na řeč.

Základ logopedické péče u pacienta s Parkinsonovou chorobou v České republice spočívá především v dechových, hlasových a artikulačních cvičeních s doplněním o oromotorická cvičení ([12]). Cílem dechových cvičení je zlepšit koordinaci a sílu dýchacích svalů a nacvičit správný způsob dýchání. Správné techniky dýchání jsou pak propojovány s fonačními cviky, kdy se postupuje od prodloužené fonace samohlásek k jejich zesilování a zeslabování až po rytmické řady, pacient si osvojuje techniku měkkých hlasových začátků, učí se koordinovat dýchání s tvorbou hlasu, pracovat se silou a výškou hlasu ([20]). V rámci artikulačních cvičení je pozornost cílena na zřetelné vyslovování slabik, slov a vět. V USA pak v 80. letech 20. stol. vytvořily logopedky Lorraine Raming a Cynthia Fox behaviorální řečovou terapii LSVT (Lee Silverman Voice Treatment), která spočívá v intenzivních individuálních terapeutických sezeních probíhajících 4x týdně po dobu jednoho měsíce, náplní jednotlivých sezení jsou hlasová cvičení, při nichž se pacient s Parkinsonovou chorobou učí hledat a využívat správnou hlasovou intenzitu spolu se správným způsobem dýchání nejprve na izolovaných slovech, poté na frekventovaných frázích, při čtení a následně při konverzaci ([20]). LSVT má pozitivní dopad nejen na fonaci, ale i na prozodii, artikulaci, mimiku a polykání.

3 METODY ANALÝZY ŘEČOVÉHO SIGNÁLU

Akustická analýza řečových signálů představuje poměrně přesný a objektivní způsob získávání a vyhodnocování parametrů mluvené řeči ([24]). Vedle subjektivních metod hodnocení řeči přináší cenné informace o řečových charakteristikách hodnoceného jedince. Nejedná se však o jednoduchou úlohu, neboť lidská řeč je velmi proměnlivou záležitostí. Rozdíly lze nalézt nejen mezi jednotlivými osobami, které promluvu realizují, neboť fyziologické vlastnosti hlasového ústrojí ani způsob artikulace a prozodické faktory nejsou stejné ([25]), avšak i v rámci promluvy u konkrétního jedince. Výsledný řečový signál, který mluvčí vysílá, je ovlivněn výškou základního tónu, hlasitostí, přízvukem a tempem realizované promluvy, která se v čase může měnit, rovněž také okolním šumem a zkreslením při zpracování signálu například vlivem frekvenčních charakteristik mikrofonu, ekvalizéru, zesilovače či vlastností přenosových médií ([26]).

Dle J. Černockého ([27]) jsou metody popisu řečového signálu členěny na neparametrický popis a parametrický popis. Neparametrický popis je založen na poznacích o zpracování signálu (banky filtrů, Fourierova transformace apod.) a parametrický popis je založený na poznacích o tvorbě řeči, přičemž tato metoda využívá mnoho technik neparametrického popisu, není tedy snadné a někdy ani žádoucí je od sebe oddělovat. Získané hodnoty z obou uvedených metod jsou označovány vždy jako parametry.

3.1 Předzpracování řečového signálu

V současné době probíhá zpracování řečového signálu téměř výlučně v číslicové podobě, právě z těchto důvodů je nutné řečový signál před samotným zpracováním ještě upravit neboli předzpracovat, aby bylo následné zpracování jednodušší a efektivnější. Metody používané na předzpracování řečového signálu jsou:

- Ustřednění
- Preemfáze
- Segmentace

3.1.1 Ustřednění

Ustřednění je jednou ze základních metod používaných na předzpracování řečového signálu. Její podstata spočívá v odečítání stejnosměrné složky od získaného signálu. Stejnosměrná složka nenesé žádné užitečné informace, nýbrž při zpracování řečového signálu působí rušivě ([27]).

$$s'[n] = s[n] - \mu_s, \quad (3.1)$$

přičemž $s'[n]$ je výsledný signál po odečtení střední hodnoty, $s[n]$ představuje výchozí signál a μ_s musíme odhadnout.

Střední hodnotu je možné počítat dvěma způsoby: off-line a on-line.

Střední hodnota off-line se počítá průměrováním signálu:

$$s'[n] = \frac{1}{N} \sum_{n=0}^{N-1} s[n], \quad (3.2)$$

kde $s'[n]$ je opět výsledný signál po odečtení střední hodnoty, $s[n]$ je výchozí signál a N je počet vzorků původního filtrovaného signálu.

Střední hodnota on-line se využívá tehdy, pokud signál neustále přibývá nebo je příliš dlouhý, střední hodnota se v tomto případě odhaduje rekurzivně:

$$s'[n] = \gamma s'[n-1] + (1-\gamma)s[n], \quad (3.3)$$

kde γ se blíží 1.

3.1.2 Preemfáze

Preemfáze představuje zdůraznění vyšších frekvencí, v číslicovém zpracování signálu se preemfáze využívá k vyrovnání působení kmitočtové charakteristiky prostředí na řečový signál, která má vliv na vyšší kmitočtové složky řečového signálu přibližně od 2kHz a vyšších, tím dochází k vyrovnávání kmitočtové charakteristiky vstupního signálu. K těmto účelům se využívá filtrace preemfázovým číslicovým filtrem typu horní propust (HP) s konečnou impulzivní odezvou s přenosovou funkcí ([25]):

$$H_{\text{pre}}(z) = 1 - a_{\text{pre}}z^{-1}, \quad (3.4)$$

kde a_{pre} je koeficient, který bývá v rozsahu od 0,9 po 1.

3.1.3 Segmentace

Pomocí segmentace dochází k dělení řečového signálu na jednotlivé úseky neboli rámce. Důvodem tohoto kroku je fakt, že získaný řečový signál je náhodný a pro metody odhadu parametrů ne příliš vhodný. Účelem segmentace je tedy snaha dosáhnout takového signálu, který bude stacionární. Délka rámce by na jedné straně měla být dostatečně malá, aby se dal signál na daném úseku považovat za stacionární, ale na straně druhé dostatečně velká, aby bylo možné poměrně přesně získat požadované parametry. Kompromisem mezi těmito dvěma podmínkami je délka respektující setrvačnost hlasového ústrojí, typicky 20 – 25 ms ([27]). Segmentaci dělíme na: segmentace bez překrývání rámců nebo segmentace s překrýváním rámců ([25]).

- Segmentace bez překrývání rámců spočívá v rychlém časovém posunu v signálu za současných malých nároků na paměť a procesor, hodnoty parametrů se však mohou od jednoho rámce ke druhému významně lišit. Při segmentaci, kdy $p_{ram} = 0$ lze počet rámců vypočítat pomocí vzorce:

$$N_{ram} = \left\lfloor \frac{N}{I_{ram}} \right\rfloor, \quad (3.5)$$

přičemž N_{ram} je počet rámců, N je délka segmentového signálu a I_{ram} je velikost rámce ve vzorcích.

- Segmentace s překrýváním rámců zajišťuje „vyhlazené“ průběhy parametrů, má však naopak pomalý časový průběh s vysokými nároky na paměť a procesor. Výsledné parametry mohou být rámec od rámce příliš podobné, což není žádoucí z důvodu detekce daných parametrů a jejich rozpoznávání ([27]). Počet rámců v případě, kdy $p_{ram} \neq 0$, lze vypočítat:

$$N_{ram} = 1 + \left\lfloor \frac{N - I_{ram}}{S_{ram}} \right\rfloor, \quad (3.6)$$

přičemž N_{ram} je opět počet rámců, N je délka segmentového signálu, I_{ram} je velikost rámce ve vzorcích a S_{ram} je velikost nepřekryté části rámců ve vzorcích. Kompromisem je běžná délka překrytí 10 ms, tedy 50% rámce ([25]).

Tvorba rámců, tedy segmentace, se realizuje tzv. okénkovou funkcí. Mezi nejčastěji používané patří pravoúhlé okno a Hammingovo okno.

Pravoúhlé okno je definováno takto:

$$w[n] = 1, \text{ pro } n = 0, 1, \dots, N - 1, \quad (3.7)$$

kde $w[n]$ je hodnota vzorku signálu po násobení oknem, pro ostatní hodnoty n je $w[n] = 0$.

Hammingovo okno je definováno takto:

$$w[n] = 0,54 - 0,46 \cos\left[n \frac{2\pi}{N}\right], \text{ pro } n = 0, 1, \dots, N - 1, \quad (3.8)$$

kde znovu platí, že pro ostatní hodnoty n je $w[n] = 0$.

Hammingovo okno bývá používáno v průměru častěji než okno pravoúhlé. Hammingovo okno má širší hlavní lalok, což s sebou přináší horší frekvenční rozlišení a širší přechodové pásmo jak okno pravoúhlé, avšak oproti němu má Hammingovo okno větší potlačení postranních laloků, čímž má lepší vlastnosti v nepropustným pásmě a do spektra signálu daného okna se nedostanou téměř žádné spektrální složky z vedlejších oken potažmo rámců ([26]).

3.2 Parametrizace

Z řečového signálu je možné získat určité hodnoty neboli parametry signálu. Parametrem může být skalár, matice čísel či vektor charakterizující určitou vlastnost řeči nebo hlasového traktu. Pomocí získaných parametrů je možné popsat plynulost řeči, tzn. její zpomalování či zrychlování, zrychlené opakování slov či slabik, neúmyslné pauzy. Rovněž lze obdržet údaje o artikulaci, tedy informace o práci s jazykem či měkkým patrem, jehož omezená funkce způsobuje hypernazalitu. Dále lze získat parametry popisující kvalitu řeči, práci s dechem, suprasegmentální rysy, v nichž lze spatřit odraz emocí řečníka, a taktéž správné fungování hlasivek, frekvenci jejich kmitu a nepřímo i tvar hlasového traktu, tedy délku a objem dutiny nosní, ústní a hrdelní ([15]).

Existuje celá řada parametrů charakterizujících řečový signál, ty jsou založeny buď na lineárních metodách a transformacích, nebo se může jednat o nelineární dynamické parametry. Parametry, které jsou nejčastěji užívány pro popis řečového signálu u osob s Parkinsonovou chorobou, lze třídit následovně:

- parametry popisující fonaci
- parametry popisující intenzitu řečového signálu
- parametry popisující tempo řečového signálu
- parametry popisující hybnost artikulačních orgánů
- parametry popisující kvalitu řečových signálů

Uvedené třídění je pouze orientační, neboť některé parametry lze zařazovat pod více kategorií.

3.2.1 Parametry popisující fonaci

Kmitočet základního tónu F_0

Kmitočet základního tónu F_0 patří vedle intenzity a tempa řeči mezi tři základní suprasegmentální rysy ([18]), přičemž kmitočet základního tónu F_0 popisuje práci hlasivek a udává základní tón, od kterého jsou odvozovány vyšší harmonické tóny ([28]), pro posluchače se projevuje ve výšce hlasu mluvčího. F_0 se liší u dětí i dospělých, ale vliv má i pohlaví daného jedince. Pomocí tohoto parametru lze velmi dobře detekovat zdravý a poškozený hlas. U osob s Parkinsonovou nemocí dochází postupem času k zvyšování střední hodnoty kmitočtu základního tónu F_0 , a to jak v případě delších mluvních celků, tak u prodloužených vokálů, tímto parametrem je tedy možné usuzovat na pokročilost onemocnění ([18]). Jeden ze vztahů sloužící k výpočtu základního hlasivkového tónu F_0 je možné vyjádřit vzorcem:

$$F_0 = \frac{f_{vz}}{k_{min}}, \quad (3.9)$$

kde f_{vz} představuje vzorkovací kmitočet a k_{min} je pozice (ve vzorcích) prvního maxima autokorelační funkce.

Jitter

Pro popis variability základního tónu F_0 se nejčastěji využívá směrodatné odchylky a jitter ([18]). Jitter vyjadřuje variabilitu frekvence základního tónu F_0 v čase ([29]). Tento parametr popisuje chvění neboli třes hlasivek ([25]). V případě, kdy jsou hlasivkové svaly ztuhlé nebo je narušena schopnost udržení hrtanových svalů ve stabilní pozici po delší dobu, dochází k dopadu na kolísání F_0 , naopak u monotónní řeči či dysprozodie F_0 nekolísá ([15]). Jitter lze vypočítat rozdílem dvou sousedních period hlasivkových pulzů dělený průměrnou délkou periody ([25]):

$$J_g = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_g[i] - T_g[i-1]|}{\frac{1}{N} \sum_{i=0}^{N-1} T_g[i]}, \quad (3.10)$$

kdy je uvažováno, že N značí počet period hlasivkových pulzů (segmentů), pro každý segment existuje základní perioda a $T_g[i]$ je i -tá hodnota základní periody.

Shimmer

Shimmer je parametr poruchy amplitudové stability ([28]). Jedná se tedy o amplitudovou nepravidelnost, kdy dochází ke kolísání amplitudy hlasivkových impulzů. Fyziologická hodnota pro shimmer by měla být menší než 5% ([30]). Stejně jako u jitteru se může amplituda hlasivkových pulzů lišit i v rozsahu velmi krátkého úseku řečového signálu. Lze jej vypočítat jako rozdíl velikosti dvou sousedních pulzů dělený jejich průměrnou velikostí ([25]):

$$S_g = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_g[i] - A_g[i-1]|}{\frac{1}{N} \sum_{i=0}^{N-1} A_g[i]}, \quad (3.11)$$

kde N je počet amplitud hlasivkových pulzů a $A_g[i]$ je i -tá hodnota amplitudy.

Pitch Period Entropy

Dalším parametrem vztahujícím se k fonaci je Pitch Period Entropy (PPE). Jedná se o parametr, jehož výhoda spočívá v nezávislosti na pohlaví řečníka, při jeho výpočtu se bere v úvahu logaritmická půltónová škála, inverzní filtrace a odhad entropie ([29]).

3.2.2 Parametry popisující intenzitu řečového signálu

Dalším suprasegmentálním rysem je intenzita řeči. U osob s Parkinsonovou nemocí bývá intenzita řeči oproti zdravé populaci nižší a příliš nekolísá, mluvčí s Parkinsonovu nemocí je schopný hlas krátkodobě zesílit, ale záhy tato intenzita opět klesá ([18]). Mezi parametry popisující intenzitu řečového signálu lze řadit např. krátkodobou energii (STE – Short-Time Energy), Teagerův-Kaiserův energický operátor (TKEO) nebo Low Short-Time Energy Ratio (LSTER).

Krátkodobá energie

K popisu hypofonie či odchylek intenzity hlasu lze využít krátkodobé energie ([29]). Pomocí ní je možné vystihnout energii řečového signálu v jednotlivých segmentech, přičemž slouží např. k rozlišování znělých hlásek vyznačujících se vyšší energií a neznělých hlásek s energií nižší. Rovněž lze využít k rozpoznávání přechodu mezi znělým a neznělým úsekem řečového signálu. K výpočtu slouží následující vzorec:

$$STE = \sum_{k=0}^{N-1} s[k]^2, \quad (3.12)$$

přičemž N je počet vzorků segmentu a $s[k]$ je vzorek diskrétního řečového signálu.

Teager-Kaiser energetický operátor

Teager-Kaiser energetický operátor se taktéž zaměřuje na popis intenzity řečového signálu, vypočítává se z celého řečového signálu nejen ze segmentů jako je tomu u krátkodobé energie. Rozdílem oproti krátkodobé energii je taktéž fakt, že TKEO bere v úvahu i frekvenci signálu ([29]). TKEO je definovaný vzorcem:

$$\psi[s(n)] = s(n)^2 - s(n+1)s(n-1), \quad (3.13)$$

kde $s(n)$ označuje vstupní řečový signál.

Low Short-Time Energy Ratio

K diferenciaci mezi řečovým a hudebním signálem slouží Low Short-Time Energy Ratio (LSTER). Jedná se o poměr počtu segmentů, v nichž je krátkodobá energie (STE) nižší než 0,5násobek průměrné krátkodobé energie v sekundovém úseku segmentového signálu ([31]). Tento parametr je založen na předpokladu, že se v mluvené řeči vyskytuje větší množství segmentů obsahujících tiché části (pauzy), než je tomu u hudebních signálů. Řeč tedy vykazuje větší odchylky než hudební signál ([29]). LSTER lze vypočítat takto:

$$LSTER = \frac{1}{2N} \sum_{n=0}^{N-1} [\text{sgn}(0,5\overline{STE} - STE(n)) + 1], \quad (3.14)$$

kde STE je krátkodobá energie segmentovaného signálu a N je počet segmentů signálu. K posouzení hlasu u osob s Parkinsonovou nemocí se využívá z důvodu vykazování vyšších hodnot u patologické produkce hlasu, především při prodloužené realizaci vokálů, kdy dochází ke kolísání výdechového proudu a tím k nestabilitě v udržení hrtanových svalů ([29]).

3.2.3 Parametry popisující tempo řečového signálu

Třetím základním suprasegmentálním rysem je tempo řeči. Tempo řeči je možné vyjádřit pomocí různých parametrů. Mezi tyto parametry řadíme Total Speech Time (TST), Total Pause Time (TPT) a Not Speech Time (NST). TST je trvání řeči, která obsahuje i tiché, např. mezislovní pauzy, TPT představuje celkovou délku pauz delších než 10 ms a NST pak odpovídá rozdílu TST a TPT, tedy ([32]):

$$NST = TST - TPT. \quad (3.15)$$

Z těchto parametrů je dále možné definovat Total Speech Rate (absolutní tempo TSR), což je poměr celkového počtu slabik vyskytujících se v řeči a TST (vyjádřeno ve slabikách za sekundu), Net Speech Rate (tempo čistě řečového signálu NSR), kde se jedná o poměr celkového počtu slabik vyskytujících se v řeči a NST (vyjádřeno ve slabikách za sekundu) a Articulation Rate (artikulační rychlost AR), což je podobné jako NSR, avšak v tomto případě jsou pauzy delší než 50 ms ([18]). Současné studie se příliš neshodují na tom, zda je tempo u pacientů s Parkinsonovou nemocí zvýšené, snížené nebo se neliší od kontrolního vzorku osob.

3.2.4 Parametry popisující hybnost artikulačních orgánů

Během průchodu výdechového proudu násadními trubicemi hlasového traktu dochází v dutině hrtanu, úst a nosu k rezonancím. Frekvence, na nichž dochází k těmto rezonancím, se označují jako formanty (F_1, F_2, F_3, \dots). Hodnoty formantů a odpovídající prostor v dutinách ovlivňuje poloha artikulačních orgánů. Na základě toho je možné pomocí formantů zaznamenávat hybnost jazyka a do jisté míry i artikulaci ([15]). Formanty F_1 a F_2 nejvíce vypovídají o produkci vokálů. U osob s Parkinsonovou nemocí jsou proto často využívány k popisu řeči. Negativem formantových kmitočtů je jejich

závislost na věku a pohlaví, z toho důvodu se přistoupilo k využívání poměru F_{2i}/F_{2u} , u kterého byla dokázána schopnost robustně odlišit zdravou řeč od řeči postižené hypokinetickou dysartrií, navíc bez závislosti na pohlaví ([18]).

3.2.5 Parametry hodnotící kvalitu řečového signálu

Kvalita hlasu bývá často snížena vlivem dysfonie, kdy je řeč chraptivá, drsná a nesrozumitelná, nebo vlivem nesprávné funkce měkkého patra. Nesprávnou funkcí měkkého patra dochází k narušení poměrů mezi oralitou a nazalitou a výsledná řeč bývá doprovázena hypernazalitou. Míru kvality řeči lze tedy popsat úrovní přítomného šumu. Existuje několik řečových parametrů, které se touto problematikou zabývají.

Zero Crossing Rate

Jedním z nejjednodušších parametrů je Zero Crossing Rate (ZRC), jež popisuje počet průchodů signálu nulovou úrovní ([15]). Průchod nulovou úrovní nastane ve chvíli, kdy dva sousední vzorky zkoumaného signálu mají různé znaménko. Z metod ZCR vychází postupy pro určení základního hlasivkového tónu nebo přibližné stanovení nejsilnějšího formantu ([26]). Tento parametr je však výrazně citlivý na šum a na posun stejnosměrné složky. Vztah pro výpočet počtu průchodu nulou v jednom segmentu řečového signálu je stanoven takto ([27]):

$$Z = \frac{1}{2} \sum_{n=1}^{l_{ram}-1} |\text{sign } x[n] - \text{sign } x[n-1]|, \quad (3.16)$$

kde $\text{sign}(x)$ je zjednodušená znaménková funkce:

$$\text{sign } x[n] = \begin{cases} +1 & \text{pro } x[n] \geq 0 \\ -1 & \text{pro } x[n] < 0 \end{cases} \quad (3.17)$$

Kdy funkce $|\text{sign } x[n] - \text{sign } x[n-1]|$ dává hodnotu 2 pro každý případ, že se od vzorku $x[n-1]$ ke vzorku $x[n]$ změnilo znaménko.

Harmonic-to-Noise ratio

Dalším z významných parametrů je Harmonic-to-Noise ratio (HNR). Jedná se o parametr podílu harmonického signálu ku šumu ([15]). Lze ho vypočítat pomocí následujícího vzorce ([26]):

$$HNR = 10 \log_{10} \left(\frac{\sum_{i=1}^N |S_i|^2}{\sum_{i=1}^N |N_i|^2} \right), \quad (3.18)$$

kde $|S_i|$ je odhad energie harmonického signálu a $|N_i|$ je odhad energie šumu.

Noise-to-Harmonic ratio

Noise-to-Harmonic ratio (NHR) lze poté spočítat jako převrácenou hodnotu HNR, tedy podle vztahu:

$$NHR = \frac{1}{HNR}. \quad (3.19)$$

4 PROGRAMOVACÍ JAZYK PYTHON

4.1 Python

Jedná se o objektově orientovaný programovací jazyk, který je celkem snadno naučitelný. Kód psaný v Pythonu je přehledný a v porovnání s ostatními programovacími jazyky také krátký. Bývá rovněž řazen mezi skriptovací jazyky. Python je dynamicky interpretovaný, což znamená, že se kód překládá za běhu programu postupně řádek po řádku. Případné chyby se tedy projeví až po spuštění. Je vyvíjen již od roku 1990 jako open-source. Zdrojový kód bývá obvykle uložen v souborech s příponou „py“ nebo „pyw“, ale některé soubory mohou mít i jiné přípony. Lze s ním pracovat na počítačích s operačním systémem Windows, Linux a rovněž v rámci Mac OS ([34]).

Na některých zařízeních s operačním systémem Linux je již Python k dispozici, je však zapotřebí se ujistit, zda se jedná o aktuální verzi, neboť některé operace by v starších verzích nemusely fungovat požadovaným způsobem, avšak byla snaha tento problém řešit přidáním nových funkcí i do starších verzí. Podobné je to i v případě operačního systému Mac OS X. Naproti tomu na počítačích s operačním systémem Windows od společnosti Microsoft se musí tento programovací jazyk nejprve stáhnout a nainstalovat. Pozornost je však nutné zaměřit i na fakt, zda dané zařízení, na které bude Python instalovaný, používá 32bitovou nebo 64bitovou architekturu ([35]). Samotná instalace ve většině případů nepředstavuje již výraznější problém. V roce 2008 byla vytvořena verze 3.0., ze které v současné době vychází nejnovější verze Python 3.6.5. Python není zpětně kompatibilní, proto byly vyvinuty nástroje pro konverzi zdrojových kódů psaných ve starší verzi do formy pro verzi 3.0 ([37]).

Jako vývojové prostředí můžeme používat výchozí pythonovský shell, jež je dobrým nástrojem pro ladění kratších programů. Pro složitější zdrojové kódy je vhodnější použít jiné IDE, které podporuje syntaxi Pythonu. Nejznámější IDE je pravděpodobně Visual Studio od Microsoftu, do kterého je třeba doinstalovat doplněk Python tools, jedná se však o editor, který podléhá zpoplatnění. Variantou může být i hojně využívané prostředí NetBeans, jehož výhoda spočívá v tom, že je dostupný zdarma. Mezi další, často využívané, spadá vývojové prostředí PyCharm, což je editor speciálně vytvořený pro programování v Pythonu ([35]). V rámci této práce došlo k využití právě tohoto vývojového prostředí.

4.1.1 Dekorátor

Dekorátor (wrapper) je funkce, která mění nebo rozšiřuje funkčnost jiné funkce či třídy ([38]). Jedná se o podobnou funkci, jakou umožňuje Aspect-Oriented Programming v jazyce Java s tím rozdílem, že dekorátory jsou mnohem jednodušší.

Python již obsahuje několik předpřipravených dekorátorů, z nichž nejpoužívanější jsou `@classmethod` a `@staticmethod` ([39]). Pro příklad lze uvést dekorátor, který se po zavolání funkce postará o výpis jejího jména.

```
def decorator(target):  
  
    def wrapper():  
        print 'Calling function "%s"' % target.__name__  
        return target()  
  
    @decorator  
    def target():  
        print 'I am the target function'
```

Obrázek 4.1: Dekorátor ([40])

Výpis pak vypadá takto:

```
>>> target()  
Calling function "target"  
I am the target function
```

Obrázek 4.2: Výpis ([40])

4.2 PyCharm

Jak již bylo zmíněno výše, PyCharm byl vytvořen hlavně pro programování v Pythonu. Lze v něm ale používat i jiné programovací jazyky, např. HTML, JavaScript, CSS nebo XML. Podporu pro ostatní jazyky, které nejsou kompatibilní s vývojovým prostředím PyCharm, je možné přidat pomocí různých pluginů.

Toto vývojové prostředí se dá používat na počítačích s operačním systémem Windows, Linux i Mac OS X, což je jeden z důvodů, proč je tolik používán. Další výhodou představuje možnost přidání webového aplikačního frameworku Django, jež se užívá hlavně k vytváření webových aplikací. PyCharm podporuje poslední verze Pythonu 3, stejně tak v něm ale lze spustit soubory vytvořené ve starším Pythonu 2. Jako většina podobných IDE, podporuje automatické doplňování. Obsahuje také integrovaný debugger a unit test ([41]).

4.3 Tvorba knihovny

Při vytváření delšího programu v programovacím jazyce Python je výhodnější vybrané definice napsat do speciálního souboru, které je možné při běhu programu kdykoliv zavolat. Těmto souborům se říká moduly ([42]). Jedná se o textové soubory s příponou „py“, kdy jméno modulu odpovídá jménu daného souboru ([43]).

Moduly jsou analogií knihoven v jazyce C. Jejich velkou výhodou není pouze možnost uložení funkcí a spustitelných příkazů, ale také celkové zpřehlednění kódu. Dalším pozitivem je možnost vytvořený modul zavolat i v jiném programu, než pro který byl původně vytvořen, nebo modulu ([42]). Každý modul má svou vlastní tabulku symbolů, která je používána jako globální tabulka všemi funkcemi definovanými v modulu.

Načtení modulu v programu se provádí příkazem:

```
import jmeno_modulu.
```

Jméno modulu se zadává bez přípony „py“.

Načtení funkce z modulu se provede po zavolání příkazu:

```
jmeno_modulu.jmeno_funkce()
```

nebo také použitím:

```
from jmeno_modulu import jmeno_funkce.
```

Obvykle jsou všechna prohlášení o importu umístěna na začátku modulu, ale uvedené pravidlo není nutností. Samozřejmostí je i možnost načtení více funkcí z modulů zároveň. Toho je možné docílit tím způsobem, že se názvy funkcí napíší za klíčové slovo import a oddělí se čárkami. Vytvoření nového modulu lze provést uděláním textového souboru s příponou „py“. Po zkompilování modulu se však změní přípona souboru na „pyc“, tudíž při druhém načítání modulu se již použije soubor s touto příponou, což urychlí celou operaci ([44]). Python ukládá do adresáře pycache ve vyrovnávací paměti (cache) kompilovanou verzi každého modulu. Při následujícím spuštění Python zkontroluje datum modifikace zdroje proti kompilované verzi, aby zjistil, zda není zastaralý a není potřeba ho znovu zkompilovat ([42]).

Moduly se obvykle umísťují do balíčků (packages). Jedná se o adresáře se soubory Pythonu. Importování balíčků a modulů, které jsou obsaženy v balíčcích, se provádí podobně jako načítání jednotlivých modulů, tedy příkazem import. Adresář musí obsahovat soubor `_init_.py`. Tento soubor naznačuje, že adresář, ve kterém je obsažen, je balíček Pythonu a lze jej tedy importovat jako modul. Samotný soubor může i nemusí být prázdný. Importování modulu z balíčku se provádí příkazem ([42]):

```
import jmeno_balicku.jmeno_modulu
```

nebo také:

```
from jmeno_balicku import jmeno_modulu
```

Moduly uvnitř balíčků se mohou na jiné moduly ve stejném balíčku odkazovat relativní cestou. Tečkou je označen aktuální balíček, dvěma tečkami je označen balíček nadřazený. Volání modulu v rámci jednoho balíčku lze tedy provést příkazem:

```
from . import jmeno_modulu,
```

pro volání modulu v nadřazeném balíčku slouží ([44]):

```
from .. import jmeno_modulu
```

5 NÁVRH A REALIZACE SYSTÉMU

5.1 Popis systému

Navrhnutý systém pro analýzu řečových signálů, který má za úkol diagnostikovat Parkinsonovu chorobu, se skládá z několika bloků:

- **Načtení vstupních dat** – první blok je zaměřen na načtení vstupních dat. Vstupními daty se v tomto případě rozumí audio nahrávky mluveného projevu lidí s Parkinsonovou nemocí. Pro srovnání parametrů řečového signálu lze samozřejmě vložit i nahrávku mluveného projevu zdravého jedince a tím posoudit přítomnost či nepřítomnost symptomů doprovázejících hypokinetickou dysartrii. Nahrávky musí mít příponu „wav“.
- **Předzpracování** – druhým blok je určen k předzpracování vstupního signálu. Jak již bylo uvedené výše, existují tři různé druhy předzpracování. Konkrétně se jedná o preemfázi, ustřednění a často používanou segmentaci. V této práci byla využita pro předzpracování signálu právě segmentace. Segmentací se rozumí rozdělení nahrávky na menší úseky, tzv. segmenty, aby bylo možné následně změřit parametry, které mohou poukázat na přítomnost symptomů hypokinetické dysartrie, potažmo Parkinsonovy nemoci. Nahrávka se dělí na rámce velké 20 ms, důvodem je snaha zajistit, aby se dal průběh signálu jednoho segmentu považovat za stacionární.
- **Parametrizace** – v této části systému již dochází k výpočtům samotných parametrů řečového signálu. Na základě výše uvedeného orientačního členění byly pro účely této práce vybrány následující parametry. Z parametrů popisujících fonaci byly zvoleny kmitočet základního tónu F_0 , jitter a shimmer. Mezi parametry hodnotící intenzitu řečového signálu byla vybrána krátkodobá energie. Další kategorie je zaměřena na tempo řeči s parametrem NST. Ze skupiny vztahující se k posouzení hybnosti artikulačních orgánů byl vybrán parametr, který počítá formantové frekvence. Ve skupině orientované na posouzení kvality řečového signálu byly zvoleny parametry ZCR, HNR a NHR.

5.2 Databáze řečových záznamů

Pro testování vytvořeného systému bylo použito několik nahrávek zdravých osob i osob s Parkinsonovou chorobou. Vzhledem k tomu, že některé parametry mohou být ovlivněny věkem i pohlavím testovaných jedinců, byly vybrány nahrávky mužů i žen různých věkových kategorií. Řečníci měli za úkol přednést krátkou báseň o čtyřech verších, kterou mohli číst z papíru. Vzorkovací frekvence nahrávek $f_{vz} = 16$ kHz.

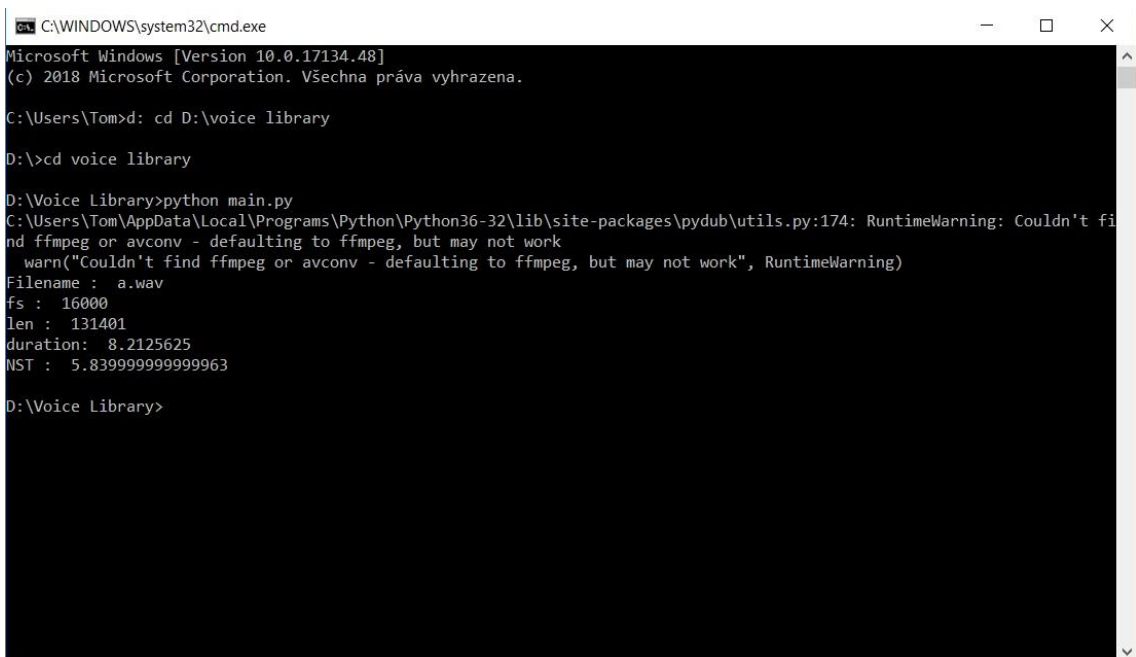
5.3 Použité parametry

Realizace systému pro rozpoznání Parkinsonovy choroby z mluveného projevu byla prováděna programovacím jazykem Python ve vývojovém prostředí PyCharm. Pro účely bakalářské práce byla vytvořena knihovna, která obsahuje několik parametrů, jež patří k základním pro posuzování přítomnosti symptomů hypokinetické dysartrie u Parkinsonovy nemoci. K výpočtu některých parametrů je využíván program Praat.exe. Před použitím je potřeba importovat některé knihovny pro práci se zvukovými signály, např. Numpy nebo Scipy. Program se spouští přes terminál zavoláním hlavní funkce main.py. Parametry, které jsou implementovány v knihovně:

- Výpočet základního tónu F_0 – pro výpočet základního hlasivkového tónu F_0 se musí zavolat skript `F0.py`. Tento skript provede vypsání přímo do konzole. Dále je přiložena i verze skriptu, která používá pro výpočet program Praat.exe.
- Jitter – k výpočtu jitteru se musí zavolat modul `parametr_jitter.py`, který vypíše přímo do konzole hodnotu jitteru.
- Shimmer – pro získání hodnoty shimmeru je třeba zavolat modul `shimmer.py`, který stejně jako jitter vypíše hodnoty do konzole.
- Krátkodobá energie – k výpočtu krátkodobé energie slouží skript `short_time_energy_and_zero_cross_rate.py`. U tohoto skriptu je třeba dbát na správný výběr parametru, jelikož obsahuje nejenom funkci pro výpočet krátkodobé energie, ale i funkci pro výpočet parametru Zero Crossing Rate.
- NST – k získání hodnoty NST slouží modul `NST.py`. V tomto modulu již jen stačí zavolat funkci `calculate_NST` a dojde k výpočtu Net Speech Time.
- Formantové frekvence – k výpočtu formantových frekvencí slouží `formants.py`. Zde je potřeba zavolat funkci `calculate_formants`.
- ZCR – jak již bylo zmíněno výše, tento parametr je uložen v modulu `short_time_energy_and_zero_cross_rate.py`, kde si lze vybrat mezi výpočtem krátkodobé energie a výpočtem parametru Zero crossing Rate.
- HNR – k dosažení výsledku Harmonic-to-Noise ratio je potřeba zavolat skript `HNR_and_NHR.py`. Zde se nachází funkce pro výpočet HNR a funkce pro výpočet NHR. Tudíž je třeba si vybrat, který z parametrů chce uživatel zkoumat. K výpočtu HNR se použije funkce `calculateHNR`.
- NHR – parametr Noise-to-Harmonic ratio je uložen stejně jako HNR ve skriptu `HNR_and_NHR.py`. K získání výsledku NHR se zavolá funkce `calculateNHR`.

5.4 Vyhodnocení

Při zpracování vzorových audionahrávek dochází k jejich členění na segmenty velké 20 ms, u každého tohoto segmentu dojde k jeho vyhodnocení v rámci daného parametru. V rámci jedné sekundy, která se skládá z 50 segmentů, získáme tedy 50 číselných údajů. V případě NST je situace rozdílná, neboť získaný parametr je počítán z celého úseku nahrávky, tudíž lze získat pouze jeden číselný údaj. Jelikož cílem této práce není zhodnocení ani porovnávání vzorových audionahrávek u osob s Parkinsonovou nemocí a bez ní, není možné získané výsledky v rozsahu této práce vyhodnocovat. Pro ukázkou lze uvést např. parametr NST.



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.17134.48]
(c) 2018 Microsoft Corporation. Všechna práva vyhrazena.

C:\Users\Tom>d: cd D:\voice library

D:\>cd voice library

D:\Voice Library>python main.py
C:\Users\Tom\AppData\Local\Programs\Python\Python36-32\lib\site-packages\pydub\utils.py:174: RuntimeWarning: Couldn't find ffmpeg or avconv - defaulting to ffmpeg, but may not work
  warn("Couldn't find ffmpeg or avconv - defaulting to ffmpeg, but may not work", RuntimeWarning)
Filename : a.wav
fs : 16000
len : 131401
duration: 8.2125625
NST : 5.839999999999963

D:\Voice Library>
```

Obrázek 5.1 Výpis NST

ZÁVĚR

Hlavním cílem bakalářské práce bylo vytvoření knihovny pro parametrizaci dysartrické řeči v jazyce Python. Dílčí cíle spočívaly v seznámení se s metodami analýzy řečového signálu u Parkinsonovy nemoci, v porozumění moderním parametrizačním technikám, které mají za úkol kvantifikovat poškození motorických aspektů řeči, a rovněž v implementaci vybraných parametrů v jazyce Python.

Bakalářská práce se skládá z pěti kapitol, první čtyři kapitoly jsou teoretické, věnují se problematice tvorby řeči a popisu řečového traktu, charakterizaci a klasifikaci dysartrie s bližším zaměřením na hypokinetickou dysartrii, která je typická pro Parkinsonovu chorobu, metodám analýzy řečového signálu, programovacímu jazyku Python, vývojovému prostředí PyCharm a tvorbě knihovny. Uvedené oblasti jsou zpracované na základě studia odborné literatury a dalších zdrojů. Části věnující se tvorbě řeči a Parkinsonově nemoci byly navíc prokonzultovány s logopedem.

Pátá kapitola je praktická, zaměřená na návrh a realizaci systému a vyhodnocení získaných údajů. V rámci páté kapitoly byl popsán systém pro analýzu řečových signálů, který má za úkol provést objektivní diagnostiku hypokinetické dysartrie. Tento systém byl rozčleněn na tři části, tedy na načtení vstupních dat, předzpracování řečového systému a samotnou parametrizaci. Na základě členění, které bylo blíže rozpracováno v teoretické části, v její třetí kapitole, byly pro účely této práce vybrány parametry popisující fonaci: kmitočet základního tónu F_0 , jitter, shimmer, parametry hodnotící intenzitu řečového signálu: krátkodobá energie, parametry popisující tempo řeči: NST, parametry vztahující se k posouzení hybnosti artikulačních orgánů: formantové kmitočty, a parametry k posouzení kvality řečového signálu: ZCR, HNR a NHR, jež patří k základním parametrům při posuzování přítomnosti symptomů hypokinetické dysartrie u Parkinsonovy nemoci.

Pro posouzení funkčnosti vytvořeného systému byly užity nahrávky zdravých osob a osob s Parkinsonovou nemocí, bližší charakteristika těchto nahrávek je rovněž uvedena v praktické části bakalářské práce.

Realizace systému byla prováděna programovacím jazykem Python ve vývojovém prostředí PyCharm. K výpočtům některých parametrů bylo zároveň nutné využít programu Praat.exe. V rámci podkapitoly „Použité parametry“ byly definovány postupy sloužící k zavolání funkce, která umožňuje výpočet stanovených parametrů.

Při zpracování audionahrávek osob s Parkinsonovou nemocí a osob zdravých dochází v systému k získání velkého množství dat, výjimkou je NST, kde lze získat pouze jeden číselný údaj. Z toho důvodu nebyly jednotlivé parametry blíže rozpracovány, neboť tato bakalářská práce nemůže postřehnout rozsah získaných dat.

Stanovené cíle této bakalářské práce byly splněny, v průběhu vytváření knihovny pro zpracování parametrizace dysartrické řeči v jazyce Python a zpracovávání textové části bakalářské práce došlo k seznámení se s metodami analýzy řečového signálu u Parkinsonovy nemoci i s moderními parametrizačními technikami, které mají za úkol kvantifikovat poškození motorických aspektů řeči, a rovněž k implementaci vybraných parametrů v jazyce Python.

LITERATURA

- [1] KLENKOVÁ, J. *Logopedie: narušení komunikační schopnosti, logopedická prevence, logopedická intervence v ČR, příklady z praxe*. Praha: Grada, 2006, 228 s. ISBN 80-247-1110-9.
- [2] KRAHULCOVÁ, B. *Dyslálie - patlavost*. Praha: Beakra, 2007, 187 s. ISBN 978-80-903863-0-3.
- [3] Mluvní orgány [online] c2012 - 2018 [cit. 20. března 2018]. Dostupné na: <https://www.czechency.org/slovník/MLUVN%C3%8D%20ORG%C3%81NY>
- [4] CHROBOK, V., KUČERA, M., FRIČ, M. *Anatomie hlasotvorného ústrojí*. In. DRŠATA, J., a kol. *Foniatric - hlas*. Havlíčkův Brod: Tobiáš, 2011, 321 s. ISBN 978-80-7311-116-8.
- [5] Dýchací soustava [online] c2015 [cit. 20. března 2018]. Dostupné na: <https://volny-cas-uceni-cz.webnode.cz/prirodoveda/a5-rocnik/clovek/soustavy/dychaci-soustava/>
- [6] VATER, W., BAUDZIO, M. *Od prvního hlasu k prvním slovům*. Praha: Tech-Market, 1996, 104 s. ISBN 80-902134-0-5.
- [7] FRIČ, M., KUČERA, M., VYDROVÁ, J., ŠVEC, J. *Fyziologie a funkce hrtanu*. In. DRŠATA, J., a kol. *Foniatric - hlas*. Havlíčkův Brod: Tobiáš, 2011, 321 s. ISBN 978-80-7311-116-8.
- [8] JEDLIČKA, I. *Narušená komunikační schopnost v důsledku poruch hlasu*. In. ŠKODOVÁ, E., JEDLIČKA, I. *Klinická logopedie*. Praha: Septima, 2007, 616 s. ISBN 978-807367-340-6.
- [9] Stručná klinická anatomie dutiny ústní a hltanu [online] c2011 - 2018 [cit. 20. března 2018]. Dostupné na: <http://www.artmedicacentrum.cz/cs/orl-usni-nosnikre/strucna-klinicka-anatomie-dutiny-ustni-a-hltanu.html>
- [10] LOVE, R. J., WEBB, W. G. *Mozek a řeč*. Praha: Portál, 2009, 376 s. ISBN 978-80-7367-464-9.
- [11] NEUBAUER, K. *Terapie dysartrie*. In. LECHTA, V., a kol. *Terapie narušené komunikační schopnosti*. Praha: Portál, 2011, 392 s. ISBN 978-80-7367-901-9.
- [12] NEUBAUER, K. *Neurogení poruchy komunikace u dospělých*. Praha: Portál, 2007, 227 s. ISBN 978-80-7367-159-4.
- [13] CSÉFALVAY, Z. *Diagnostika dysartrie*. In. LECHTA, V. *Diagnostika narušené komunikační schopnosti*. Praha: Portál, 2003, 360 s. ISBN 80-7178-801-5.
- [14] NEUBAUER, K. *Dysartrie*. In. ŠKODOVÁ, E., JEDLIČKA, I. *Klinická logopedie*. Praha: Septima, 2007, 616 s. ISBN 978-807367-340-6.

- [15] CSÉFALVAY, Z., MEKYSKA, J., KOŠŤÁLOVÁ, M. *Diagnostika dysartrie*. In: CSÉFALVAY, Z., LECHTA, V. *Diagnostika narušené komunikační schopnosti u dospělých*. Praha: Portál, 2013, 232 s. ISBN 978-80-262-0364-3.
- [16] SMÉKAL, Z., MEKYSKA, J., GALÁŽ, Z., MŽOUREK, Z., REKTOROVÁ, I., FAÚNDEZ ZANUY, M. *Analysis of phonation in patients with Parkinson' s disease using empirical mode decomposition*. In *2015 International Symposium on Signals, Circuits and Systems (ISSCS)*. 2015. ISBN: 978-1-4673-7487- 3.
- [17] REKTOROVÁ I. *Současné možnosti diagnostiky a terapie Parkinsonovy nemoci*. *Neurol. prax* 2009; 10 (Suppl 2): 5–36.
- [18] MEKYSKA, J., SMÉKAL, Z., KOŠŤÁLOVÁ, M., MRAČKOVÁ, M., SKUTILOVÁ, S., REKTOROVÁ, I. *Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení*. *Cesk Slov Neurol*. 2011, (6): 74/107.
- [19] SKODDA, S., Wenke V., SCHLEGEL, U. *Gender-Related Patterns of Dysprosody in Parkinson Disease and Correlation Between Speech Variables and Motor Symptoms*. *Journal of Voice* [online]. 2011, 25(1), 76-82 [cit. 2018-04-29]. DOI: 10.1016/j.jvoice.2009.07.005. ISSN 08921997. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0892199709001131>
- [20] ZAMIŠKOVÁ, G., RESSNER, P., DLOUHÁ, J., ŠIGUTOVÁ, D. *Poruchy řeči u Parkinsonovy nemoci*. *Neurologie pro praxi*. 2010, 11(2): 112/116.
- [21] RŮŽIČKOVÁ, H. *Poruchy hlasu, řeči a komunikace u Parkinsonovy nemoci*. In: ROTH, J., SEKYROVÁ, M., RŮŽIČKOVÁ, E. et al. *Parkinsonova nemoc*. Praha: Maxdorf, 2009, 107 – 121. ISBN: 978-80-7345-178-3.
- [22] ROUBÍČKOVÁ, J. et al. *Test 3F – dysartrický profil*. Praha: Galén, 2011, 86 s. ISBN 978-80-7262-714-1
- [23] *Parkinsonova choroba: Léčba, možnosti a postupy* [online] c2016 [cit. 25. března 2018]. Dostupné na: <http://www.parkinsonovachoroba.cz/lecba/>
- [24] TSANAS, A. *Acoustic analysis toolkit for biomedical speech signal processing: concepts and algorithms*. Institute of Biomedical Engineering, Department of Engineering Science, Oxford Centre for Industrial and Applied Mathematics (OCIAM), Mathematical Institute, University of Oxford, UK, 2013.
- [25] SMÉKAL, Z. *Číslíkové zpracování signálů (MCSI)*. Elektronická skripta pro magisterská studia. Ústav telekomunikací, FEKT, VUT Brno, 2009.
- [26] MUCHA, J. *Akustická analýza přednesu básně u pacientů s Parkinsonovou nemocí: diplomová práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016, 68 s. Vedoucí práce byl Ing. Zoltán Galáž

- [27] ČERNOCKÝ, J. *Zpracování řečových signálů*. Ústav počítačové grafiky a multimédií, FIT, VUT Brno, 2006
- [28] KUČERA, M., FRIČ, M., HALÍŘ, M. *Praktický kurz hlasové rehabilitace a reedukace*. Opočno: M. Kučera, ORL ambulance – centrum hlasových poruch v Rychnově nad Kněžnou, 2010, 58 s. ISBN 978-80-254 – 6592 - 9.
- [29] MEKYSKA, J. *Analýza řečových promluv pro IT diagnostiku neurologických onemocnění: disertační práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2014, 160 s. Vedoucí práce byl prof. Ing. Zdeněk Smékal, CSc.
- [30] FRIČ, M., DRŠATA, J. *Akustické metody vyšetření hlasu*. In. DRŠATA, J., a kol. *Foniatric - hlas*. Havlíčkův Brod: Tobiáš, 2011, 321 s. ISBN 978-80-7311-116-8.
- [31] ČAPEK, K. *Vytvoření webové aplikace pro objektivní analýzu hypokineticke dysartrie ve frameworku Django: diplomová práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2014, 160 s. Vedoucí práce byl Ing. Zoltán Galáž.
- [32] SKODDA, S., RINSCH, H., SCHLEGEL, U. *Progression of Dysprosody in Parkinson's Disease Over Time – A Longitudinal Study. Movement Disorders*, [online]. 2009, 24, 2009, s. 716 - 722 [cit. 29. dubna 2018]. DOI: 10.1002/mds.22430. ISSN 08853185. Dostupné z: <https://www.ncbi.nlm.nih.gov/pubmed/19117364>
- [33] SILVA, D., OLIVEIRA, L.C., ANDREA, M. *Jitter Estimation Algorithms for Detection of Pathological Voices*. EURASIP Journal on Advances in Signal Processing [online] 2009, 2009(1), - [cit. 19. května 2018]. DOI: 10.1155/2009/567875. ISSN 1687-6180. Dostupné z: <https://asp-urasipjournals.springeropen.com/articles/10.1155/2009/567875>
- [34] The Python Tutorial [online] c2001 - 2018 [cit. 25. března 2018]. Dostupné na: <https://docs.python.org/3/tutorial/index.html>
- [35] PILGRIM, M. *Ponořme se do Python(u) 3*. Praha: CZ.NIC, 2010, 430 s. ISBN 978-80-904248-2-1
- [36] KOTEK, L. *Python – programovací jazyk pro výuku algoritmizace a programování*. Praha: Univerzita Karlova v Praze, Pedagogická fakulta, Katedra informačních technologií a technické výchovy, 2011, 73s. Vedoucí práce byl Ing. Jaroslav Novák, Ph.D.
- [37] Python 3000 Status Update [online] c1996 - 2018 [cit. 29. dubna 2018]. Dostupné na: <https://www.artima.com/weblogs/viewpost.jsp?thread=208549>
- [38] Dekorátory [online] c2015 [cit. 25. května 2018]. Dostupné na: <https://www.sallyx.org/sally/python/decorators.php>

- [39] Decorators [online] c2018 [cit. 25. května 2018]. Dostupné na: <http://python-3-patterns-idioms-test.readthedocs.io/en/latest/PythonDecorators.html>
- [40] Python Decorators [online] c2013 [cit. 25. května 2018]. Dostupné na: <http://www.siafoo.net/article/68>
- [41] Meet PyCharm: Quick Start Guide [online] c2000 - 2018 [cit. 29. dubna 2018]. Dostupné na: <https://www.jetbrains.com/help/pycharm/quick-start-guide.html>
- [42] Modules [online] c2001 - 2018 [cit. 25. května 2018]. Dostupné na: <https://docs.python.org/3/tutorial/modules.html>
- [43] Modules and Packages [online] c2018 [cit. 25. května 2018]. Dostupné na: https://www.learnpython.org/en/Modules_and_Packages
- [44] Moduly [online] c2015 [cit. 25. května 2018]. Dostupné na: <https://www.sallyx.org/sally/python/python11.php>

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

AR	Articulation Rate
CSS	Cascading Style Sheets
F_0	kmitočet základního hlasivkového tónu
F_1, F_2, F_3	formantové kmitočty
F_{2i}	druhý formantový kmitočet hlásky [i]
F_{2u}	druhý formantový kmitočet hlásky [u]
HNR	Harmonic to Noise ratio
HP	horní propust
HTML	Hypertext Markup Language
iCOMT	inhibitor Catechol-O-Methyltransferase prodlužující efekt L-dopa
IDE	integrated development environment
iMAO	inhibitor monoaminoxydázy zabraňující odbourávání dopaminu
L-dopa	L-3,4-dihydroxyphenylalanin, dopaminární léčba Parkinsonovy nemoci
LSTER	Low Short-Time Energy Ratio
LSVT	Lee Silverman Voice Treatment
NHR	Noise to Harmonic ratio
NSR	Net Speech Rate
NST	Net Speech Time
PPE	Pitch Period Entropy
p_{ram}	překrývání rámců
STE	Short-Time Energy

TKEO	Teager-Kaiser Energy operator
TPT	Total Pause Time
TSR	Total Speech Rate
TST	Total Speech Time
XML	Extensible Markup Language
ZCR	Zero-Crossing rate

SEZNAM PŘÍLOH

A OBSAH PŘILOŽENÉHO CD.....49

A OBSAH PŘILOŽENÉHO CD

Přiložené CD obsahuje soubory vytvořené v rámci této bakalářské práce. Jedná se o soubory vytvořené ve vývojovém prostředí PyCharm programovacím jazykem Python. Testovací audio nahrávky nejsou součástí tohoto CD.

Lze na tomto disku najít i elektronickou verzi bakalářské práce.