

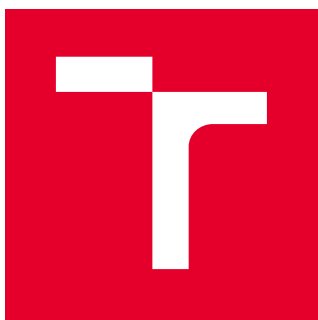
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2020

Bc. Daniel Kováč



**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

**DIFERENČNÍ ANALÝZA MULTILINGVÁLNÍHO ŘEČOVÉHO  
KORPUSU PACIENTŮ S NEURODEGENERATIVNÍMI  
ONEMOCNĚNÍMI**

DIFFERENTIAL ANALYSIS OF MULTILINGUAL CORPUS IN PATIENTS WITH NEURODEGENERATIVE DISEASES

**DIPLOMOVÁ PRÁCE**

MASTER'S THESIS

**AUTOR PRÁCE**

AUTHOR

**Bc. Daniel Kováč**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. Jiří Mekyska, Ph.D.**

**BRNO 2020**

# Diplomová práce

magisterský navazující studijní obor **Audio inženýrství**

Ústav telekomunikací

**Student:** Bc. Daniel Kováč

**ID:** 170614

**Ročník:** 2

**Akademický rok:** 2019/20

**NÁZEV TÉMATU:**

## **Diferenční analýza multilingválního řečového korpusu pacientů s neurodegenerativními onemocněními**

**POKYNY PRO VYPRACOVÁNÍ:**

V rámci této práce bude provedena parametrizace řečových a hlasových úloh zdravých řečníků a pacientů s Parkinsonovou a Alzheimerovou nemocí. Bude se jednat např. o prodlouženou fonaci, diadochokinetickou úlohu, monolog atd. Parametry budou kvantifikovat různé poruchy spojené s hypokinetickou dysartrií, např. monotónnost, nepřesnou artikulaci atd. Následně bude využito statistické zpracování a modely strojového učení. Pomocí těchto analýz bude zjištěno, které parametry mají vysokou diskriminační sílu a které jsou závislé na konkrétním jazyku mluvčího (bude uvážována čeština, americká angličtina a maďarština).

**DOPORUČENÁ LITERATURA:**

[1] BRABENEC, L., J. MEKYSKA, Z. GALAZ a Irena REKTOROVA. Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *Journal of Neural Transmission*. 2017, 124(3), 303-334. DOI: 10.1007/s00702-017-1676-0. ISSN 0300-9564.

[2] RUSZ, J., R. CMEJLA, H. RUZICKOVA a E. RUZICKA. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America*. 2011, 129(1), 350-367. DOI: 10.1121/1.3514381. ISSN 0001-4966.

**Termín zadání:** 3.2.2020

**Termín odevzdání:** 1.6.2020

**Vedoucí práce:** Ing. Jiří Mekyska, Ph.D.

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

**UPOZORNĚNÍ:**

Autor diplomové práce nesmí při vytváření diplomové 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**

Diplomová práce se zabývá automatizovanou diagnózou hypokinetické dysartrie v multilingválním řečovém korpusu. Jedná se o poruchu motorické realizace řeči vyskytující se u pacientů s neurodegenerativními onemocněními jako je například Parkinsonova nemoc. Automatizovaná diagnóza probíhá na základě akustické analýzy řeči a následným použitím matematických modelů. Tato metoda je na vzestupu díky její objektivitě a možné nezávislosti na národnosti. Cílem práce je zjistit, které akustické parametry mají vysokou diskriminační sílu a které jsou nezávislé na konkrétním jazyku mluvčího. K tom je využita statistická analýza parametrizovaných řečových úloh a následné modelování metodami strojového učení. Analýzy proběhly pro češtinu, americkou angličtinu, maďarštinu a všechny jazyky dohromady. Bylo zjištěno, že pouze některé parametry podporující diagnózu hypokinetické dysartrie a jsou nezávislé na jazyku mluvčího. Nejlepší výsledky vykazuje parametr relF2SD a po něm parametr NST. Při klasifikaci mluvčích všech jazyků dohromady model dosahuje přesnosti 59 % a senzitivity 72 %.

## **KLÍČOVÁ SLOVA**

Hypokinetická dysartrie, Parkinsonova nemoc, akustická analýza, akustické parametry, řečová úloha, XGboost, klasifikace.

## **ABSTRACT**

This diploma thesis focuses on the automated diagnosis of hypokinetic dysarthria in the multilingual speech corpus, which is a motor speech disorder that occurs in patients with neurodegenerative diseases such as Parkinson's disease. The automatic speech recognition approach to diagnosis is based on the acoustic analysis of speech and subsequent use of mathematical models. The popularity of this method is on the rise due to its objectivity and the possibility of working simultaneously on different languages. The aim of this work is to find out which acoustic parameters have high discriminative power and are universal for multiple languages. To achieve this, a statistical analysis of parameterized speech tasks and subsequent modelling by machine learning methods was used. The analyses were performed for Czech, American English, Hungarian and all languages together. It was found that only some parameters enable the diagnosis of the hypokinetic disorder and are, at the same time, universal for multiple languages. The relF2SD parameter shows the best results, followed by the NST parameter. When classifying speakers of all the languages together, the model achieves accuracy of 59 % and sensitivity of 72 %.

## **KEYWORDS**

Hypokinetic dysarthria, Parkinson's disease, acoustic analysis, acoustic parameters, speech task, XGboost, classification.

KOVÁČ, Daniel. *Diferenční analýza multilingválního řečového korpusu pacientů s neurodegenerativními onemocněními*. Brno, 2020, 54 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Jiří Mekyska, Ph.D.

## PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Diferenční analýza multilingválního řečového korpusu pacientů s neurodegenerativními onemocněními“ jsem vypracoval samostatně pod vedením vedoucího diplomové 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 diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové 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 autorského 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.

Brno .....

.....

podpis autora

## PODĚKOVÁNÍ

Rád bych poděkoval především vedoucímu diplomové práce panu Ing. Jiřímu Mekyskovi, Ph.D. za odborné vedení, cenné rady, vstřícný přístup a ochotu. Dále bych rád poděkoval svým rodičům Gabriele a Pavlovi Kováčovým za jejich trpělivost a podporu během svého celého studia. Mé poděkování také patří všem, kdo přišli a zúčastnili se výzkumu, bez kterého by tato práce nemohla vzniknout.

# Obsah

<b>Úvod</b>	<b>11</b>
<b>1 Parkinsonova nemoc</b>	<b>12</b>
1.1 Fáze Parkinsonovy nemoci . . . . .	12
<b>2 Hypokinetická dysartrie</b>	<b>13</b>
2.1 Vliv bazálních ganglií na produkci řeči . . . . .	13
2.2 Projevy Hypokinetické dysartrie . . . . .	14
2.3 Diagnóza hypokinetické dysartrie . . . . .	16
2.3.1 Percepční analýza . . . . .	16
2.3.2 Akustická analýza . . . . .	16
<b>3 Automatizovaná diagnóza hypokinetické dysartrie</b>	<b>17</b>
<b>4 Kohorta a metodologie</b>	<b>21</b>
4.1 Nahrávání řeči . . . . .	21
4.2 Řečové úlohy . . . . .	21
4.3 Popis databáze . . . . .	22
4.4 Segmentace nahrávek . . . . .	24
4.5 Akustické parametry . . . . .	24
4.6 Statistická analýza . . . . .	26
4.7 Vícerozměrná analýza . . . . .	27
4.7.1 Metoda eXtreme Gradient Boosting . . . . .	27
<b>5 Výsledky</b>	<b>29</b>
5.1 Jednorozměrná analýza . . . . .	29
5.2 Vícerozměrná analýza . . . . .	38
<b>6 Diskuze</b>	<b>41</b>
6.1 Jednorozměrná analýza . . . . .	41
6.2 Vícerozměrná analýza . . . . .	42
<b>Závěr</b>	<b>44</b>
<b>Literatura</b>	<b>45</b>
<b>Seznam symbolů, veličin a zkratk</b>	<b>52</b>



Seznam příloh	53
A Obsah přiloženého DVD	54

## Seznam obrázků

2.1	Anatomie bazálních ganglií . . . . .	14
4.1	Segmentace nahrávek na jednotlivé řečové úlohy . . . . .	24
4.2	Matice záměn modelu XGBoost. . . . .	28
5.1	Rozdíl parametru relF2SD z úlohy TSK1 mezi PN a ZK pro cs . . . .	30
5.2	Rozdíl parametru relF0SD z úlohy TSK1 mezi PN a ZK pro en-US .	30
5.3	Rozdíl parametru NST z úlohy TSK2 mezi PN a ZK pro en-US . . .	31
5.4	Rozdíl parametru relF2SD z úlohy TSK2 mezi PN a ZK pro cs . . . .	31
5.5	Rozdíl parametru relF0SD z úlohy TSK2 mezi PN a ZK pro en-US .	31
5.6	Rozdíl parametru SPIR z úlohy TSK2 mezi PN a ZK pro en-US . . .	32
5.7	Rozdíl parametru SPIR z úlohy TSK2 mezi PN a ZK pro vj . . . . .	32
5.8	Rozdíl parametru mean HNR z úlohy TSK3 mezi PN a ZK pro en-US	32
5.9	Rozdíl parametru median HNR z úlohy TSK3 mezi PN a ZK pro en-US	33
5.10	Rozdíl parametru relF2SD z úlohy TSK3 mezi PN a ZK pro vj . . . .	33
5.11	Rozdíl parametru median relF2SD z úlohy TSK4 mezi PN a ZK pro cs	33
5.12	Rozdíl parametru shimmer z úlohy TSK6 mezi PN a ZK pro en-US .	34
5.13	Rozdíl parametru VAI ze všech úloh mezi PN a ZK pro en-US . . . .	34
5.14	Důležité parametry datasetu cs . . . . .	39
5.15	Důležité parametry datasetu en-US . . . . .	39
5.16	Důležité parametry datasetu hu . . . . .	40
5.17	Důležité parametry datasetu vj . . . . .	40

# Seznam tabulek

1.1	Fáze Parkinsonovy nemoci . . . . .	12
3.1	Studie zabývající se automatizovanou diagnózou HD . . . . .	17
4.1	Seznam řečových úloh . . . . .	22
4.2	Zastoupení pohlaví v každé skupině v kombinaci s diagnózou . . . . .	22
4.3	Počet účastníků a jejich věkové zastoupení pro češtinu . . . . .	23
4.4	Počet účastníků a jejich věkové zastoupení pro americkou angličtinu . . . . .	23
4.5	Počet účastníků a jejich věkové zastoupení pro maďarštinu . . . . .	23
4.6	Seznam akustických parametrů . . . . .	25
4.7	Mřížka pro optimalizaci hyperparametrů . . . . .	27
5.1	Výsledky statistických testů akustických parametrů . . . . .	29
5.2	Výsledky deskriptivní statistiky: aritmetický průměr . . . . .	35
5.3	Výsledky deskriptivní statistiky: medián . . . . .	36
5.4	Výsledky deskriptivní statistiky: směrodatná odchylka . . . . .	37
5.5	Optimální hyperparametry modelů jednotlivých jazykových skupin . . . . .	38
5.6	Úspěšnosti modelů strojového učení v % . . . . .	38
6.1	Přehled signifikantních parametrů na základě statistické analýzy . . . . .	42

# Úvod

Parkinsonova nemoc je dnes stále častěji vyskytující se neurodegenerativní onemocnění postihující převážně starší lidi. Mimo věk jsou hlavními faktory genetická predispozice a pohlaví. Stále nebyl objeven lék, který by nemoc vyléčil, pouze je možné zmírnit její průběh a ten závisí na včasné diagnóze nemoci. Jedna z významných oblastí, kde se nemoc projevuje je řeč a poruchy s ní spojené označované jako hypokinetická dysartrie. Tyto poruchy je možné detekovat již v rané fázi onemocnění a proto se jedná o důležitý neinvazivní biologický marker. Hypokinetická dysartrie se projevuje především poruchami fonace, artikulace a prozodie. Dnes se pro její diagnózu využívají dvě metody. V první se jedná o subjektivní posuzování klinickým logopedem a druhá využívá analýzu akustických parametrů vypočítaných z nahraného signálu řeči pacienta. Druhá metoda je dnes na vzestupu a to hlavně díky její objektivitě a možné nezávislosti na národnosti. Dnes je předmětem mnoha studií, přičemž se testují stále nové parametry, které by dokázaly poruchy řeči správně kvantifikovat. Akustická analýza řeči má velký potenciál v telemedicině a snahou je najít systém se standardizovaným setem akustických parametrů, který by dokázal automaticky rozpoznávat a monitorovat Parkinsonovu nemoc na dálku za pomoci běžně dostupných elektronických zařízení, jako je chytrý telefon, hodinky nebo tablet a to nezávisle na tom, jakým jazykem dotyčný hovoří.

Protože existuje jen pár studií analyzujících závislost akustických parametrů hypokinetické dysartrie na národnosti, zabývá se tato práce především tímto tématem. **Cílem je zjistit, které akustické parametry hypokinetické dysartrie mají vysokou diskriminační sílu a zároveň, které jsou nezávislé na konkrétním jazyku mluvčího.** Z hlediska jazyka bude pozornost věnována češtině, americké angličtině a maďarštině a to z důvodu rozdílných jazykových skupin, do kterých spadají. Dále budou natrénovány modely strojového učení, které budou schopny zjistit přítomnost hypokinetické dysartrie s určitou senzitivitou a specificitou.

V práci je nejdříve popsána Parkinsonova nemoc a její fáze. Následně se zabývá hypokinetickou dysartrií, v jakých oblastech se projevuje a jakým způsobem se diagnostikuje. Další kapitola obsahuje soubor studií do roku 2020 které se zabývají automatizovanou diagnózou hypokinetické dysartrie. V následující kapitole je popsána kohorta a metodologie práce. Nakonec jsou uvedeny výsledky, které jsou diskutovány v poslední kapitole.

# 1 Parkinsonova nemoc

*Parkinsonova nemoc* (PN) je chronická neurodegenerativní porucha charakterizovaná progresivní degenerací dopaminergních neuronů v černé substantii (substantia nigra) [1]. Nedostatek dopaminu v této části středního mozku způsobuje potíže spojené se zhoršením motorických i nemotorických funkcí, poruchami spánku, řeči, emocionálními změnami a řadou dalších deficitů [2].

Riziko onemocnění PN s rostoucím věkem stoupá. Dalšími významnými faktory jsou pohlaví a genetická predispozice. Rovněž je zřejmé, že počet onemocněných přibývá s rostoucí populací a střední délkou života [3].

Nemoc poprvé popsal v roce 1817 anglický lékař James Parkinson [4]. Přestože věda udělala od této chvíle značný pokrok, dodnes neznáme pravou příčinu PN a nejsme ji schopni vyléčit. Je možné však zmírnit její průběh a to buď neinvazivně např. užitím levodopy, aminokyseliny, která se v mozku přeměňuje na dopamin, a nebo rizikověji chirurgicky [5]. Včasné zjištění a zahájení léčby má zásadní vliv na další průběh nemoci.

## 1.1 Fáze Parkinsonovy nemoci

V Tabulce 1.1 je popsáno pět fází PN podle Hoehna a Yahra jakožto škála, kterou lékaři na celém světě používají ke klasifikaci pacientů ve výzkumných studiích [6].

Tab. 1.1: Fáze Parkinsonovy nemoci

Fáze PN	Závažnost	Symptomy
1 - raná	mírná	třes jedné ruky, tuhost, neohrabaná jedna noha, ztráta výrazu na jedné straně tváře
2 - raná	mírná	<b>řečové abnormality</b> , ztráta výrazu tváře na obou stranách, zhoršené mrkání, ztuhlost svalů
3 - střední	střední	problémy s rovnováhou způsobené zpomalenými reflexy, všechny ostatní symptomy
4 - pokročilá	vážná	neschopnost vykonávat každodenní potřeby, potřebná asistence
5 - pokročilá	vážná	riziko pádu při vstávání a otáčení těla, zamrznutí při chůzi, možné halucinace nebo bludy

## 2 Hypokinetická dysartrie

Jeden z nevýraznějších problémů postihujících onemocněné Parkinsonovou nemocí jsou poruchy řeči, kterými trpí až 90 % pacientů [7]. Tyto poruchy jsou spojené zejména se špatnou artikulací, fonací, prozodií a plynulostí řeči, mají zásadní vliv na komunikační schopnost člověka a souhrně se označují pod pojmem *hypokinetická dysartrie* (HD) [8]. Z 87 % se tyto poruchy vyskytují u pacientů s neurodegenerativními onemocněními, jako je PN nebo různé typy demence. V ostatních případech se na vzniku HD může podílet i cévní mozková příhoda, hypoxie, trauma nebo jiné příčiny [9].

### 2.1 Vliv bazálních ganglií na produkci řeči

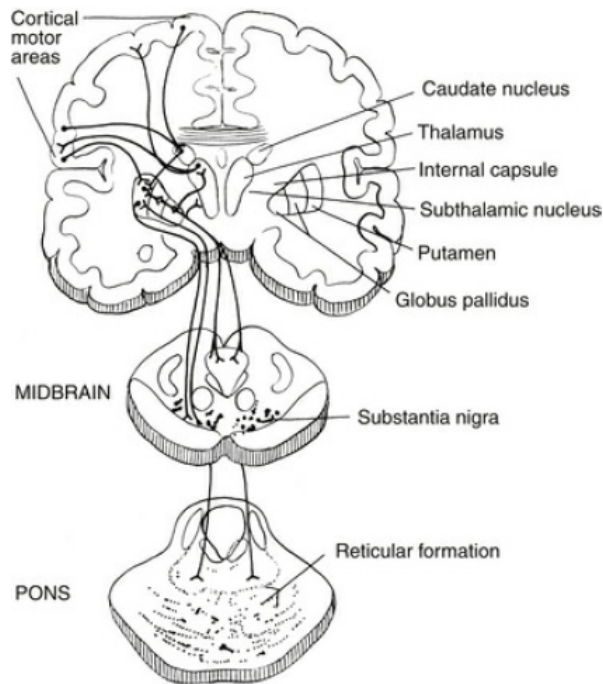
Řeč vzniká za pomoci motorických řečových procesů rozdělených do třech fází:

1. Nejprve myšlenka nebo pocit generují impuls verbálně komunikovat. Tyto informace musí být uspořádány a převedeny do kódu, který je řízen pravidly příslušného jazyka. Tato aktivita se označuje jako kognitivně-lingvistický proces.
2. Další činností je motoricko-řečové plánování a řízení, jež spočívá ve výběru určitých sensorimotorických programů, které aktivují řečové svaly v odpovídající intenzitě a čase.
3. V poslední fázi se provede inervace dýchacích, fonačních, rezonančních a artikulačních svalů. Centrální a periferní nervový systém vede nervový vzruch do mluvních orgánů. Svalové kontrakce produkují akustický signál, který nese danou informaci.

Motorický řečový systém zasahuje do více úrovní nervového systému člověka. Jedná se o komplexní síť struktur a signálových cest, které řídí veškerý pohyb svalů. Kognitivně-lingvistické procesy v mozku probíhají v Brocově centru řeči. Zde také vznikají vzorce, které jsou posílány jednotlivým svalům potřebným k vyslovení slova.

Sensorimotorické interakce jsou však také nezbytnou součástí správné tvorby řeči. Pomocí zobrazovacích metod bylo při zkoumání mozku zjištěno významné zapojení bazálních ganglií (Obr. 2.1). Okruhy bazálních ganglií mají silnou funkční vazbu jak na různé oblasti mozkové kůry, tak na extrapyramidový systém, který je zodpovědný za udržení svalového napětí. Tyto součásti šedé hmoty koncového mozku, ke kterým se řadí i černá substance, hrají klíčovou roli v motorických činnostech člověka a poskytují jistou signálovou platformu pro vznik obecných mechanismů akustické komunikace lidí se schopností artikulovat [10].

Hlavní funkcí tohoto okruhu je správná regulace svalového napětí upravováním informací vystupujícího z mozkové kůry. Tlumí kortikální signál tak, aby odpovídal původním pohybovým cílům. Bez této modulace by byl nad rámec potřebného pro přesné vykonání původně zamýšleného pohybu. U hypokinetické dysartrie je však tento tlumící účinek příliš vysoký, což se přiřazuje nerovnováze v neurotransmitrech – v případě onemocnění PN se jedná o úbytek dopaminu v této oblasti. Poškození okruhu tímto způsobem buďto snižuje pohyb nebo brání v jeho vykonání.



Obr. 2.1: Anatomie bazálních ganglií [9]

## 2.2 Projevy Hypokinetické dysartrie

Hypokinetická dysartrie úzce souvisí se špatnou motorickou funkcí člověka. Pacienti s HD často zaznamenávají klidový třes končetin, hlavy, ale i jazyka, čelisti nebo rtů. Vše je doprovázeno špatnou faciokinezí. Takového člověka můžeme vnímat bez emocí a s maskovanou tváří, přestože to nemusí odpovídat vnitřnímu emocionálnímu stavu. Jeden z nejvýraznějších projevů HD je poškození hlasu a poruchy řeči s tím spojené. Špatná artikulace a fonace, ze které se částečně odvíjí dysprozodie, jsou základními znaky poskytující užitečné informace a jsou zásadní pro odlišení typu dysartrie. Charakteristické rysy hlasu člověka postiženého HD jsou: chrapot, hrubost, chvění nebo

velké zastoupení slyšitelného dechu. V řeči se dále projevuje hypernazálností, opakovatelností slabik, jejich zkracováním nebo prodlužováním, nadměrným výskytem pauz a špatnou artikulací, což vede k nesrozumitelnosti. Celkový řečový projev je spíš tichý s malou variací hlasitosti i výšky hlasu. Přestože ostatní na nich pozorují tyto neobvyklosti, lidé zasažení HD tyto změny můžou popírat. Sami však uvádí, že jim občas dělá problém s řečí začít [9].

## **Fonace**

Hlas vzniká rozkmitáním hlasivek v důsledku působení síly proudícího vzduchu z plic, což souvisí se správnou respirací. Ta je u člověka postiženého HD nepravdělná kvůli nesprávné funkci respiračních svalů a tím vznikají v řeči pacienta neočekávané pauzy. Síla vydechovaného vzduchu bývá často slabá, což zapříčiňuje úbytek hlasitosti ve zvukovém projevu. Ochablé hlasivkové svaly hrtanu způsobují dysfonii projevující se chraptivostí hlasu a hlasovým třesem. Někdy pacienti projevují známky hypernazality, jejíž příčinou je velký objem vzduchu procházející nosní dutinou v poměru k té ústní [8] [9].

## **Artikulace**

Artikulační svaly jako jsou hrtan, rty nebo jazyk vykazují rigiditu a snížený rozsah pohybu. To se nejvíce projevuje ve špatném vyslovování souhlásek. Největším problémem se stávají ploviva, která někdy pacienti nejsou schopni vyslovit a to především ve chvíli, kdy se rychle opakují. Výsledná řeč se tak stává nesrozumitelná a nejví se přirozeně. U samohlásek jsou důležité především rezonance, které se mění v závislosti na tvaru ústní, nosní a hrdelní dutiny. Tvar dutin závisí také na činnosti artikulačních svalů a proto vznikající formanty jsou další oblastní, kde se HD projevuje [8] [9].

## **Prozodie**

Pro malý obsah vzduchu v plicích se věty a fráze řečníka stávají uspěchané, tempo řeči klesá nebo roste a počáteční slabiky slov jsou často opakovány. Někdy se opakují celá slova nebo věty na konci řeči – takovému jevu se říká palilalie. Intonace, rytmus, přízvuk, počet pauz a tempo řeči jsou hlavní suprasegmentální jevy, které zkoumá prozodie. U pacientů s HD je většina těchto znaků ovlivněna ztuhlostí svalů a to má vliv na sémantiku řeči, ale i na špatném vyjadřování emocí. V takovém případě se hovoří o tzv. dysprozodii [8] [9].



## 2.3 Diagnóza hypokinetické dysartrie

Pro diagnózu dysartrie je potřeba zaměřit se na vyšetření řeči a to provádí logoped buď subjektivně poslechem nebo vyhodnocováním naměřených akustických parametrů. Tyto dvě metody se mohou kombinovat. Při posuzování je však třeba nejdříve vyloučit ostatní poruchy řeči, jako je afázie, orální nebo verbální apraxie a jiné. Poté se může pokročit k samotné diagnóze dysartrie [11].

### 2.3.1 Percepční analýza

Jedná se o subjektivní hodnocení na základě klinických dojmů logopeda. Existuje několik diagnostických postupů, přičemž v Česku se nejvíc používá tzv. Test 3F Dysartrický profil [12]. Zkratka 3F vznikla ze tří testovaných funkčních celků:

- F1 – Faciokineze: rty, čelist, jazyk.
- F2 – Fonorespirace: respirace, respirace při fonaci, fonace.
- F3 – Fonetika: artikulace, prozodie, srozumitelnost.

V každém celku jsou tedy 3 oblasti a každá z nich vyšetřována pěti úkoly. Každý je hodnocen maximálně dvěma body za správné vykonání a po sečtení pacient obdrží tzv. dysartrický index. Tím je dána míra dysartrického postižení:

- 85–90 bez poruchy
- 74–85 velmi lehká dysartrie
- 57–73 lehká dysartrie
- 36–56 středně těžká dysartrie
- 17–35 těžká dysartrie
- 0–17 velmi těžká dysartrie -- anartrie.

### 2.3.2 Akustická analýza

Jedná se o objektivní hodnocení na základě analýzy akustických parametrů vypočítaných z nahrávek řečového signálu pacientů. Tyto parametry kvantifikují míru postižení HD v mnohem větší škále než při percepční analýze. Mezi konvenční používané parametry se řadí např. základní frekvence řeči, jitter, shimmer, podíl tonální složky a šumu atd. Nahrazují tak jednotlivé konkrétní poruchy jako je třeba monotónnost nebo hrubost řeči. Dnes existuje již řada dalších nekonvenčních akustických parametrů a nové se stále testují [13]. Další výhodou tohoto přístupu je nejen to, že odstraňuje subjektivitu, ale navíc některé parametry mohou být nezávislé na národnosti. Tím stejné analýzy mohou probíhat v různých zemích, které nemají percepční testy pro svůj vlastní jazyk. Rizikem zde naopak může být zkreslení hodnot parametrů v důsledku nesprávného nahrávání signálu řeči [14].

### 3 Automatizovaná diagnóza hypokineticé dysartrie

Již několik let se vědci intenzivně snaží definovat vliv akustické analýzy řeči na diagnóze HD v PN za použití metod strojového učení. Cílem je představit akustickou analýzu jako neinvazivní marker PN, který lze automaticky rozpoznávat. Proces automatické diagnózy lze popsat ve dvou krocích:

1. Použitím akustické analýzy se z řeči vypočítají akustické parametry, které kvantifikují HD ve všech projevujících se oblastech.
2. Aplikují se matematické modely, které využívají tyto akustické parametry k zjištění přítomnosti HD s určitou klasifikační přesností (ACC), senzitivitou (SEN) a specificitou (SPE) [15].

V Tabulce 3.1 jsou vypsané studie týkající se automatizované diagnózy HD do roku 2020.

Tab. 3.1: Studie zabývající se automatizovanou diagnózou HD

Studie a rok	Jazyk a databáze	Řečová úloha	Parametry	Výsledky
Şule Yücelbaş (2020) [16]	tr - PN:188, ZK: 64	prodloužená fonace [a]	KONV/ NEKONV	ACC = 89 % (muži) ACC = 87 % (ženy)
Moro-Velazquez et al. (2019) [17]	es - PN:47, ZK: 32 es - PN:50, ZK: 50 cs - PN:20, ZK: 14	DDK, čtení, monolog	NEKONV	multilingvální studie; ACC = 85 - 94 %
Vásquez-Correa et al. (2019) [18]	es - PN:50, ZK: 50 de - PN:88, ZK: 88 cs - PN:50, ZK: 50	čtení, věty, monolog, DDK	KONV	multilingvální studie; ACC = 72 % (es) ACC = 77 % (de) ACC = 73 % (cs)
Sztahó et al. (2019) [19]	hu - PN:55, ZK: 33	prodloužená fonace [a], čtení, věty, slova DDK, monolog	KONV/ NEKONV	ACC = 89 % SEN = 90 % SPE = 88 %
Dahmani a Guerti (2018) [20]	de - PN:40, ZK: 40	prodloužená fonace [a], [i], [u], věty	KONV/ NEKONV	ACC = 89 % SEN = 90 %
Moro-Velazquez et al. (2018) [21]	es - PN:50, ZK: 50 es - PN:47, ZK: 32 de - PN:88, ZK: 88 cs - PN:20, ZK: 16	DDK	NEKONV	multilingvální studie; ACC = 74(de) - 94(cs) % SEN = 74(de) - 90(cs) % SPE = 74(de)- 100(cs) %
Tündik et al. (2017) [22]	hu - PN:49, ZK: 49	čtení	NEKONV	ACC = 86 %

Tab. 3.1: Pokračování z předchozí strany

Studie a rok	Jazyk a databáze	Řečová úloha	Parametry	Výsledky
Hsu et al. (2017) [23]	zh - PN:11, ZK: 7	čtení	KONV	explorativní analýza; PD = nižší VAI oproti en
Lahmiri et al. (2017) [24]	en - PN:51, ZK: 27	čtení, monolog	KONV/ NEKONV	ACC = 92 % SEN = 95 % SPE = 91 %
Mucha et al. (2017) [25]	cs - PN:99, ZK: 53	čtení básně	KONV	ACC = 80 % SEN = 83 % SPE = 74 %
Godino- Llorente et al. (2017) [26]	es - PN:50, ZK: 50	DDK	NEKONV	ACC = 85 % SEN = 82 % SPE = 88 %
Naranjo et al. (2016) [27]	es - PN:40, ZK: 40	prodloužená fonace [a]	KONV/ NEKONV	ACC = 75 % SEN = 72 % SPE = 79 %
Orozco- Arroyave et al. (2016) [28]	es - PN:50, ZK: 50 de - PN:88, ZK: 88 cs - PN:20, ZK: 16	DDK, věty, čtení, slova	KONV/ NEKONV	multilingvální studie; ACC = 85 - 99 %
Mekyska et al. (2016) [29]	cs - PN:83, ZK: 49	prodloužená fonace [a], [e], [i], [o], [u]	KONV/ NEKONV	ACC = 92 % SEN = 93 % SPE = 92 %
Galaz et al. (2016) [30]	cs - PN:98, ZK: 51	čtení rytmického textu	KONV	parametry prozodie; důležitost rytmiky a přízvuku v řeči
Smekal et al. (2015) [31]	cs - PN:84, ZK: 49	prodloužená fonace [a], [e], [i], [o], [u]	KONV/ NEKONV	ACC = 94 % SEN = 96 % SPE = 90 %
Mekyska et al. (2015) [32]	cs - PN:84, ZK: 49	prodloužená fonace [a], [e], [i], [o], [u]	KONV/ NEKONV	ACC = 89 % SEN = 92 % SPE = 84 %
Vásquez- Correa et al. (2015) [33]	es - PN:14, ZK: 14	monolog, čtení	KONV/ NEKONV	algoritmus, který zlepšuje řečové nahrávky – přesnost se zvýšila o 11 %
Musa et. al (2015) [34]	en - PN:23, ZK: 8	prodloužená fonace [a],	KONV/ NEKONV	ACC = 98 % SEN = 99 % SPE = 99 %
Arora et al. (2015) [35]	en - PN:10, ZK: 10	prodloužená fonace [a]	KONV/ NEKONV	SEN = 96 % SPE = 97 %
Orozco- Arroyave et al. (2015) [36]	es - PN:50, ZK: 50 de - PN:88, ZK: 88 cs - PN:42, ZK: 42	čtení, monolog	NEKONV	multilingvální studie; čtení - ACC = 91 - 98 % monolog - ACC >98 %

Tab. 3.1: Pokračování z předchozí strany

Studie a rok	Jazyk a databáze	Řečová úloha	Parametry	Výsledky
Benba et al. (2015) [37]	tr - PN:17, ZK: 14	prodloužená fonace [a]	KONV/ NEKONV	ACC = 91.17 %
Hariharan et al. (2014) [38]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV/ NEKONV	ACC = 100 % SEN = 100 % SPE = 100 %
Orozco- Arroyave et al. (2014) [39]	es - PN:50, ZK: 50 de - PN:88, ZK: 89 cs - PN:42, ZK: 42	DDK, věty, monolog, slova	KONV/ NEKONV	multilingvální studie; slova - ACC = 91 - 98 % (es) slova - ACC = 84 - 96 % (de) DDK - ACC = 97- 99 % (de)
Novotny et al. (2014) [40]	cs - PN:24, ZK: 22	DDK	KONV	ACC = 88 % SEN = 86 % SPE = 73 %
Villa-Canas et al. (2014) [41]	es - PN:50, ZK: 50	prodloužená fonace [a], [e], [i], [o], [u]	KONV	vokál [i] - ACC = 71 % vokál [i] - SEN = 61 % vokál [i] - SPE = 73 %
Orozco- Arroyave et al. (2014) [42]	es - PN:50, ZK: 50	prodloužená fonace [a], [e], [i], [o], [u]	KONV/ NEKONV	ACC = 81 % SEN = 83 % SPE = 81 %
Orozco- Arroyave et al. (2013) [43]	es - PN:20, ZK: 20	prodloužená fonace [a], [e], [i], [o], [u]	NEKONV	ACC = 70 - 76 % vokál [i] - ACC = 76,81 %
Rusz et al. (2013) [44]	cs - PN:20, ZK: 15	prodloužená fonace [a], [o], [u], opakování vět, čtení, monolog	KONV	monolog - ACC = 76,81 %
Belalcazar- Bolaños et al. (2013) [45]	es - PN:100, ZK: 100	prodloužená fonace [a], [e], [i], [o], [u]	KONV/ NEKONV	vokál [i] - ACC = 67 %
Sekar et al. (2013) [46]	tr - PN:20, ZK: 20	prodloužená fonace [a], [e], [i], [o], [u], čísla, slova, krátké věty	KONV	ACC = 85 % SEN = 80 % SPE = 90 %
Shahbakhti et al. (2013) [47]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV/ NEKONV	ACC = 91,5 % SEN = 90 % SPE = 91 %
Shahbakhti et al. (2013) [48]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV/ NEKONV	ACC = 96 %
Belalcazar- Bolaños et al. (2013) [49]	es - PN:40, ZK: 40	prodloužená fonace [a], [e], [i], [o], [u]	KONV	ACC = 80 % vokál [e] - ACC = 92 %
Orozco- Arroyave et al. (2014) [50]	es - PN:20, ZK: 20	prodloužená fonace [a], [e], [i], [o], [u]	NEKONV	ACC = 81 % SEN = 82 % SPE = 79 %

Tab. 3.1: Pokračování z předchozí strany

Studie a rok	Jazyk a databáze	Řečová úloha	Parametry	Výsledky
Tsanas et al. (2012) [51]	en - PN:33, ZK: 10	prodloužená fonace [a]	KONV/ NEKONV	ACC = 90 %
Hazan et al. (2012) [52]	en - PN:16, ZK: 14 de - PN:68, ZK: 30	prodloužená fonace [a], [e], [i], [o], [u]	KONV	multilingvální studie; separátně - ACC = 85 % vj - ACC = 75 %
Baker et al. (2011) [53]	en - PN:25, ZK: 25	prodloužená fonace [a]	KONV/ NEKONV	ACC = 94,2 %
Aström a Koker (2011) [54]	en - PN:25, ZK: 25	prodloužená fonace [a]	KONV/ NEKONV	ACC = 91 % SEN = 91 % SPE = 93 %
Rusz et al. (2011) [55]	cs - PN:25, ZK: 25	prodloužená fonace [a], [e], [i], [o], [u], DKK, čtení, monolog, rytmické četní	KONV/ NEKONV	78 % pacientů PN mělo HD, nejvíce byla postižena prozodie, pak artikulace
Bocklet et al. (2011) [56]	cs - PN:23, ZK: 23	prodloužená fonace [i], DDK, čtení, monolog	KONV/ NEKONV	ACC = 91 %
Skodda et al. (2011) [57]	hu - PN:68, ZK: 32	čtení	KONV	snížené VAI v poměru k ZK
Shirvan a Tahami (2011) [58]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV	ACC = 98 %
Spadoto et al. (2010) [59]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV/ NEKONV	ACC = 75,37 %
Sakar a Kursun (2009) [60]	en - PN:23, ZK: 8	prodloužená fonace [a]	KONV/ NEKONV	ACC = 81,53 %

Tabulka obsahuje 45 studií, přičemž 7 z nich je multilingválních (cs, en, de, es) a zbylých 38 je zaměřeno na jeden konkrétní jazyk (cs, en, hu, es, de, tr, zh). Studie s největší databází pro jeden jazyk obsahuje 252 nahrávek (188 PN, 64 ZK) [16] a s nejmenší databází 18 nahrávek (11 PN, 7 ZK) [23]. Nejčastěji použitá řečová úloha je prodloužená fonace vokálu [a]. Multilingvální studie se nejvíce zabývaly testováním modelů natrénovaných stejnými parametry pro každý jazyk zvlášť a pak byly porovnány jejich klasifikační přesnosti. V jednom případě byl model trénován dvěma jazyky a testován jedním [17]. Ve dvou případech byl zvolen přístup natrénování modelu databází s jedním jazykem a pak testování na jazyku druhém, kde byl zjištěn pokles klasifikační přesnosti [28] [52]. Pouze jedna se zabývala trénováním modelu více jazyky a testováním více jazyky [52]. V pracích je rovněž pouze jednou obsaženo přímé porovnání parametrů vypočítaných z databází různých jazyků [23].

## 4 Kohorta a metodologie

### 4.1 Nahrávání řeči

Nahrávání řeči pacientů provádí neurologové příslušných národností, kteří s pacienty běžně pracují. Aby se zabránilo rozdílům v kvalitě nahrávek, byla sepsána jistá doporučení, kterými se neurologové mají při nahrávání řídit:

- Řeč bude nahrávána obyčejným stolním počítačem nebo notebookem.
- Bude použit kondenzátorový mikrofon s kardioidní směrovou charakteristikou.
- Mikrofon bude namontován na stojan a postaven do přibližné vzdálenosti 20 cm od pusy řečníka. V případě nahrávání obrazu kamerou zaměřenou na tvář pacienta je možné mikrofon umístit výše.
- Mikrofon by měl být opatřen odpruženým držákem, aby se předešlo snímání nežádoucích vibrací.
- Mikrofon bude připojen k externímu zvukovému rozhraní, které podporuje 48 kHz vzorkovací frekvenci. Zařízení bude dále připojeno k počítači nebo notebooku.
- Zvukový signál bude nahráván přes software Audacity jako jeden kanál vzorkovaný s vzorkovací frekvencí 48 kHz.
- Je třeba zajistit, aby nahrávání proběhlo v tiché místnosti s co nejmenším odrazem zvuku a nízkou hladinou okolního šumu jako je: zvuk ulice, klimatizace, větrák v počítači, hodiny, telefon, různé zdravotnické zařízení, kroky na schodišti atd. Pokud je vnímán velký dozvuk místnosti, lze využít určitých závěsů nebo jiných textilií k potlačení echa. Doporučuje se především takto zakrýt okna v místnosti, protože sklo silně odráží akustické vlnění.

Před samotným nahráváním je požádáno o zaslání zkušebních testových nahrávek, aby byla kvalita zvuku ověřena. Všechny nahrávky jsou pak odeslány na sdílený FTP server a informace o nahrávce poté vyplněny do sdílené tabulky v Google dokumentu.

### 4.2 Řečové úlohy

Aby se mohly analyzovat všechny oblasti řeči, kde se HD vyskytuje, bylo vytvořeno několik řečových úloh, které každý účastník postupně vykonává. Úlohy jsou vymyšleny tak, aby se vyšetřila fonace, artikulace i prozodie – zaměřují se na jednotlivé oblasti, ale zároveň každá úloha může být prostředkem pro zkoumání více oblastí. V Tabulce 4.2 je seznam všech vykonávajících úloh, jejich označení, název a popis.

Tab. 4.1: Seznam řečových úloh

Označení	Řečová úloha	Popis
TSK1	Monolog	Nejméně 90 s dlouhý monolog bez přerušení klinikem. Účastníci hovoří o svých koníčcích, rodině, zaměstnání, aktuální činnosti atd.
TSK2	Čtení	Čtení krátkého textu. Pacientovi je umožněno přečíst si text předem.
TSK3	Prodloužená fonace	Přibližně třísekundová (ne delší než 5 s) prodloužená samohláska [a] v pohodlné výšce a hlasitosti. Vykonáno na jeden nádech.
TSK4	Prodloužená fonace	Přibližně třísekundová (ne delší než 5 s) prodloužená samohláska [i] v pohodlné výšce a hlasitosti. Vykonáno na jeden nádech.
TSK5	Prodloužená fonace	Přibližně třísekundová (ne delší než 5 s) prodloužená samohláska [u] v pohodlné výšce a hlasitosti. Vykonáno na jeden nádech.
TSK6	Prodloužená fonace	Samohláska [a] v pohodlné výšce a hlasitosti udržena co nejdéle a co nejmonotónněji. Vykonáno na jeden nádech.
TSK7	Diadochokinetická úloha	Rychlé a rovnoměrné opakování slabik [pa]-[ta]-[ka] co nejdéle (alespoň 10 opakování) a co nejmonotónněji. Vykonáno na jeden nádech.
TSK8-17	Opakování víceslabičných slov	Opakování deseti víceslabičných slov podle lékaře. Šest z nich by mělo vždy mít alespoň tři slabiky a pro první dvě slova by měli mít strukturu KVKV (K – konsonanta, V – Vokál).

### 4.3 Popis databáze

Celý korpus obsahuje nahrávky rozdělené do tří skupin podle národnosti a příslušného jazyka: *čeština* (cs), *americká angličtina* (en-US) a *maďarština* (hu). V databázi jsou pro každou jazykovou skupinu obsaženy nahrávky pacientů s PN, ale také zdravých lidí bez známky HD nesoucí označí: *zdravá kontrola* (ZK). Počty zastoupených mužů a žen v kombinaci s diagnózou pro každou skupinu je vypsán v Tabulce 4.2.

Tab. 4.2: Zastoupení pohlaví v každé skupině v kombinaci s diagnózou

	cs			en-US			hu			dohromady
	muži	ženy	celkem	muži	ženy	celkem	muži	ženy	celkem	
PN	3	4	7	6	3	9	8	2	10	26
ZK	10	22	32	3	9	12	2	6	8	52
celkem	13	26	39	9	12	21	10	8	18	78

V databázi je také vždy uveden věk nahrávaného člověka. Jedná se o znak, který může ovlivnit výsledky výstupních hodnot pozdějších analýz a proto je třeba jej brát v potaz. V tabulkách 4.3, 4.4 a 4.5 jsou vypočítány následující statistické veličiny, které zde popisují věkovou strukturu jednotlivých jazykových skupin: *rozsah souboru* ( $n$ ), *aritmetický průměr* ( $\bar{x}$ ), *výběrová směrodatná odchylka* ( $s$ ), *nejmenší hodnota* ( $X_{min}$ ), *první kvartil* ( $Q_1$ ), *medián* ( $Me$ ), *třetí kvartil* ( $Q_3$ ), *největší hodnota* ( $X_{max}$ ) a *variální rozpětí* ( $R$ ).

Tab. 4.3: Počet účastníků a jejich věkové zastoupení pro češtinu

cs							
	PN			ZK			dohromady
	muži	ženy	celkem	muži	ženy	celkem	
$n$	3	4	7	10	22	32	39
$\bar{x}$	70,33	66,50	68,14	65,80	67,00	66,625	66,90
$s$	8,08	5,07	6,23	5,65	6,51	6,18922	6,06
$X_{min}$	63	60	60	60	59	59	59
$Q_1$	66,00	63,75	64,00	61,25	63,00	62,00	62,50
$Me$	69,00	67,50	69,00	64,50	66,00	66,00	66,00
$Q_3$	74,00	70,25	70,50	70,50	69,00	69,25	70,00
$X_{max}$	79	71	79	75	85	85	85
$R$	16	11	19	15	26	26	26

Tab. 4.4: Počet účastníků a jejich věkové zastoupení pro americkou angličtinu

en-US							
	PN			ZK			dohromady
	muži	ženy	celkem	muži	ženy	celkem	
$n$	6	3	9	3	9	12	21
$\bar{x}$	68,83	68,00	68,56	67,67	74,11	72,50	70,81
$s$	4,96	2,00	4,07	15,70	4,81	8,37	6,84
$X_{min}$	59	66	59	50	68	50	50
$Q_1$	69,25	67,00	68,00	61,50	71,00	70,25	68,00
$Me$	70,50	68,00	70,00	73,00	74,00	73,50	71,00
$Q_3$	71,75	69,00	71,00	76,50	78,00	78,00	74,00
$X_{max}$	72	70	72	80	82	82	82
$R$	13	4	13	30	14	32	32

Tab. 4.5: Počet účastníků a jejich věkové zastoupení pro maďarštinu

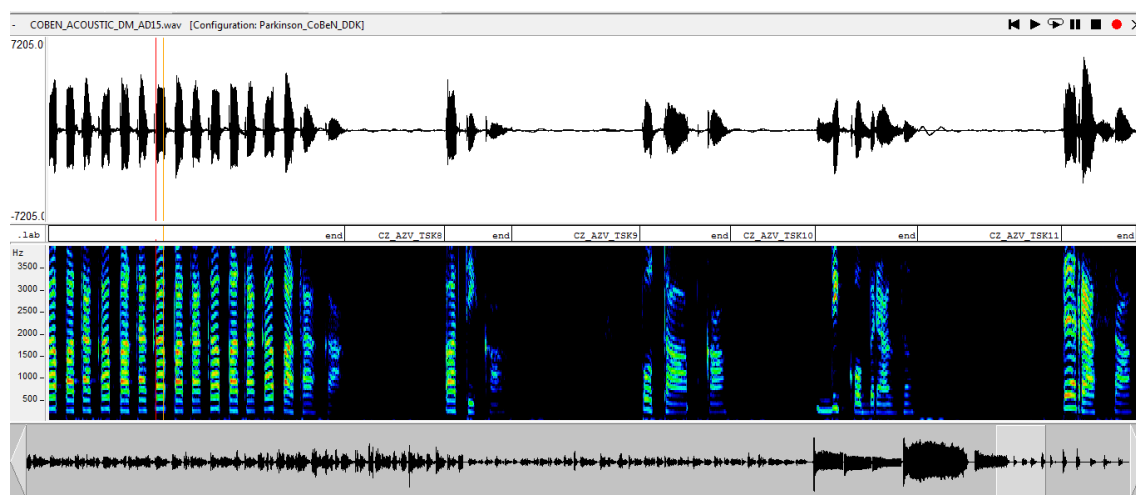
hu							
	PN			ZK			dohromady
	muži	ženy	celkem	muži	ženy	celkem	
$n$	8	2	10	2	6	8	18
$\bar{x}$	68,50	56,50	66,10	62,50	66,83	66,83	65,94
$s$	6,55	16,26	9,40	3,54	5,74	5,74	7,26
$X_{min}$	62	45	45	60	58	58	45
$Q_1$	63,00	50,75	63,00	61,25	64,00	64,00	63,00
$Me$	67,00	56,50	67,00	62,50	67,50	64,00	65,94
$Q_3$	71,00	62,25	70,25	63,75	71,75	67,50	71,00
$X_{max}$	81	68	81	65	72	72	81
$R$	19	23	36	5	14	14	36



Z výsledků lze vyčíst, že největší zastoupení v databázi mají česky hovořící lidé, na druhém místě je americká angličtina a nejméně nahrávek je s maďarštinou. Z hlediska věkového průměru lze tvrdit, že nejstarší skupinu tvoří americká angličtina, poté čeština a nejmladší jsou maďaři.

## 4.4 Segmentace nahrávek

Nahrávky bylo nejdřív potřeba nasegmentovat na jednotlivé řečové úlohy, aby se mohli samostatně analyzovat (viz Obrázek 4.1). K tomu posloužil zvukový editor WaveSurfer. Každá nahrávka byla poslechem zkontrolována a poté označena na časové ose tam, kde každá úloha začínala a končila. Podle označení se také následující střížené úseky všech nahrávek automaticky rozdělily do skupin konkrétních řečových úloh. Diadochokinetickou úlohu bylo ještě třeba rozdělit na jednotlivé slabiky.



Obr. 4.1: Segmentace nahrávek na jednotlivé řečové úlohy

## 4.5 Akustické parametry

Byly vypočítány vybrané akustické parametry v různých kombinacích s řečovými úlohami pro každého účastníka. Seznam všech parametrů spolu s poruchami, které prošetřují je uveden v Tabulce 4.6. Parametry jsou rozděleny do několika oblastí. Výsledky byly uloženy jako 3 matice parametrů rozdělených podle národnosti do tabulkového procesoru Excel, kde každý řádek obsahuje hodnoty jednoho parametru a sloupce zastupují testované jedince hovořící stejným jazykem.

S informacemi o věku a pohlaví jsou s maticí parametrů propojené skrz *identifikátor uživatele* (UID).

Tab. 4.6: Seznam akustických parametrů

Oblast projevu HD a konkrétní porucha	Řečová úloha	Akustický parametr	Definice parametru
<b>Fonace</b>			
Nedostatečná respirace	(TSK6)	MPT	Maximální doba fonace, aerodynamická účinnost vokálního traktu měřená jako maximální doba trvání prodloužené samohlásky.
Nepravidelné kolísání výšky tónu	(TSK3–6)	relF0SD	Směrodatná odchylka základní frekvence ve vztahu k její střední hodnotě, změna frekvence hlasivek.
Mikroperturbace ve frekvenci	(TSK3–6)	Jitter (PPQ)	Míra variace rozsahu hlasu. Jitter je definován jako variabilita základního tónu řeči (F0) mezi sousedními vokálními cykly.
Mikroperturbace v amplitudě	(TSK3–6)	Shimmer (APQ)	Hrubost řeči. Shimmer je definován jako sekvence maximálního rozsahu amplitudy signálu v každém vokálním cyklu.
Zvýšený šum	(TSK3–6)	HNR	Množství šumu v řečovém signálu převážně v důsledku neúplného uzavření hlasivek. Jde o poměr mezi tonální složkou řeči a amplitudou šumu.
Aperiodicita	(TSK3–6)	DUV	Podíl neznělých segmentů.
Chvění brady a jazyka	(TSK3–6)	relF1SD, relF2SD	Směrodatná odchylka prvního (F1) a druhého (F2) formantu vzhledem k jeho střední hodnotě. Formanty jsou spojeny s rezonancemi ústní, nosní a hrdelní dutiny a jsou modifikovány polohou jazyka a čelisti.
<b>Artikulace</b>			
Snížená hybnost jazyka (nepřesnost samohlásek)	(TSK1–5)	VAI	Index artikulace samohlásek definovaný jako $VAI = (F1a + F2i) / (F1i + F1u + F2a + F2u)$ .
Rigidita jazyka a čelisti	(TSK1–2)	relF1SD, relF2SD	Směrodatná odchylka prvního (F1) a druhého (F2) formantu vzhledem k jeho střední hodnotě.
Nepřesnost souhlásek	(TSK7–8)	VOT	Doba nástupu hlasu definovaná jako délka celé souhlásky od počátečního impulsu po začátek samohlásky.
Snížené tempo při diadochokinezi	(TSK7)	DDK rate	Diadochokinetická rychlost představující počet slabikálních vokalizací za sekundu. Uvažuje se prvních 30 slabik.
Nepravidelné tempo diadochokineze	(TSK7)	DDK reg	Diadochokinetická pravidelnost definovaná jako směrodatná odchylka délky slabik. Uvažuje se prvních 30 slabik.
Nestabilita diadochokinetického tempa	(TSK7)	DDK instab	Diadochokinetická nestabilita definovaná jako podíl směrodatné odchylky trvání 4. - 10. DDK cyklu a průměrnou dobou trvání prvních tří cyklů.
<b>Prozodie</b>			
Snížená variabilita hlasitosti	(TSK1–2)	relSEOSD	Změna hlasitosti řeči definovaná jako směrodatná odchylka krátkodobé energie vzhledem k její střední hodnotě po odstranění pauz přesahujícího 50 ms.

Tab. 4.6: Pokračování z předchozí strany

Oblast projevu HD a konkrétní porucha	Řečová úloha	Akustický parametr	Definice parametru
Monotónnost	(TSK1-2)	relF0SD	Variace výšky tónu definovaná jako směrodatná odchylka základního tónu F0 vzhledem k jeho střední hodnotě.
Nevhodné pauzy	(TSK2)	SPIR	Počet pauz vzhledem k celkové době promluvy po odstranění pauz trvající méně než 50 ms.
Větší podíl ticha	(TSK2)	PPR	Procentuální poměr pauz definovaný jako poměr celkové doby pauz (delších než 50 ms) a celkové doby promluvy.
Delší doba pauz	(TSK2)	DurMED	Střední doba pauz delších než 50 ms.
Vyšší variabilita trvání pauz	(TSK2)	DurMAD	Střední absolutní odchylka trvání pauz (pauzy delší než 50 ms).
Nepřirozená rychlost řeči	(TSK2)	AR	Počet hlásek řečníka za sekundu po odstranění pauzy delší než 50 ms.
Nížší tempo	(TSK2)	NST	Celková délka čistě řečového signálu.
Nížší tempo	(TSK2)	TST	Celková délka pauz delších než 50 ms.

## 4.6 Statistická analýza

Veškerá statistická analýza proběhla za využití programového prostředí MATLAB verze 9.6. Z dat nejprve byly odstraněny nečíselné prvky, které vznikly z příčiny neprovedené nebo nesprávně provedené řečové úlohy (např. účastník špatně viděl na text v úloze, při které je potřeba číst). Také byl odstraněn parametr DUV, který obsahoval příliš mnoho nulových hodnot, aby mohl být předmětem statistických testů. Protože parametry byly závislé na věku a pohlaví, musel se tento vliv z našich dat odstranit, aby byl parametr závislý pouze na míře postižení HD. K odstranění kovariát došlo za pomoci lineární regrese. Následně se otastovalo normální rozdělení parametrů Kolmogorovým-Smirnovovým testem. Protože některé z nich normální rozdělení nevykazovaly, byl pro následnou analýzu diskriminační síly jednotlivých parametrů využit Mann-Whitneyho U test. Pro zjištění závislosti parametrů na národnosti byl zvolen Kruskal-Wallisův test a korekce p-hodnot proběhla využitím adjusted FDR. Z hodnot parametrů byl nakonec vypočítán aritmetický průměr, medián a směrodatná odchylka pro PN a ZK všech jazyků.

## 4.7 Vícerozměrná analýza

Vícerozměrná analýza proběhla využitím programovacího jazyka Python verze 3.7.7 v prostředí Spyder s verzí 4.1.3. Veškeré vypočítané parametry v kombinaci s řečovou úlohou byly po odstranění kovariátů použity jako vstupní hodnoty pro natrénování modelů strojového učení, které jsou do určité míry schopny identifikovat HD a s tím spojenou PN. Analýza proběhla pro každý jazyk zvlášť a také pro všechny dohromady. Jako klasifikátor byla vybrána metoda náhodných lesů a metoda *eXtreme Gradient Boosting* (XGBoost). Druhá metoda vykazovala lepší výsledky během prvotních testů a proto se další práce stavěla pouze na tomto klasifikátoru.

### 4.7.1 Metoda eXtreme Gradient Boosting

Na rozdíl od přístupu bagging, do kterého spadá metoda náhodných lesů a rozhodovací stromy jsou zde tvořeny z náhodně vybraných parametrů, využívá XGBoost přístup zvaný boosting. Zde se na začátku vytvoří pouze jeden rozhodovací strom a každý nový se tvoří na základě chybovosti stromu předchozího. Takový klasifikátor má výhodu ve větší kontrole modelu volbou vhodných hyperparametrů. Část hyperparametrů modelu byla nastavena následovně:

<b>objective:</b>	binary: logistic,	<b>importance_type:</b>	gain,
<b>eval_metric:</b>	error,	<b>max_delta_step:</b>	0,
<b>seed:</b>	42,	<b>num_parallel_tree:</b>	1,
<b>n_estimators:</b>	500,	<b>reg_alpha:</b>	0,
<b>base_score:</b>	0.5,	<b>reg_lambda:</b>	1,
<b>colsample_bynode:</b>	1,	<b>random_state:</b>	42.

V Tabulce 4.7 jsou pak další důležité hyperparametry, které byly vybírány tak, aby jejich kombinace byla optimální pro dosažení nejlepšího možného výkonu.

Tab. 4.7: Mřížka pro optimalizaci hyperparametrů

<b>Learning_rate</b>	0,001	0,001	0,1	0,2	0,3	0,4		
<b>gamma</b>	0	0,1	0,15	0,25	0,5			
<b>max_depth</b>	6	8	10	12	15	20	30	40
<b>subsample</b>	0,5	0,6	0,7	0,8	0,9	1		
<b>colsample_byLevel</b>	0,4	0,5	0,6	0,7	0,8	0,9	1	
<b>colsample_byTree</b>	0,4	0,5	0,6	0,7	0,8	0,9	1	
<b>min_child_weight</b>	0,5	1	3	5	7	10		
<b>scale_pos_weight</b>	1	3	5	6	7	9		

Pro nalezení optimálních hyperparametrů byla zvolena metoda random search. Počet iterací byl nastaven na 500 a optimalizace probíhala na základě *balanced accuracy* ( $ACC_{\text{bal}}$ ) s 10-fold křížovou krosvalidací:

$$ACC_{\text{bal}} = \left( \frac{SP}{SP + FN} + \frac{SN}{SN + FP} \right) \cdot \frac{1}{2}, \quad (4.1)$$

kde  $SP$  značí počet účastníků vyhodnocených jako skutečně pozitivní,  $SN$  skutečně negativní,  $FP$  falešně pozitivní a  $FN$  falešně negativní. Na Obrázku 4.2 je vyzobrazena matice záměn pro tento model.

**P ř e d p o v ě ě**

		<b>0</b>	<b>1</b>	
S k u t e č n o s t	<b>0</b>	<b>SN</b>	<b>FP</b>	0 = ZK 1 = PN
	<b>1</b>	<b>FN</b>	<b>SP</b>	

Obr. 4.2: Matice záměn modelu XGBoost

Následně byla vybrána ta kombinace hodnot hyperparametrů, pro kterou vycházela  $ACC_{\text{bal}}$  nejvyšší, a s tou byl model znovu natrénován. Pro vyhodnocení modelu byla použita 10-fold křížová krosvalidace s 20 opakováním a jako metriky zvoleny *přesnost* (ACC), *senzitivita* (SEN) a *specifická* (SPE):

$$ACC = \frac{SP + SN}{SP + SN + FP + FN} \cdot 100, \quad (4.2)$$

$$SEN = \frac{SP}{SP + FN} \cdot 100, \quad (4.3)$$

$$SPE = \frac{SN}{SN + FP} \cdot 100. \quad (4.4)$$

Ve výsledku tedy vyšlo 200 hodnot ACC, 200 hodnot SEN a 200 hodnot SPE, pro které byl vypočítán aritmetický průměr a směrodatná odchylka.

## 5 Výsledky

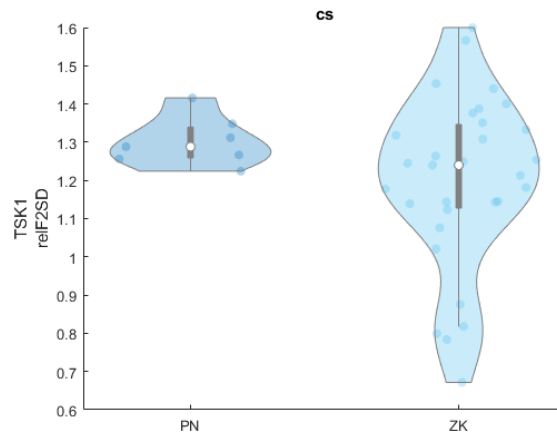
### 5.1 Jednorozměrná analýza

V Tabulce 5.1 jsou vypsány výsledky statistických testů pro různé kombinace

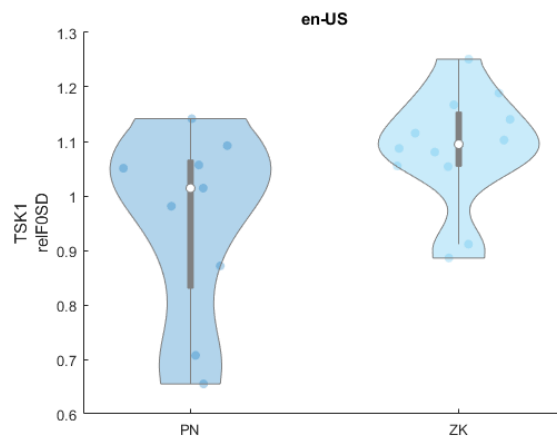
Tab. 5.1: Výsledky statistických testů akustických parametrů

Úloha	Parametr	Kolmogorovův-Smirnovův test				Mann-Whitneyho U test				Kruskal-Wallisův test p - hodnoty po korekci
		1 = nevykazuje normální rozdělení				p - hodnoty				
		cs	en-US	hu	vj	cs	en-US	hu	vj	
TSK1	relF1SD	0	0	0	0	0,956	0,500	0,829	0,521	<b>0,001</b>
TSK1	relF2SD	0	0	0	0	<b>0,005</b>	0,500	0,408	0,886	<b>0,000</b>
TSK1	relF0SD	0	0	0	0	0,390	<b>0,004</b>	0,203	0,987	0,054
TSK1	relSEOSD	0	0	0	1	0,840	0,303	0,573	0,268	<b>0,003</b>
TSK2	AR	0	0	0	0	0,103	0,546	0,541	0,309	<b>0,000</b>
TSK2	DurMAD	1	1	1	1	0,159	1,000	0,606	0,223	0,171
TSK2	DurMED	1	0	1	1	0,153	0,696	0,139	0,675	0,189
TSK2	NST	0	0	0	1	0,498	0,145	0,606	<b>0,017</b>	<b>0,000</b>
TSK2	PPR	0	0	0	0	0,370	0,189	0,423	0,238	<b>0,000</b>
TSK2	relF1SD	0	0	0	0	0,701	0,696	0,815	0,284	<b>0,000</b>
TSK2	relF2SD	0	0	0	0	<b>0,046</b>	0,804	0,606	0,411	<b>0,000</b>
TSK2	relF0SD	0	0	0	0	0,370	<b>0,006</b>	0,963	0,364	<b>0,030</b>
TSK2	relSEOSD	1	0	0	1	0,985	0,696	0,815	0,961	0,178
TSK2	SPIR	0	0	0	0	0,621	<b>0,003</b>	0,815	<b>0,005</b>	<b>0,000</b>
TSK2	TPT	0	0	0	0	0,111	0,456	0,236	0,375	<b>0,003</b>
TSK2	TST	0	0	0	1	0,129	0,414	0,963	0,097	<b>0,000</b>
TSK3	shimmer (APQ)	1	0	0	1	0,386	0,214	0,762	0,694	<b>0,000</b>
TSK3	mean HNR	0	0	0	0	0,851	<b>0,004</b>	0,696	0,910	<b>0,001</b>
TSK3	median HNR	0	0	0	0	0,910	<b>0,006</b>	0,696	0,936	<b>0,001</b>
TSK3	jitter (PPQ)	1	1	0	1	0,522	0,749	0,460	0,979	0,645
TSK3	relF1SD	0	0	0	0	0,175	0,414	0,237	0,144	0,594
TSK3	relF2SD	1	0	0	1	0,065	0,500	0,360	<b>0,036</b>	0,441
TSK3	relF0SD	1	1	1	1	0,970	0,859	0,274	0,234	0,554
TSK4	shimmer (APQ)	1	0	0	1	0,522	0,271	0,965	0,448	<b>0,000</b>
TSK4	mean HNR	0	0	0	0	0,347	0,110	0,965	0,859	<b>0,000</b>
TSK4	median HNR	0	0	0	0	0,214	0,082	0,965	0,842	<b>0,000</b>
TSK4	jitter (PPQ)	1	1	0	1	0,735	0,972	0,696	0,759	<b>0,013</b>
TSK4	relF1SD	0	0	0	1	0,407	0,594	0,633	0,970	<b>0,000</b>
TSK4	relF2SD	0	0	0	1	<b>0,032</b>	0,303	0,696	0,404	<b>0,014</b>
TSK4	relF0SD	0	1	0	1	0,451	0,166	0,237	0,060	0,178
TSK5	shimmer (APQ)	1	0	0	1	0,880	0,303	0,274	0,702	0,062
TSK5	mean HNR	0	0	0	0	0,572	0,189	0,237	0,800	<b>0,049</b>
TSK5	median HNR	0	0	0	0	0,651	0,189	0,237	0,817	0,052
TSK5	jitter (PPQ)	1	0	0	1	0,522	0,214	0,146	0,751	0,288
TSK5	relF1SD	1	1	0	1	0,243	0,594	0,360	0,825	<b>0,004</b>
TSK5	relF2SD	1	0	0	1	0,707	0,271	0,203	0,663	<b>0,001</b>
TSK5	relF0SD	1	0	0	1	0,851	0,241	0,237	0,052	0,645
TSK6	shimmer (APQ)	0	0	0	1	0,370	<b>0,030</b>	0,696	0,675	<b>0,000</b>
TSK6	jitter (PPQ)	0	0	0	1	0,674	0,051	0,573	0,600	0,463
TSK6	relF1SD	1	0	0	1	0,956	0,594	0,897	0,488	0,448
TSK6	relF2SD	0	0	0	1	0,194	0,915	0,146	0,074	0,448
TSK6	relF0SD	0	0	0	1	1,000	0,972	0,315	0,763	0,645
TSK6	MPT	0	0	0	0	0,194	0,456	0,696	0,160	0,759
TSK7	PR	0	0	0	0	0,763	0,696	0,146	0,448	<b>0,000</b>
TSK7	COV	0	0	0	1	0,628	0,859	0,696	0,396	0,052
TSK7	RI	1	0	0	1	0,328	1,000	0,237	0,206	<b>0,018</b>
TSK7	PA	0	0	0	0	0,398	0,374	0,963	0,110	0,257
TSK7	RA	0	0	0	1	0,763	0,271	0,274	0,176	0,849
GLOBAL	VAI	0	0	0	0	0,175	<b>0,043</b>	0,315	0,324	<b>0,000</b>

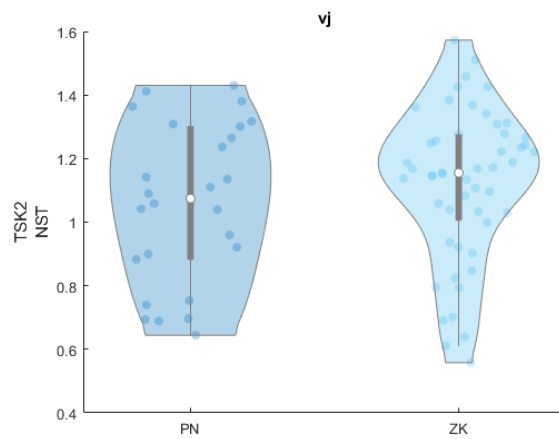
akustických parametrů s řečovými úlohami. P-hodnoty Mann-Whitneyho U testu jsou zvýrazněny zeleně, pokud jsou jejich hodnoty nižší než hladina významnosti stanovena na  $\alpha = 0.05$  – v tomto případě parametr vykazuje signifikantní rozdíl mezi hodnotami PN a ZK. Na obrázcích 5.1 až 5.13 jsou vykresleny houslové grafy těchto významných kombinací parametrů, úloh a jazyka v závislosti na diagnóze. Byla provedena i korekce p-hodnot pro vícenásobné porovnávání, ale protože se následně nedostaly pod hladinu významnosti, byly ponechány bez korekce. U Kruskal-Wallisova testu jsou rovněž zvýrazněny p-hodnoty tentokrát již s korekcí. Tůčné vyznačení vyjadřuje významnou závislost daného parametru na národnosti.



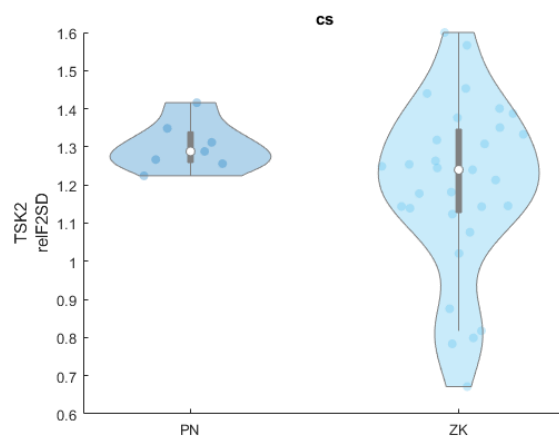
Obr. 5.1: Rozdíl parametru relF2SD z úlohy TSK1 mezi PN a ZK pro cs



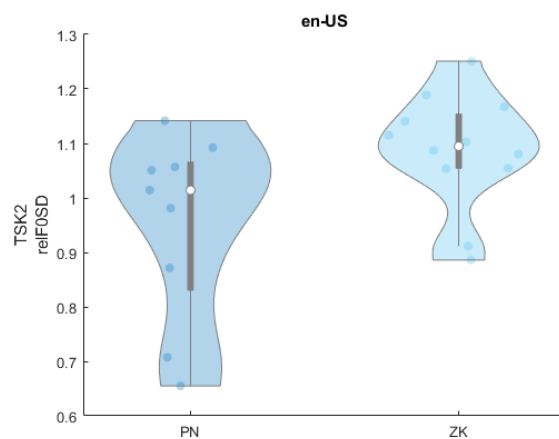
Obr. 5.2: Rozdíl parametru relF0SD z úlohy TSK1 mezi PN a ZK pro en-US



Obr. 5.3: Rozdíl parametru NST z úlohy TSK1 mezi PN a ZK pro *vj*

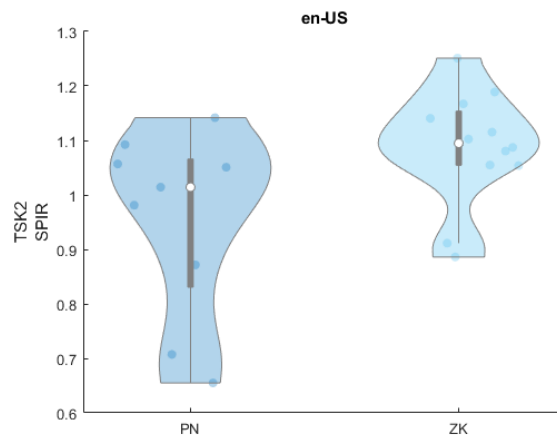


Obr. 5.4: Rozdíl parametru relF2SD z úlohy TSK2 mezi PN a ZK pro *cs*

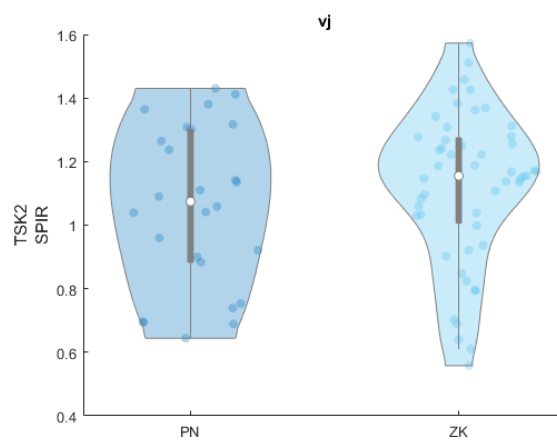


Obr. 5.5: Rozdíl parametru relF0SD z úlohy TSK2 mezi PN a ZK pro *en-US*

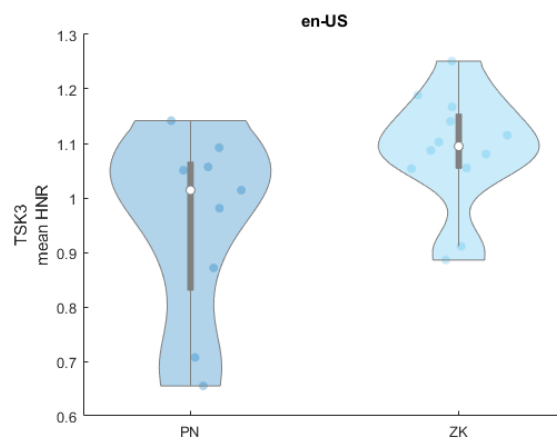




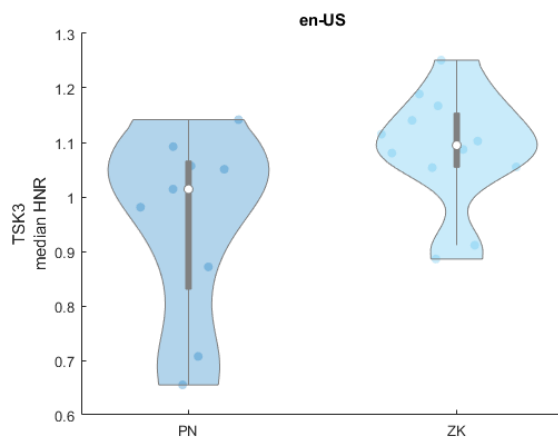
Obr. 5.6: Rozdíl parametru SPIR z úlohy TSK2 mezi PN a ZK pro en-US



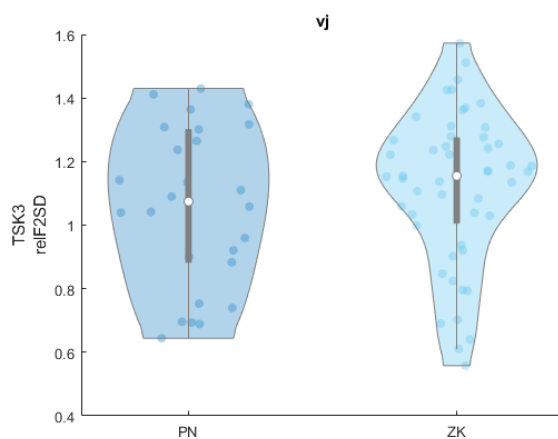
Obr. 5.7: Rozdíl parametru SPIR z úlohy TSK2 mezi PN a ZK pro vj



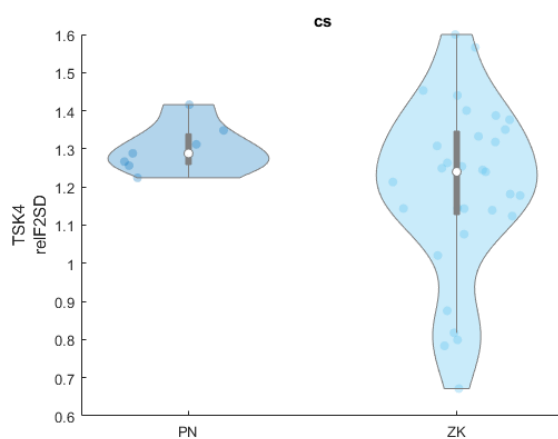
Obr. 5.8: Rozdíl parametru mean HNR z úlohy TSK3 mezi PN a ZK pro en-US



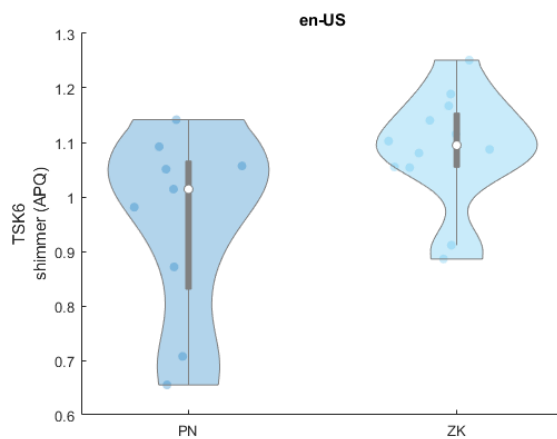
Obr. 5.9: Rozdíl parametru median HNR z úlohy TSK3 mezi PN a ZK pro en-US



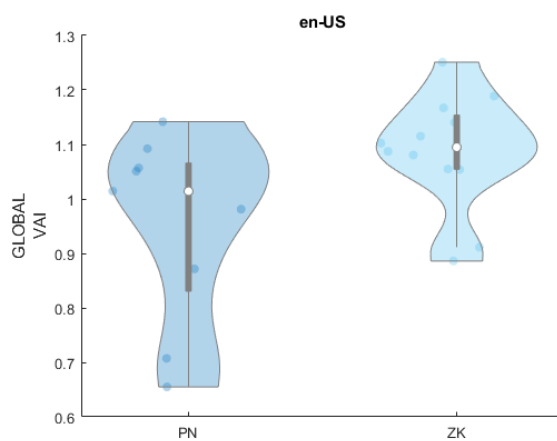
Obr. 5.10: Rozdíl parametru relF2SD z úlohy TSK3 mezi PN a ZK pro vj



Obr. 5.11: Rozdíl parametru median relF2SD z úlohy TSK4 mezi PN a ZK pro cs



Obr. 5.12: Rozdíl parametru shimmer z úlohy TSK6 mezi PN a ZK pro en-US



Obr. 5.13: Rozdíl parametru VAI ze všech úloh mezi PN a ZK pro en-US

V tabulkách 5.2 až 5.4 jsou obsaženy výsledky deskriptivní statistiky, kde rovněž platí, že zelené zvýraznění vyjadřuje signifikantní rozdíl mezi PN a ZK a tučně vyznačené řádky zde znamenají signifikantní závislost parametru na národnosti.

Tab. 5.2: Výsledky deskriptivní statistiky: aritmetický průměr

Úloha	Parametr	Aritmetický průměr							
		cs		en-US		hu		vj	
		PN	ZK	PN	ZK	PN	ZK	PN	ZK
TSK1	relF1SD	0,550	0,555	0,543	0,512	0,390	0,377	0,484	0,519
TSK1	relF2SD	0,323	0,285	0,261	0,269	0,247	0,238	0,274	0,273
TSK1	relF0SD	0,273	0,242	0,235	0,329	0,256	0,204	0,254	0,256
TSK1	relSEOSD	1,608	1,661	1,151	1,341	0,978	0,980	1,226	1,473
TSK2	AR	18,073	22,408	31,335	29,326	35,652	33,668	27,988	26,307
TSK2	DurMAD	0,047	0,044	0,047	0,038	0,067	0,057	0,055	0,044
TSK2	DurMED	0,121	0,115	0,113	0,103	0,131	0,123	0,123	0,113
TSK2	NST	38,081	35,457	38,453	35,936	72,474	68,564	49,058	41,400
TSK2	PPR	0,488	0,460	0,402	0,425	0,318	0,354	0,406	0,431
TSK2	relF1SD	0,482	0,503	0,499	0,507	0,342	0,332	0,437	0,478
TSK2	relF2SD	0,322	0,300	0,254	0,254	0,240	0,240	0,270	0,279
TSK2	relF0SD	0,253	0,234	0,232	0,341	0,221	0,209	0,238	0,253
TSK2	relSEOSD	1,106	1,258	1,068	1,046	0,953	0,987	1,058	1,157
TSK2	SPIR	0,037	0,038	0,033	0,039	0,023	0,028	0,032	0,036
TSK2	TPT	38,614	33,076	28,238	28,485	35,623	39,299	34,467	32,654
TSK2	TST	78,386	70,629	69,051	67,223	111,724	111,140	86,165	76,490
TSK3	shimmer (APQ)	3,272	4,381	6,908	13,986	9,069	8,283	6,670	7,299
TSK3	mean HNR	20,397	19,987	19,195	14,159	14,329	16,115	17,859	17,900
TSK3	median HNR	20,503	20,259	19,158	14,168	14,504	16,177	17,940	18,078
TSK3	jitter (PPQ)	0,375	0,571	0,281	0,697	0,851	0,536	0,542	0,587
TSK3	relF1SD	0,076	0,107	0,058	0,106	0,120	0,081	0,083	0,105
TSK3	relF2SD	0,015	0,066	0,049	0,071	0,100	0,103	0,056	0,075
TSK3	relF0SD	0,060	0,068	0,019	0,048	0,074	0,030	0,052	0,057
TSK4	shimmer (APQ)	1,775	2,888	7,499	13,933	6,736	6,048	5,500	6,066
TSK4	mean HNR	27,390	25,858	20,830	17,742	20,963	21,419	23,103	23,019
TSK4	median HNR	27,905	26,004	20,872	17,750	21,099	21,607	23,277	23,156
TSK4	jitter (PPQ)	0,464	0,209	0,406	1,036	0,496	0,340	0,425	0,440
TSK4	relF1SD	0,026	0,034	0,045	0,051	0,119	0,108	0,063	0,052
TSK4	relF2SD	0,015	0,055	0,072	0,045	0,127	0,121	0,076	0,064
TSK4	relF0SD	0,012	0,011	0,019	0,016	0,013	0,011	0,014	0,012
TSK5	shimmer (APQ)	10,311	7,417	8,104	12,145	6,618	4,919	7,827	8,291
TSK5	mean HNR	25,461	24,396	21,833	19,955	19,819	24,046	22,651	22,982
TSK5	median HNR	25,290	24,347	21,820	19,932	19,941	24,269	22,661	22,975
TSK5	jitter (PPQ)	0,458	0,301	0,254	0,310	0,339	0,255	0,324	0,305
TSK5	relF1SD	0,029	0,082	0,119	0,070	0,159	0,120	0,104	0,088
TSK5	relF2SD	0,061	0,114	0,079	0,148	0,389	0,270	0,193	0,147
TSK5	relF0SD	0,054	0,018	0,010	0,015	0,012	0,010	0,021	0,017
TSK6	shimmer (APQ)	3,158	4,350	8,745	15,888	10,287	9,737	7,638	7,940
TSK6	jitter (PPQ)	0,594	0,657	0,532	1,181	0,765	0,606	0,655	0,762
TSK6	relF1SD	0,113	0,149	0,177	0,180	0,159	0,152	0,149	0,158
TSK6	relF2SD	0,045	0,125	0,102	0,078	0,113	0,197	0,084	0,129
TSK6	relF0SD	0,143	0,125	0,107	0,114	0,127	0,070	0,126	0,113
TSK6	MPT	15,573	18,475	14,510	17,470	16,578	17,765	15,511	18,175
TSK7	PR	5,464	5,314	4,955	4,863	6,371	5,950	5,571	5,341
TSK7	COV	4,344	4,256	7,178	6,884	3,721	3,904	4,762	4,999
TSK7	RI	0,030	0,028	0,045	0,043	0,030	0,029	0,034	0,032
TSK7	PA	-2,043	0,172	-5,454	-1,015	-0,779	-0,705	-2,913	-0,220
TSK7	RA	0,000	0,000	-0,006	0,003	0,000	-0,002	-0,002	0,001
GLOBAL	VAI	1,302	1,197	0,952	1,087	0,891	0,991	1,058	1,120

Tab. 5.3: Výsledky deskriptivní statistiky: medián

Úloha	Parametr	Medián							
		cs		en-US		hu		vj	
		PN	ZK	PN	ZK	PN	ZK	PN	ZK
TSK1	relF1SD	0,550	0,546	0,551	0,502	0,359	0,368	0,509	0,506
TSK1	relF2SD	0,332	0,286	0,254	0,268	0,247	0,229	0,275	0,276
TSK1	relF0SD	0,255	0,224	0,231	0,334	0,215	0,197	0,253	0,243
TSK1	relSEOSD	1,535	1,468	1,122	1,240	0,988	0,935	1,082	1,150
TSK2	AR	19,191	22,153	32,139	28,265	36,534	34,014	28,515	24,667
TSK2	DurMAD	0,049	0,039	0,042	0,040	0,043	0,053	0,044	0,042
TSK2	DurMED	0,118	0,111	0,095	0,104	0,105	0,117	0,103	0,105
TSK2	NST	37,763	36,075	39,554	34,352	75,053	71,264	42,208	36,385
TSK2	PPR	0,463	0,457	0,376	0,424	0,307	0,326	0,391	0,424
TSK2	relF1SD	0,470	0,499	0,450	0,495	0,291	0,344	0,454	0,469
TSK2	relF2SD	0,316	0,294	0,256	0,252	0,254	0,227	0,269	0,275
TSK2	relF0SD	0,255	0,234	0,206	0,336	0,223	0,221	0,219	0,249
TSK2	relSEOSD	1,066	1,057	0,991	1,041	0,979	0,955	1,028	1,032
TSK2	SPIR	0,037	0,038	0,032	0,040	0,024	0,024	0,032	0,038
TSK2	TPT	35,735	30,710	25,795	28,460	34,424	38,358	34,779	30,562
TSK2	TST	77,122	70,104	68,502	65,585	114,913	112,163	81,204	72,850
TSK3	shimmer (APQ)	2,581	3,319	7,367	11,087	10,537	6,339	5,681	3,991
TSK3	mean HNR	20,111	20,641	19,234	13,657	12,649	18,324	18,276	18,963
TSK3	median HNR	20,157	21,078	18,890	13,738	12,734	18,136	18,736	18,903
TSK3	jitter (PPQ)	0,286	0,333	0,420	0,360	0,454	0,325	0,305	0,338
TSK3	relF1SD	0,061	0,105	0,066	0,085	0,090	0,067	0,059	0,098
TSK3	relF2SD	0,014	0,053	0,052	0,053	0,046	0,074	0,026	0,054
TSK3	relF0SD	0,022	0,020	0,032	0,000	0,021	0,017	0,017	0,012
TSK4	shimmer (APQ)	1,527	1,684	7,685	10,537	6,323	6,289	4,275	3,020
TSK4	mean HNR	27,141	26,559	21,279	18,931	20,924	21,821	23,976	24,328
TSK4	median HNR	27,786	26,761	21,344	18,629	21,250	21,846	24,725	24,149
TSK4	jitter (PPQ)	0,168	0,206	0,652	0,178	0,437	0,359	0,212	0,233
TSK4	relF1SD	0,027	0,027	0,036	0,044	0,074	0,106	0,044	0,039
TSK4	relF2SD	0,010	0,033	0,052	0,001	0,106	0,086	0,040	0,033
TSK4	relF0SD	0,009	0,010	0,010	0,017	0,013	0,010	0,010	0,012
TSK5	shimmer (APQ)	3,231	4,416	6,711	10,473	5,465	3,913	5,548	5,813
TSK5	mean HNR	28,739	26,354	22,268	19,648	20,354	21,805	22,979	22,307
TSK5	median HNR	28,457	26,073	22,195	19,793	20,637	22,271	23,107	23,030
TSK5	jitter (PPQ)	0,258	0,188	0,207	0,261	0,318	0,279	0,200	0,193
TSK5	relF1SD	0,036	0,055	0,056	0,087	0,138	0,095	0,074	0,067
TSK5	relF2SD	0,072	0,071	0,098	0,120	0,378	0,322	0,139	0,092
TSK5	relF0SD	0,014	0,013	0,008	0,012	0,011	0,008	0,010	0,011
TSK6	shimmer (APQ)	2,931	3,919	6,056	13,874	10,416	8,704	6,149	5,151
TSK6	jitter (PPQ)	0,278	0,555	0,453	1,067	0,479	0,476	0,476	0,577
TSK6	relF1SD	0,095	0,094	0,122	0,182	0,122	0,118	0,101	0,106
TSK6	relF2SD	0,057	0,093	0,099	0,074	0,070	0,169	0,069	0,105
TSK6	relF0SD	0,087	0,104	0,070	0,084	0,103	0,052	0,081	0,085
TSK6	MPT	13,946	17,961	13,487	18,198	17,198	15,301	15,071	17,108
TSK7	PR	5,440	5,524	4,785	5,218	6,585	5,888	5,452	5,482
TSK7	COV	3,712	3,625	5,509	5,376	3,577	3,258	3,272	4,061
TSK7	RI	0,027	0,022	0,036	0,036	0,030	0,026	0,029	0,025
TSK7	PA	-2,080	0,238	-4,140	-1,955	-0,464	-1,544	-2,304	-0,308
TSK7	RA	0,002	0,000	-0,006	0,004	0,000	-0,002	0,000	0,001
GLOBAL	VAI	1,288	1,240	1,014	1,095	0,879	0,998	1,074	1,155

Tab. 5.4: Výsledky deskriptivní statistiky: směrodatná odchylka

Úloha	Parametr	Směrodatná odchylka							
		cs		en-us		hu		vj	
		PN	ZK	PN	ZK	PN	ZK	PN	ZK
TSK1	relF1SD	0,082	0,156	0,134	0,107	0,142	0,088	0,153	0,157
TSK1	relF2SD	0,018	0,032	0,023	0,027	0,037	0,048	0,044	0,039
TSK1	relF0SD	0,090	0,081	0,052	0,065	0,087	0,081	0,078	0,090
TSK1	relSEOSD	0,714	0,776	0,521	0,508	0,094	0,160	0,529	0,739
TSK2	AR	3,235	7,402	8,283	5,150	4,068	10,219	9,494	10,255
TSK2	DurMAD	0,013	0,027	0,033	0,008	0,074	0,029	0,049	0,025
TSK2	DurMED	0,019	0,035	0,046	0,015	0,093	0,026	0,062	0,030
TSK2	NST	5,355	7,409	6,610	5,704	10,245	13,732	18,015	15,846
TSK2	PPR	0,066	0,113	0,119	0,056	0,055	0,099	0,108	0,117
TSK2	relF1SD	0,060	0,116	0,125	0,099	0,106	0,080	0,128	0,127
TSK2	relF2SD	0,023	0,031	0,012	0,016	0,035	0,046	0,045	0,041
TSK2	relF0SD	0,049	0,050	0,095	0,063	0,069	0,065	0,079	0,074
TSK2	relSEOSD	0,223	0,631	0,255	0,142	0,113	0,142	0,208	0,517
TSK2	SPIR	0,004	0,005	0,003	0,006	0,007	0,012	0,008	0,008
TSK2	TPT	10,007	11,924	10,612	4,662	6,257	8,518	11,053	10,935
TSK2	TST	10,404	11,739	5,940	7,187	9,357	8,723	22,078	17,969
TSK3	shimmer (APQ)	2,108	3,934	4,023	11,272	4,864	6,033	4,444	8,084
TSK3	mean HNR	3,705	4,422	2,303	3,448	5,330	5,717	4,815	5,041
TSK3	median HNR	3,493	4,525	2,282	3,611	5,270	5,731	4,712	5,179
TSK3	jitter (PPQ)	0,417	0,745	0,235	1,174	0,841	0,575	0,649	0,846
TSK3	relF1SD	0,104	0,081	0,070	0,117	0,069	0,054	0,083	0,091
TSK3	relF2SD	0,042	0,070	0,104	0,091	0,125	0,099	0,103	0,085
TSK3	relF0SD	0,112	0,114	0,024	0,141	0,099	0,056	0,087	0,115
TSK4	shimmer (APQ)	0,962	2,960	3,646	12,427	4,909	2,331	4,340	8,409
TSK4	mean HNR	2,126	4,360	1,823	3,988	5,304	4,264	5,036	5,651
TSK4	median HNR	1,828	4,309	1,766	4,089	5,255	4,544	5,064	5,688
TSK4	jitter (PPQ)	0,871	0,182	0,758	2,698	0,275	0,099	0,574	1,375
TSK4	relF1SD	0,028	0,023	0,024	0,024	0,130	0,071	0,101	0,047
TSK4	relF2SD	0,022	0,054	0,083	0,073	0,105	0,127	0,092	0,080
TSK4	relF0SD	0,013	0,005	0,036	0,007	0,004	0,004	0,023	0,005
TSK5	shimmer (APQ)	13,929	7,785	4,155	8,355	4,483	3,513	8,004	8,112
TSK5	mean HNR	10,414	6,243	4,207	3,690	5,406	4,374	7,232	6,101
TSK5	median HNR	10,894	6,170	4,264	3,687	5,395	4,309	7,418	6,067
TSK5	jitter (PPQ)	0,537	0,420	0,213	0,163	0,117	0,104	0,326	0,345
TSK5	relF1SD	0,043	0,111	0,181	0,064	0,088	0,078	0,133	0,098
TSK5	relF2SD	0,047	0,153	0,149	0,098	0,233	0,203	0,227	0,167
TSK5	relF0SD	0,120	0,041	0,007	0,009	0,005	0,004	0,064	0,032
TSK6	shimmer (APQ)	1,270	2,787	6,836	7,851	3,388	3,003	5,068	6,857
TSK6	jitter (PPQ)	0,556	0,650	0,516	0,843	0,585	0,520	0,503	0,743
TSK6	relF1SD	0,070	0,145	0,182	0,128	0,140	0,094	0,142	0,133
TSK6	relF2SD	0,085	0,139	0,141	0,058	0,134	0,147	0,124	0,131
TSK6	relF0SD	0,148	0,104	0,117	0,122	0,099	0,060	0,120	0,102
TSK6	MPT	8,756	7,549	4,093	7,781	5,857	8,305	6,020	7,852
TSK7	PR	0,588	0,879	0,759	1,015	0,848	0,500	0,919	0,907
TSK7	COV	4,283	3,115	6,632	5,458	1,402	2,904	4,719	3,908
TSK7	RI	0,017	0,025	0,031	0,024	0,007	0,015	0,020	0,024
TSK7	PA	5,124	4,554	9,358	5,537	3,487	3,020	6,407	4,818
TSK7	RA	0,005	0,007	0,011	0,018	0,004	0,003	0,006	0,011
GLOBAL	VAI	0,064	0,226	0,172	0,105	0,241	0,115	0,252	0,239

## 5.2 Vícerozměrná analýza

V Tabulce 5.5 jsou vypsány nalezené optimální hyperparametry modelu pomocí metody random search pro každý ze čtyř datasetů. Tyto hyperparametry při použití vykazují nejvýše dosažitelnou hodnotu  $ACC_{\text{bal}}$ .

Tab. 5.5: Optimální hyperparametry modelů jednotlivých jazykových skupin

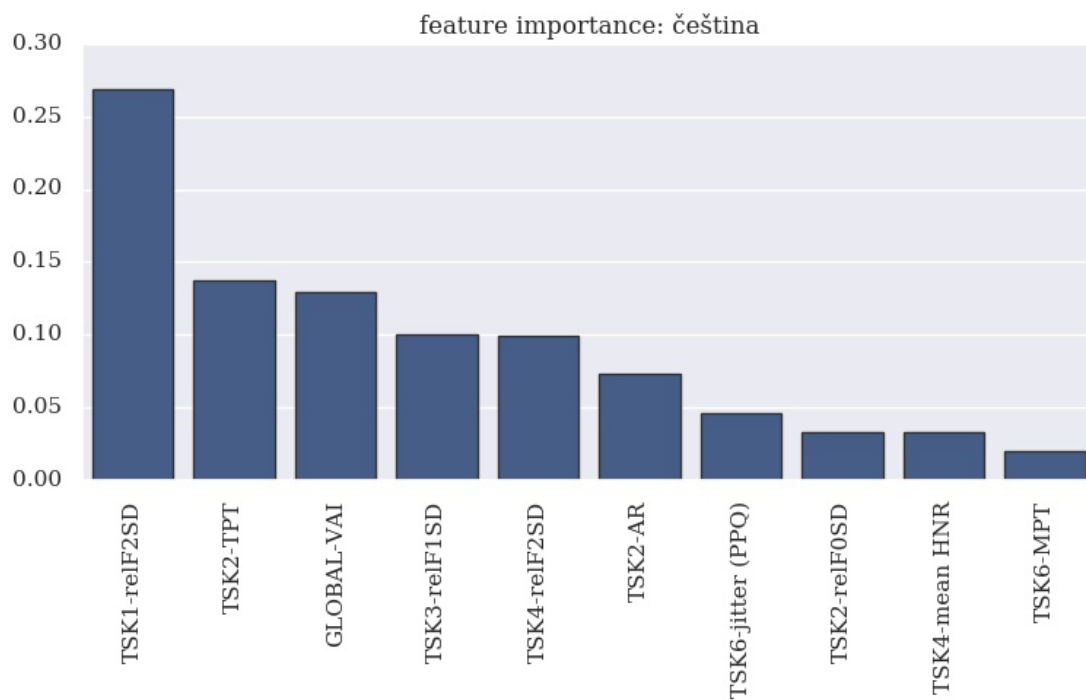
	cs	en-US	hu	vj
<b>learning_rate</b>	0,001	0,1	0,001	0,001
<b>gamma</b>	0,1	0,5	0,1	0,25
<b>max_depth</b>	12	12	15	40
<b>subsample</b>	1	0,5	0,7	0,6
<b>colsample_byLevel</b>	0,6	0,6	0,4	0,5
<b>colsample_byTree</b>	0,9	0,8	0,7	0,9
<b>min_child_weight</b>	5	1	5	5
<b>scale_pos_weight</b>	6	1	7	3

Tabulka 5.6 udává výsledky křížové validace po natrénování modelu s optimálními hyperparametry. Pro každou metriku každého datasetu je uveden aritmetický průměr a směrodatná odchylka.

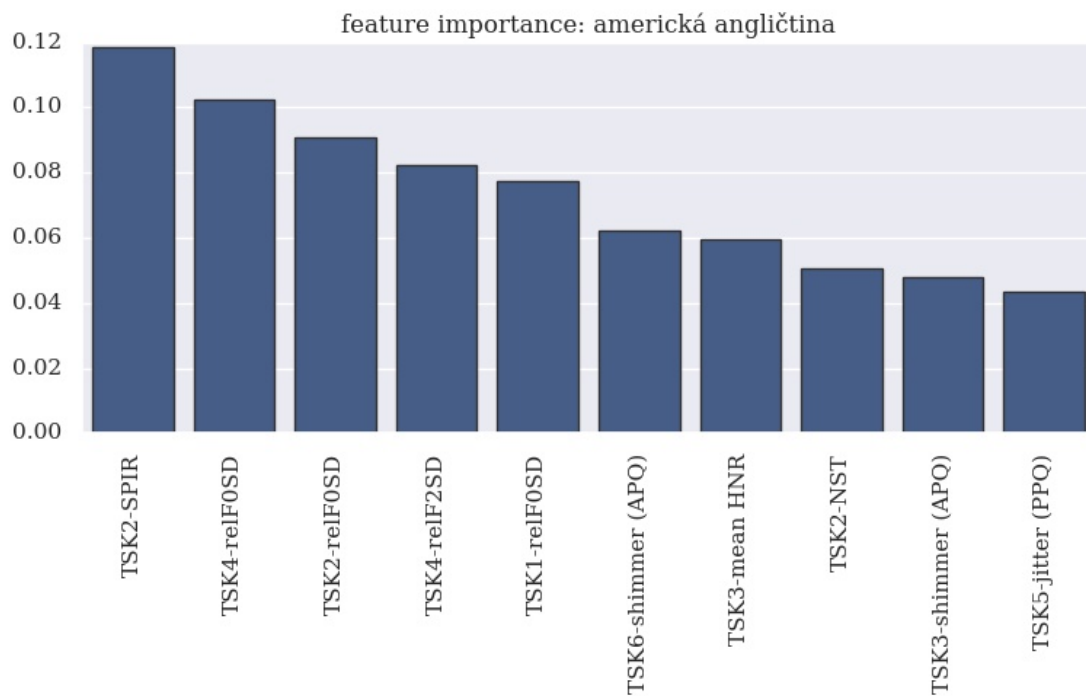
Tab. 5.6: Úspěšnosti modelů strojového učení v %

	cs		en-US		hu		vj	
	$\bar{x}$	$s$	$\bar{x}$	$s$	$\bar{x}$	$s$	$\bar{x}$	$s$
<b>ACC</b>	<b>75</b>	22	<b>79</b>	29	<b>54</b>	7	<b>59</b>	16
<b>SEN</b>	61	49	67	47	100	0	72	30
<b>SPE</b>	72	25	84	34	0	0	53	22

Na obrázcích 5.14 až 5.17 je pro každý dataset vyjádřeno deset nejdůležitějších kombinací parametru s řečovou úlohou pro daný model. Jako míra důležitosti je zde myšlena váha, která představuje poměr počtu výskytů konkrétního parametru v rozhodovacích stromech modelu k celkovému počtu výskytů všech parametrů. Výskyt parametru je v metodě XGBoost závislý na tom, do jaké míry jeho použití v jakékoli větvi jakéhokoli rozhodovacího stromu ovlivňuje celkovou pravděpodobnost s jakou model klasifikuje vstupní hodnoty do dvou výstupních (v našem případě jako PN nebo ZK). Parametr je tedy tím důležitější, čím vícekrát byl v modelu použit.

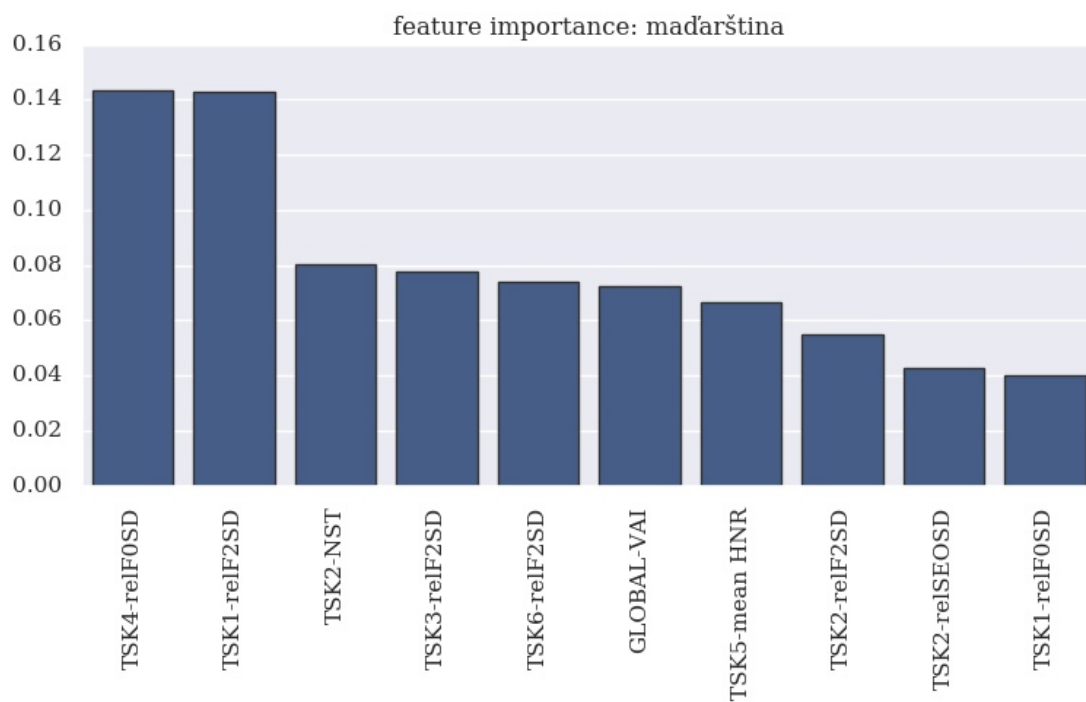


Obr. 5.14: Důležité parametry datasetu cs

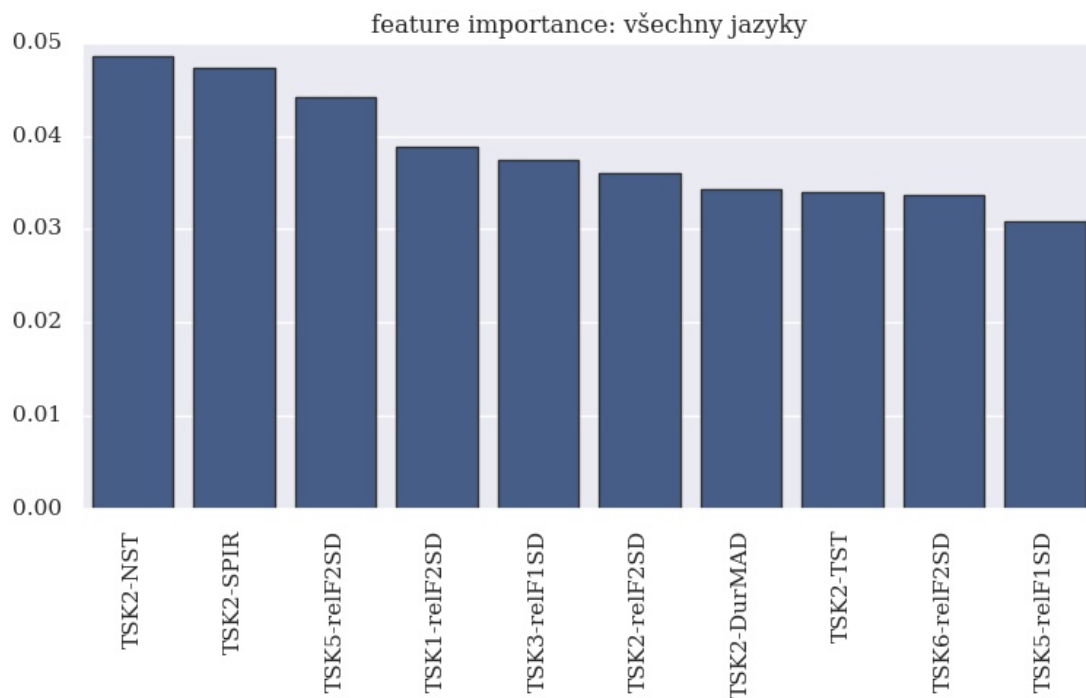


Obr. 5.15: Důležité parametry datasetu en-US





Obr. 5.16: Důležité parametry datasetu hu



Obr. 5.17: Důležité parametry datasetu vj

## 6 Diskuze

Testovaná databáze obsahovala 78 nahrávek, z toho 39 lidí hovořilo česky, 21 americkou angličtinou a 18 maďarsky. V porovnání se studii do roku 2020, kde nejvyšší počet nahrávek v databázi pro jeden jazyk obsahoval 252 nahrávek, je to o 213 nahrávek méně pro dataset cs, o 231 pro en-US a o 234 pro hu. Dataset s hu je stejně velký jako nejmenší existující, pro ten však nebyla provedena vícerozměrná analýza. Z toho vyplývá, že testovaná databáze je vzhledem k doposud publikovaným relativně malá. Celkově bylo testováno a analyzováno 49 kombinací akustických parametrů a řečových úloh, což je při porovnání s počtem parametrů analyzovaných v ostatních publikovaných pracích relativně velké množství. Analýza proběhla pro 4 jazykové skupiny: cs, en-US, hu a vj a vstupovaly do ní parametry, u kterých byl odstraněn vliv věku a pohlaví mluvčího. Takové předzpracování dat skýtá výhodu v odstranění faktorů, které negativně ovlivňují věrohodnost výsledků analýz.

### 6.1 Jednorozměrná analýza

Z celkového množství 196 případů 53 nevykazovalo normální rozdělení (16/49 pro cs, 6/49 pro en-US, 3/49 pro hu a 28/49 pro vj). Je však nutné podotknout, že výsledky Kolmogorova-Smirnovova testu mohou být závislé nejen na samotném parametru, ale také na počtu testovaných subjektů, což souvisí s velikostí testované databáze.

Ze 196 případů bylo dle Mann-Whitneyho U testu pouze 13 signifikantních. Přehled těchto parametrů se souvisejícími řečovými úlohami, ze kterých byly vypočítány, je vypsán v Tabulce 6.1. Dále je zde popsáno, jak se každý parametr změnil u člověka diagnostikovaného s PN. Změna je zde vyjádřena nárůstem nebo poklesem aritmetického průměru, mediánu a směrodatné odchylky. Tabulka také vždy uvádí oblast projevu HD a konkrétní řečovou poruchu spojenou s parametrem. Nejvíce těchto konvenčních parametrů spadá do fonace (5/13) a prozodie (5/13) a zbytek je z oblasti artikulace (3/13). Tyto parametry, pro které lze tvrdit, že vykazují signifikantní rozdíl mezi PN a ZK, jsou ovšem různé pro každý jazyk. Pro maďarštinu nevyšel ani jeden parametr signifikantní a z toho plyne, že pro tento jazyk se řeč příliš nemění v závislosti na nemoci. Nutno však brát v potaz fakt, že databáze s nahrávkami maďarštiny je ze všech nejmenší, a určitou roli zde také může hrát vysoká úroveň šumu, která se právě v těchto nahrávkách vyskytuje.

Z celkového množství 49 kombinací akustických parametrů a řečových úloh bylo v Kruskal-Wallisově testu 28 signifikantních, což znamená, že se liší v závislosti na národnosti.

Tab. 6.1: Přehled signifikantních parametrů na základě statistické analýzy

Parametr	Řečová úloha	Změna s PN	Oblast	Konkrétní porucha
<b>cs</b>				
<b>relF2SD</b>	TSK1	$\bar{x}:\uparrow$ $Me:\uparrow$ $s:\downarrow$	Artikulace	Rigidita jazyka a čelisti
<b>relF2SD</b>	TSK2	$\bar{x}:\uparrow$ $Me:\uparrow$ $s:\downarrow$	Artikulace	Rigidita jazyka a čelisti
<b>relF2SD</b>	TSK4	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\downarrow$	Fonace	Chvění brady a jazyka
<b>en-US</b>				
<b>relF0SD</b>	TSK1	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\downarrow$	Prozodie	Monotónnost
<b>relF0SD</b>	TSK2	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\uparrow$	Prozodie	Monotónnost
<b>SPIR</b>	TSK2	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\downarrow$	Prozodie	Nevhodné pauzy
<b>mean HNR</b>	TSK3	$\bar{x}:\uparrow$ $Me:\uparrow$ $s:\downarrow$	Fonace	Zvýšený šum
<b>mdeian HNR</b>	TSK3	$\bar{x}:\uparrow$ $Me:\uparrow$ $s:\downarrow$	Fonace	Zvýšený šum
<b>shimmer (APQ)</b>	TSK6	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\downarrow$	Fonace	Mikroperturbace v amplitudě
<b>VAI</b>	GLOBAL	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\uparrow$	Artikulace	Snížená hybnost jazyka
<b>vj</b>				
<b>NST</b>	TSK2	$\bar{x}:\uparrow$ $Me:\uparrow$ $s:\uparrow$	Prozodie	Nižší tempo
<b>SPIR</b>	TSK2	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:-$	Prozodie	Nevhodné pauzy
<b>relF2SD</b>	TSK3	$\bar{x}:\downarrow$ $Me:\downarrow$ $s:\uparrow$	Fonace	Chvění brady a jazyka

Z výsledků statistické analýzy je patrné, že pro každý jazyk je důležitý jiný parametr. Jsou ale i parametry, které mají vysokou diskriminační sílu při porovnávání PN a ZK všech jazyků dohromady a těmi jsou prozodické parametry NST a SPIR a fonační parametr **relF2SD** počítán z úlohy prodloužené fonace samohlásky [a] (TSK3). Poslední z nich se navíc neliší v závislosti na národnosti a proto je z hlediska vysoké diskriminační síly a nezávislosti na jazyku nejzajímavější.

## 6.2 Vícerozměrná analýza

Pro klasifikaci řečníků na PN a ZK byla vybrána metoda XGBoost, protože v porovnání s metodou náhodných lesů vykazovala lepší výsledky.

Optimální kombinace hyperparametrů pro dosažení nejvyšší hodnoty  $ACC_{bal}$  je dle random search pro každý jazyk odlišná (viz Tab. 5.5). S tím souvisí odlišná výsledná stabilita modelu každého jazyku. Zároveň se však nedá říct, že by hyperparametry modelu jednoho datasetu vykazovaly větší předpoklad vytvářet konzervativnější model než hyperparametry modelů ostatních jazyků.

Z hlediska klasifikační přesnosti (viz Tab 5.6) je nejúspěšnější model pro en-US, kde ACC dosahuje 79 %. Tento model je však zároveň nejméně stabilní se směrodatnou odchylkou 29 %. Druhý nejúspěšnější model je s datasetem cs, jehož ACC dosahuje 75 % se směrodatnou odchylkou 22 %. Pro model s hu dosahuje ACC pouze 54 %, ale za to se 100 % SEN. To může být způsobeno malou velikostí tohoto datasetu nebo kvalitou maďarských nahrávek, které obsahují nežádoucí šum.

Úspěšnost modelu pro vj se od ostatních zásadně neliší. Přestože se jeho klasifikační přesnost od modelů cs nebo en-US snížila na 59 %, senzitivita vzrostla na 72 %. Pro specifitu je tomu naopak, ale zatímco je mnohem důležitější přesněji klasifikovat řečníky s PN jako PN než řečníky se ZK jako ZK, hraje zde senzitivita velkou roli a její nárůst je pozitivním zjištěním. Z tohoto důvodu má smysl zabývat se dále analýzou a sestavováním modelů pro všechny jazyky dohromady tak, aby akustická analýza za cílem rozpoznat PN mohla probíhat v různých zemích, které například nemají sestaveny percepční testy pro svůj jazyk. Tento model strojového učení by mohl být vylepšen například tím, že by do něj vstupovaly pouze parametry, které se nemění v závislosti na jazyku. Při porovnání s dosavadními studiemi, kde úspěšnosti modelů dosahují až 100 % a to pro ACC, SEN i SPE, se tyto modely jeví méně úspěšné. To znamená, že je stále co vylepšovat a to především rozšířením databáze s nahrávkami řeči. Na druhou stranu jsou tyto modely relativně stabilní a úspěšnosti modelů jsou určovány 10-fold křížovou krosvalidací s 20 opakováním a proto jsou výsledky řádně ověřeny.

Z grafů 5.14 až 5.17 lze vyzorovat, že pro každý dataset jsou parametry z hlediska schopnosti modelu klasifikovat řečníky různě důležité. Pro češtinu je nejdůležitější parametr relF2SD počítán z úlohy TSK1, který razantně převyšuje ty zbylé. Téměř 27 % všech rozhodovacích stromů v modelu bylo tvořeno právě tímto parametrem. Pro americkou angličtinu je nejdůležitější SPIR z TSK2. Pro maďarštinu je to relF0SD z TSK4 a relF2SD z TSK1 vyskytující se v modelu z 14 %. V modelu pro všechny jazyky hraje největší roli parametr NST počítán z TSK2. Všechny tyto parametry jsou zároveň signifikantní na základě předchozí statistické analýzy a tím se potvrzuje jejich důležitost. Až na jeden globální parametr VAI jsou všechny důležité parametry počítány z úlohy monolog, čtení nebo prodloužená fonace.

Nejčastěji se v modelech objevuje parametr **relF2SD** a proto je možné jej považovat za nejdůležitější. Tento parametr popisuje, jak se mění druhý format. Z úlohy monolog (TSK1), kde kvantifikuje rigiditu jazyka a čelisti, je důležitý pro cs, hu a vj. Stejná porucha se také vyskytuje v úloze čtení (TSK2) pro hu a vj. Jako fonační parametr je důležitý pro každou jazykovou skupinu, přičemž signifikantní změnu chvění brady a jazyka kvantifikuje z úlohy prodloužená fonace samohlásky [a] (TSK3) pro hu, samohlásky [i] (TSK4) pro cs nebo en-US a samohlásky [u] (TSK5) pro vj. Důležitý je také pro hu a en-US v kombinaci s TSK6. Druhým nejdůležitějším parametrem z pohledu vícerozměrné analýzy je NST z TSK2 a to kvůli svému častému výskytu v modelech en-US, hu a vj.

# Závěr

Tato práce se zabývala automatizovanou diagnózou hypokinetické dysartrie se zaměřením na závislost akustických parametrů na národnosti. Nejdříve byla popsána Parkinsonova nemoc a související hypokinetická dysartrie, je zde uvedeno v jakých oblastech se projevuje a jakým způsobem se diagnostikuje. Následně byla provedena rešerše na zmíněné téma a sepsány zjištěné důležité poznatky.

Hlavním cílem práce bylo zjistit, které akustické parametry mají vysokou diskriminační sílu a které jsou závislé na konkrétním jazyku mluvčího. Proběhla tedy parametrizace řečových úloh a na základě statistické analýzy bylo zjištěno, že z celkových 49 kombinací akustických parametrů a řečových úloh se 28 liší v závislosti na národnosti. Také bylo zjištěno, že pouze 13 z těchto kombinací testovaných pro každý jazyk zvlášť a dohromady vykazuje signifikantní rozdíl mezi parkinsoniky a zdravými kontrolami. V průběhu práce bylo také potřeba znát, jestli parametry vykazují normální rozdělení pravděpodobnosti, aby mohl být zvolen správný statistický test. Tím bylo navíc zjištěno, že ze všech 196 případů vykazovalo normální rozdělení pouze 143.

Dále proběhlo matematické modelování na základě strojového učení. Vytvořené modely jsou schopny zjistit přítomnost hypokinetické dysartrie na základě vypočítaných akustických parametrů. Nejúspěšnější byl se svou 79% klasifikační přesností model pro americkou angličtinu a hned za ním model pro češtinu. Model pro maďarštinu byl nejméně přesný. Proběhlo také modelování pro všechny jazyky dohromady a úspěšnost se od ostatních modelů výrazně nelišila. Na základě vícerozměrné analýzy těchto modelů bylo zjištěno, které parametry jsou důležité pro který model a výsledky se do značné míry shodují s výsledky dříve zmíněné statistické analýzy. Pozornost byla dále zaměřena na obecně nejdůležitější parametry z hlediska univerzálnosti pro jazyk.

Je nutné zmínit fakt, že testovaná databáze byla vzhledem k doposud publikovaným relativně malá, což může mít vliv na výsledky testů. V potaz je bráno také to, že špatná parametrizace řečových úloh může vést ke klamným a zavádějícím závěrům. Tomu se snažilo předejít odstraněním vlivu věku a pohlaví, avšak mohou existovat další faktory, které ovlivňují akustickou analýzu řeči, jako třeba nálada mluvčího, medikace nebo jeho celkový zdravotní stav. V nahrávkách s maďarštinou se navíc objevuje nežádoucí šum, což může mít také určitý vliv na akustickou analýzu.

Tato práce je prozatím jediná multilingvální studie s databází obsahující češtinu, americkou angličtinu a maďarštinu. Výsledky lze uplatnit ve snaze najít standardizovaný set akustických parametrů pro rozpoznání Parkinsonovy nemoci nezávisle na jazyku mluvčího. Vytvořený program je univerzální a lze tak přidávat do databáze nahrávky s různými nemocemi nebo národnostmi a pokračovat tak v dalších studiích.

## Literatura

- [1] O. Hornykiewicz, "Biochemical aspects of Parkinson's disease," *Neurol. Suppl.* 51(2), S2–S9 (1998).
- [2] Narabayashi H. The neural mechanisms and progressive nature of symptoms of Parkinson's disease-based on clinical, neurophysiological and morphological studies. *J Neural Transm [P-D Sect]* 1995;10:63-75.
- [3] DeMaagd G., Philip A. Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis. *Pharm. Ther.* 2015;40:504–532.
- [4] Parkinson J. *An Essay on the Shaking Palsy.* London: Sherwood, Neely, and Jones; 1817. pp. 1–16.
- [5] Goetz C. G. (2011). The history of Parkinson's disease: early clinical descriptions and neurological therapies. *Cold Spring Harb. Perspect. Med.* 1:a008862. doi:10.1101/cshperspect.a008862
- [6] OWNWARD, Emily. What Are the Stages of Parkinson's Disease? [online]. [cit. 2019-12-20]. Dostupné z: <https://parkinsonsdisease.net/basics/stages/>
- [7] Ho AK, Ianssek R, Marigliani C, Bradshaw JL, Gates S (1999) Speech impairment in a large sample of patients with Parkinson's disease. *J Behav Neurol* 11:131–137
- [8] MEKYSKA, J.; SMÉKAL, Z.; KOŠTÁLOVÁ, M.; MRAČKOVÁ, M.; SKUTILOVÁ, S.; REKTOROVÁ, I. Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení. *Česká a slovenská neurologie a neurochirurgie*, 2011, vol. 74, no. 6, p. 662-668. ISSN: 1210- 7859.
- [9] DUFFY, Joseph R. *Motor speech disorders: substrates, differential diagnosis, and management.* Third edition. St. Louis, Missouri: Elsevier, [2013]. ISBN 978-0-323-07200-7.
- [10] Ackermann, H., Hage, S. R., Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences*, 37(06), 529–546. doi:10.1017/s0140525x13003099.
- [11] KOŠTÁLOVÁ, M, M MRAČKOVÁ, R MAREČEK, et al. Test 3F Dysartrický profil – normativní hodnoty řeči v češtině: 2; I. Eliášová 3; E. Janoušová 4; J.

- Roubíčková 5; J. Bednařík 1,2; I. Rektorová 2,3. Česká a slovenská neurologie a neurochirurgie. 2013, 2013(5), 614-618.
- [12] ROUBÍČKOVÁ, Jaroslava. Test 3F: dysartrický profil. 3., dopl. a přeprac. vyd., (V nakl. Galén 1.). Praha: Galén, c2011. ISBN 978-80-7262-714-1.
- [13] KOWOLOWSKI, Alexander. Vývoj moderních akustických parametrů kvantifikujících hypokinetickou dysartrii. Brno, 2019. Dostupné také z: <https://www.vutbr.cz/studenti/zavprace/detail/118146>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce Zoltán Galáž.
- [14] RUSZ, J., R. CMEJLA, H. RUZICKOVA a E. RUZICKA. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America*. 2011, 129(1), 350-367. DOI: 10.1121/1.3514381. ISSN 0001-4966.
- [15] BRABENEC, L., J. MEKYSKA, Z. GALAZ a Irena REKTOROVA. Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *Journal of Neural Transmission*. 2017, 124(3), 303-334. DOI: 10.1007/s00702-017-1676-0. ISSN 0300-9564.
- [16] Yücelbaş, Ş. (2020). Simple Logistic Hybrid System Based on Greedy Stepwise Algorithm for Feature Analysis to Diagnose Parkinson's Disease According to Gender. *Arabian Journal for Science and Engineering*. doi:10.1007/s13369-020-04357-1
- [17] Moro-Velazquez, L., Gomez-Garcia, J.A., Godino-Llorente, J.I. et al. Phonetic relevance and phonemic grouping of speech in the automatic detection of Parkinson's Disease. *Sci Rep* 9, 19066 (2019). <https://doi.org/10.1038/s41598-019-55271-y>
- [18] Vasquez, Juan Arias-Vergara, T. Rios-Urrego, Cristian Schuster, Maria Rusz, Jan Orozco-Aroyave, Juan Noeth, Elmar. (2019). Convolutional Neural Networks and a Transfer Learning Strategy to Classify Parkinson's Disease from Speech in Three Different Languages.
- [19] Sztaho, D., Valalik, I., Vicsi, K. (2019). Parkinson's Disease Severity Estimation on Hungarian Speech Using Various Speech Tasks. 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD). doi:10.1109/sped.2019.8906277

- [20] M. Dahmani and M. Guerti, "Glottal signal parameters as features set for neurological voice disorders diagnosis using K-Nearest Neighbors (KNN)," 2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP), Algiers, 2018, pp. 1-5.
- [21] Moro-Velazquez, L., Gomez-Garcia, J. A., Godino-Llorente, J. I., Rusz, J., Skodda, S., Grandas, F., ... Dehak, N. (2018). Study of the Automatic Detection of Parkinson's Disease Based on Speaker Recognition Technologies and Allophonic Distillation. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2018.8512562
- [22] Tundik, M. A., Kiss, G., Sztaho, D., Szaszak, G. (2017). Assessment of pathological speech prosody based on automatic stress detection and phrasing approaches. 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom). doi:10.1109/coginfocom.2017.8268218
- [23] Hsu, S.-C., Jiao, Y., McAuliffe, M. J., Berisha, V., Wu, R.-M., Levy, E. S. (2017). Acoustic and perceptual speech characteristics of native Mandarin speakers with Parkinson's disease. *The Journal of the Acoustical Society of America*, 141(3), EL293–EL299. doi:10.1121/1.4978342
- [24] Lahmiri, S., Dawson, D. A., Shmuel, A. (2017). Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures. *Biomedical Engineering Letters*, 8(1), 29–39. doi:10.1007/s13534-017-0051-2
- [25] Mucha, J., Galaz, Z., Mekyska, J., Kiska, T., Zvoncak, V., Smekal, Z., ... Alonso-Hernandez, J. B. (2017). Identification of hypokinetic dysarthria using acoustic analysis of poem recitation. 2017 40th International Conference on Telecommunications and Signal Processing (TSP). doi:10.1109/tsp.2017.8076086
- [26] Godino-Llorente, J. I., Shattuck-Hufnagel, S., Choi, J. Y., Moro-Velázquez, L., Gómez-García, J. A. (2017). Towards the identification of Idiopathic Parkinson's Disease from the speech. New articulatory kinetic biomarkers. *PLOS ONE*, 12(12), e0189583. doi:10.1371/journal.pone.0189583
- [27] Naranjo L, Pérez CJ, Campos-Roca Y, Martín J (2016) Addressing voice recording replications for Parkinson's disease detection. *Expert Syst Appl* 46:286–292
- [28] Orozco-Aroyave JR, Hönig F, Arias-Londono JD, Vargas-Bonilla JF, Dargrouq K, Skodda S, Rusz J, Nöth E (2016) Automatic detection of Parkinson's disease in running speech spoken in three different languages. *J Acoust Soc Am* 139(1):481–500



- [29] Mekyska J, Smekal Z, Galaz Z, Mzourek Z, Rektorova I et al (2016) Perceptual features as markers of Parkinson's Disease: the issue of clinical interpretability. In: Esposito A, Faundez-Zanuy M, Esposito AM, Cordasco G, Casals JS et al (eds) Recent advances in nonlinear speech processing. Springer International Publishing, New York, pp 83–91
- [30] Galaz Z, Mekyska J, Mzourek Z, Smekal Z, Rektorova I, Eliasova I, Kostalova M, Mrackova M, Berankova D (2016) Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Comput Methods Programs Biomed.* doi:10.1016/j.cmpb.2015.12.011
- [31] Smekal Z, Mekyska J, Galaz Z, Mzourek Z, Rektorova I, Faundez- Zanuy M (2015) Analysis of phonation in patients with Parkinson's disease using empirical mode decomposition. In: 2015 International Symposium on Signals, Circuits and Systems (ISSCS), Iasi, pp 1–4
- [32] Mekyska J, Galaz Z, Mzourek Z, Smekal Z, Rektorova I, et al. (2015a) Assessing Progress of Parkinson's Disease Using Acoustic Analysis of Phonation. In: International Work Conference on Bioinspired Intelligence (IWOBI 2015), pp 115–122
- [33] Vásquez-Correa JC, Arias-Vergara T, Orozco-Arroyave JR, Vargas- Bonilla JF, Arias-Londono JD, Nöth E (2015) Automatic detection of Parkinson's disease from continuous speech recorded in non-controlled noise conditions. In: Proceedings of 16th INTERSPEECH, Dresden, Germany, pp 105–109
- [34] Musa P, Baha S, Delen D (2015) Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm. *J Healthc Eng* 6(3):281–302
- [35] Arora S, Venkataraman V, Zhan A, Donohue S, Biglan KM, Dorsey ER, Little MA (2015) Detecting and monitoring the symptoms of Parkinson's disease using smartphones: a pilot study. *Parkinsonism Relat Disord* 21(6):650–653
- [36] Orozco-Arroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Skodda S, Rusz J, Nöth E (2015b) Voiced/unvoiced transitions in speech as a potential bio-marker to detect Parkinson's disease. In: Proceedings of 16th INTERSPEECH, Dresden, Germany, pp 95–99
- [37] Benba A, Jilbab A, Hammouch A, Sandabad S (2015) Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. In: 2015 International Conference on Electrical and Information Technologies (ICEIT), Marrakech, pp 300–304

- [38] Hariharan M, Polat K, Sindhu R (2014) A new hybrid intelligent system for accurate detection of Parkinson’s disease. *Comput Methods Program Biomed* 113(3):904–913
- [39] Orozco-Aroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Skodda S, Rusz J, Nöth E (2014a) Automatic detection of Parkinson’s disease from words uttered in three different languages. In: *Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTER-SPEECH)*, Singapore, pp 1573–1577
- [40] Novotny M, Rusz J, Cmejla R, Ruzicka E (2014) Automatic evaluation of articulatory disorders in Parkinson’s disease. *IEEE/ACM T Audio Speech Lang Process* 22:1366–1378
- [41] Villa-Canas T, Orozco-Aroyave JR, Vargas-Bonilla JF, Arias-Londoño JD (2014) Modulation spectra for automatic detection of Parkinson’s disease. In: *2014 XIX Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, Armenia, pp 1–5
- [42] Orozco-Aroyave JR, Belalcázar-Bolaños EA, Arias-Londoño JD, Vargas-Bonilla JF, Haderlein T, Nöth E (2014b) Phonation and articulation analysis of spanish vowels for automatic detection of Parkinson’s disease. *Lecture notes in artificial intelligence*, vol 8655. Springer, pp 389–296
- [43] Orozco-Aroyave J, Arias-Londoño JD, Vargas-Bonilla JF, Nöth E (2013a) Analysis of speech from people with Parkinson’s disease through nonlinear dynamics. *Lect Notes Artif Intell* 7911:112–119
- [44] 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. *J Acoust Soc Am* 134:2171–2181
- [45] Belalcázar-Bolaños EA, Orozco-Aroyave JR, Vargas-Bonilla JF, Arias-Londoño JD, Castellanos-Domínguez CG, Nöth E (2013b) New Cues in Low-Frequency of Speech for Automatic Detection of Parkinson’s Disease. In José Manuel Ferrández de Vicente, José Ramón Álvarez Sánchez, Félix de la Paz López, Javier Toledo-Moreo, F (eds) ‘*IWINAC (1)*’. Springer, pp 283–292
- [46] Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgun F, Delil S, Apaydin H, Kursun O (2013) Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE J Biomed Health Inform* 17(4):828–834

- [47] Shahbakhti M, Taherifar D, Zareei Z (2013a) Combination of PCA and SVM for diagnosis of Parkinson’s disease. In: 2013 2nd International Conference on Advances in Biomedical Engineering (ICABME), Tripoli, pp 137–140
- [48] Shahbakhti M, Taherifar D, Sorouri A (2013b) Linear and non-linear speech features for detection of Parkinson’s disease. In: 2013 6th Biomedical Engineering International Conference (BMEiCON), Amphur Muang, pp 1–3
- [49] Belalcázar-Bolaños EA, Orozco-Arroyave JR, Vargas-Bonilla JF, Arias-Londoño JD, Castellanos-Domínguez CG, Nöth E (2013b) New Cues in Low-Frequency of Speech for Automatic Detection of Parkinson’s Disease. In José Manuel Ferrández de Vicente, José Ramón Álvarez Sánchez, Félix de la Paz López, Javier Toledo-Moreo, F (eds) ‘IWINAC (1)’. Springer, pp 283–292
- [50] Orozco-Arroyave J.R., Arias-Londoño J.D., Vargas-Bonilla J.F., Nöth E. (2013b) Perceptual Analysis of Speech Signals from People with Parkinson’s Disease. In: Ferrández Vicente J.M., Alvarez Sánchez J.R., de la Paz López F., Toledo Moreo F.J. (eds) Natural and Artificial Models in Computation and Biology. IWINAC 2013. Lecture Notes in Computer Science, vol 7930. Springer, Berlin, Heidelberg
- [51] Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson’s disease. *IEEE T Biomed Eng* 59(5):1264–1271
- [52] Hazan H, Hilu D, Manevitz L, Ramig LO, Sapir S (2012) Early diagnosis of Parkinson’s disease via machine learning on speech data. In: 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel (IEEEI), Eilat, pp 1–4
- [53] Bakar ZA, Ibrahim NF, Sahak R, Tahir NM (2012) Parkinson’s disease feature subset selection based on voice samples. In: 2012 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE), Kota Kinabalu, pp 163–166
- [54] Aström F, Koker R (2011) A parallel neural network approach to prediction of Parkinson’s Disease. *Expert Syst Appl* 38(10):12470–12474
- [55] Rusz J, Cmejla R, Ruzickova H, Ruzicka E (2011) Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson’s disease. *J Acoust Soc Am* 129:350–367

- [56] Bocklet T, Nöth E, Stemmer G, Ruzickova H, Rusz J (2011) Detection of persons with Parkinson's disease by acoustic, vocal, and prosodic analysis. In: 2011 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Waikoloa, HI, pp 478–483
- [57] Skodda S, Gronheit W, Schlegel U (2011c) Intonation and speech rate in Parkinson's disease: general and dynamic aspects and responsiveness to levodopa admission. *J Voice*. doi:10.1016/j.jvoice.2010.04.007
- [58] Shirvan RA, Tahami E (2011) Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classification method. In: 2011 18th Iranian Conference of Biomedical Engineering (ICBME), Tehran, pp 278–283
- [59] Spadoto AA, Guido RC, Papa JP, Falcaõ AX (2010) Parkinson's disease identification through optimum-path forest. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Buenos Aires, pp 6087–6090
- [60] Sakar CO, Kursun O (2009) Tlediagnosis of Parkinson's disease using measurements of dysphonia. *J Med Syst* 34(4):591–599

## Seznam symbolů, veličin a zkratek

<b>ACC</b>	přesnost
$ACC_{\text{bal}}$	balanced accuracy
<b>cs</b>	čeština
<b>de</b>	němčina
<b>es</b>	španělština
<b>en</b>	angličtina
<b>en-US</b>	americká angličtina
<b>FN</b>	falešně negativní
<b>FP</b>	falešně pozitivní
<b>GLOBAL</b>	globální parametr
<b>HD</b>	hypokinetická dysartrie
<b>hu</b>	maďarština
<b>KONV</b>	konvenční parametry
$Me$	medián
$n$	rozsah souboru
<b>NEKONV</b>	nekonvenční parametry
<b>PN</b>	Parkinsonova nemoc
$Q_1$	první kvartil
$Q_3$	třetí kvartil
$R$	variační rozpětí
$s$	výběrová směrodatná odchylka
<b>SEN</b>	senzitivita
<b>SN</b>	skutečně negativní
<b>SP</b>	skutečně pozitivní
<b>SPE</b>	specifická
<b>tr</b>	turečtina
<b>UID</b>	identifikátor uživatele
<b>XGBoost</b>	eXtreme Gradient Boosting
$X_{\min}$	nejmenší hodnota
$X_{\max}$	největší hodnota
$\bar{x}$	aritmetický průměr
<b>zh</b>	čínština
<b>ZK</b>	zdravá kontrola

# Seznam příloh

A Obsah přiloženého DVD

54

# A Obsah příloženého DVD

```
DVD
├── matlab
│   ├── corpus-COBEN-001-speech-results-edited-CZ.xlsx
│   ├── corpus-COBEN-001-speech-results-edited-EN.xlsx
│   ├── corpus-COBEN-001-speech-results-edited-HU.xlsx
│   ├── corpus-COBEN-001-speech-tablein.xlsx
│   ├── main.m
│   ├── adjust.m
│   ├── fdr-bh.m
│   ├── KS-test.m
│   ├── missing-reconstruct.m
│   ├── missing-solve.m
│   ├── read-demograf.m
│   ├── read-parameters.m
│   ├── Violin.m
│   ├── violinplot.m
│   ├── matice-test.xlsx
│   └── matice-stat.xlsx
├── python
│   ├── main.py
│   ├── ML-cs.xlsx
│   ├── ML-en.xlsx
│   ├── ML-hu.xlsx
│   └── ML-vj.xlsx
├── řečové úlohy
│   ├── TSK2-Czech.docx
│   ├── TSK2-English.docx
│   ├── TSK2-Hungarian.docx
│   ├── TSK8-Czech.docx
│   ├── TSK8-English.docx
│   └── TSK8-Hungarian.docx
└── DP-Kovac.pdf
```