

HODNOCENÍ DŮRAZU, EMOCÍ, RYTMU, ARTIKULAČNÍ RYCHLOSTI A PRAVIDELNOSTI U PARKINSONOVY NEMOCI

J. Rusz^{1,2)}, R. Čmejla¹⁾, H. Růžičková³⁾

¹⁾ Katedra teorie obvodů, Fakulta elektrotechnická, České vysoké učení technické v Praze

²⁾ Univerzita Karlova v Praze, Laboratoř biokybernetiky a počítačové podpory výuky, Ústav patologické fyziologie 1. LF UK

³⁾ Univerzita Karlova v Praze, Neurologická klinika 1. LF UK a VFN

Abstract

Cílem této práce je vyhodnocení tradičních měření artikulační rychlosti a pravidelnosti provedených s využitím rychlého opakování kombinace slabik známé jako diadochokinetická (DDK) úloha, a dále prezentace nových měření prosodie provedených s využitím nestandardních vokálních úloh zahrnujících důraz, emoce a rytmus. Řečová data byla pořízena od 46 českých rodilých mluvčích, 23 z nich bylo diagnostikováno s výskytem Parkinsonovy nemoci (PN). DDK úloha umožňuje měřit schopnost jednotlivce opakovat kombinaci souhláska-samohláska při rychlé, konstantní a zároveň rytmické artikulaci. Mezi měření prosodie patří řečové nahrávky: (Důraz) čtený text složený z různých vět se zvýšeným důrazem na předem stanovených slovech, což může vést ke zvýšení rozdílů ve variacích základní hlasivkové frekvence a intenzity hlasitosti mezi skupinou zdravých lidí a pacientů s výskytem PN; (Emoce) čtené věty dle definovaných emocí při běžné hlasitosti, které mohou zvýšit rozdíly ve variacích základní hlasivkové frekvence a prvních dvou formantových frekvencích mezi oběma skupinami; (Rytmus) jeden rytmicky čtený text opakováný podle řečového terapeuta pro měření schopnosti napodobit příslušný rytmus. Pro všechna zde prezentovaná měření byly navrženy automatické skripty v prostředí MATLAB. Hlas pacientů s výskytem PN ve srovnání se skupinou zdravých lidí ukázal statisticky významné rozdíly pro měření DDK rychlosti, variace základní hlasivkové frekvence a intenzity u důrazu, variace základní hlasivkové frekvence a druhého formantu u emocí, a rytmu.

1 Úvod

Termín hypokinetická dysartrie (HP) je nejčastěji používán k popisu poruchy řeči u Parkinsonovy nemoci (PN) [1-2]. HP se projevuje omezením rozsahu, rychlosti a pravidelnosti pohybu orofaciálních svalů a deficitem v respiraci, fonaci a fonetice.

V předchozích studiích bylo ukázáno, že při výskytu PN již v raných stádiích dochází k poruchám řeči v různých řečových charakteristikách. V oblasti fonace vzniká větší nepravidelnost kmitání hlasivek při měření amplitudového a frekvenčního kolísání hlasu na udržované hlásce [3]. Další měření z této oblasti vedlo k pozorování, že u pacientů s PN dochází v jejich promluvách k celkovému oslabení okluzí [4]. Při měření rychlé artikulace dochází u pacientů s PN k omezení artikulační přesnosti [5]. Zároveň také klesá intenzita hlasitosti [6]. Měření prosodie byla provedena na čtených textech a v samostatné promluvě. Výsledky ukázaly, že pacienti s výskytem PN mají ve čtených textech menší počet pauz, které jsou ovšem delší, a že mají sníženou artikulační rychlost [6]. Zároveň také u pacientů s výskytem PN dochází k většímu kolísání intenzity a ztrátě melodičnosti [3]. U pacientů s PN byl také nalezen zmenšený rozsah hlasu při zpívané stupnici a vyslovování tázací či rozkazovací věty [3]. Celkově tedy lze tvrdit, že pacienti s výskytem PN inklinují k monotónní promluvě, což se projevuje snížením variace intenzity a základní hlasivkové frekvence. To také může být jedním z významných ukazatelů jejich zhoršené řeči. Ze studie vyplynulo, že dalším důležitým

ukazatelem zhoršené řeči v různých stádiích PN je snížení vnímavosti při emocionální prosodii [7], ale také zhoršená vnímavost důrazu a rytmu [8].

V této práci jsou analyzovány vokalické úlohy emocionální prosodie, intonace důrazu a držení rytmu u pacientů již v raném stádiu PN. Změny při správné intonaci důrazu na předem vybraných slovech se projevují změnou variace základní hlasivkové frekvence a intenzity hlasitosti. Oproti tomu, změny při vyslovování vět s emocionálním kontextem mohou být lépe zachyceny měřením základní hlasivkové frekvence a formantů F1 a F2. Práce se dále zabývá návrhem nového algoritmu pro hodnocení vnímání rytmu.

Pro doplnění tradičních akustických analýz se práce také věnuje návrhu algoritmu určeného k měření přesnosti a pravidelnosti artikulace při diachokinetické (DDK) úloze.

Objektivní analýza hlasových nahrávek může být důležitá z hlediska určení stupně dysartrie, pro následné monitorování zlepšení stavů pacientů, ale hlavně jako zpětná vazba při terapii řeči.

2 Metody

2.1 Účastníci

Této studii se účastnilo 46 Českých pacientů. Dvacet tři jednotlivců (19 mužů a 4 ženy) bylo diagnostikováno s dosud farmakologicky neléčenou PN [průměrný věk 61.74 (\pm SD 12.6), stádium dle Hoehnové a Yakra 1-2], ve srovnání s 23 účastníky KS [průměrný věk 58.1 (\pm SD 12.9)]. KS byla složena z 16 mužů a 7 žen a všichni její účastníci byli bez známek neurologických onemocnění a hlasových poruch.

2.2 Nahrávání promluv

Nahrávky pro všechny pacienty s PN byly pořízeny v odhlučněné místnosti (< 45dB) Neurologické kliniky 1. lékařské fakulty Univerzity Karlovy v Praze ve vysoké kvalitě s vzorkovací frekvencí 48 kHz, s 16 bitovým rozlišením. Účastníci byli požádáni o: (1) rychlé opakování trojice slabik /pa/-/ta/-/ka/ v co nejdelší možné délce; (2) přečtení textu sestaveného z 6 vět se zvýšeným důrazem na předem stanovených slovech; (3) 10 po sobě jdoucích vět podle předem definovaných emocí; (4) jeden rytmicky opakováný text o délce osmi veršů podle referenční nahrávky logopeda.

2.3 Akustické analýzy

Nahrávky pro všechny pacienty s PN byly pořízeny v odhlučněné místnosti (< 45dB) Neurologické kliniky 1. lékařské fakulty Univerzity Karlovy v Praze ve vysoké kvalitě s vzorkovací frekvencí 48 kHz, s 16 bitovým rozlišením. Účastníci byli požádáni o: (1) rychlé opakování trojice slabik /pa/-/ta/-/ka/ v co nejdelší možné délce; (2) přečtení textu sestaveného z 6 vět se zvýšeným důrazem na předem stanovených slovech; (3) 10 po sobě jdoucích vět podle předem definovaných emocí; (4) jeden rytmicky opakováný text o délce osmi veršů podle referenční nahrávky logopeda.

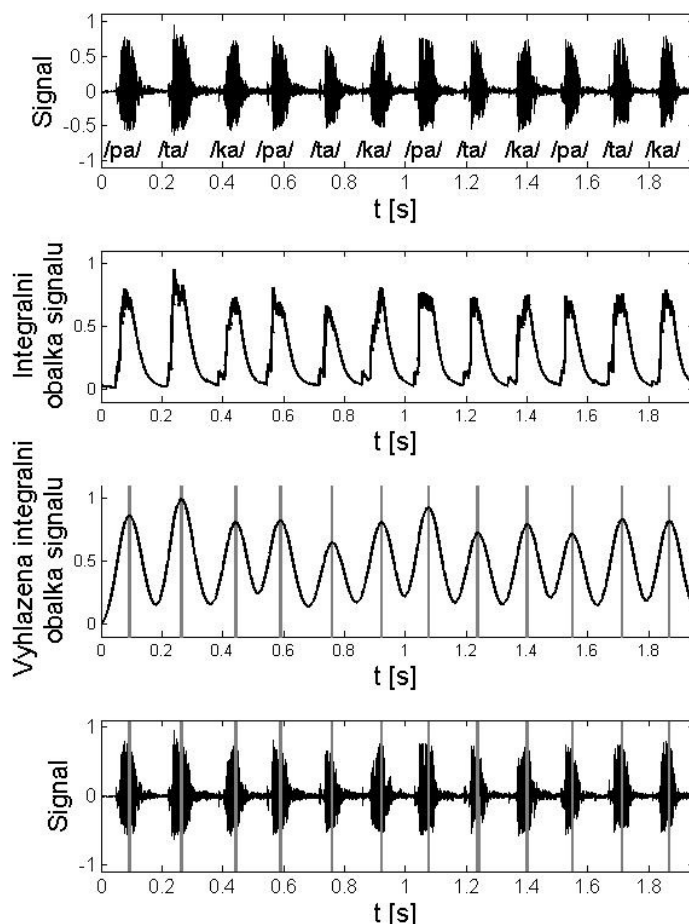
DDK úloha

Tato úloha je jedním z tradičních měření používaných v souvislosti s výskytem PN při vyšetření artikulačních schopností jednotlivce a je založena na rychlém opakování kombinace slabik /pa/-/ta/-/ka/ [8-12]. Tato kombinace souhlásek rovnoměrně zatěžuje hlasový trakt od artikulace oboustranné okluzivy /p/ přes předodásňové /t/ až po měkkopatrové /k/.

DDK úloha obvykle obsahuje dvě měření. Průměrná *DDK rychlost* je měřena jako počet slabik za sekundu. *DDK pravidelnost* měří míru variace mezi jednotlivými slabikami a tak vyhodnocuje schopnost udržet konstantní rychlost vyslovování jednotlivých slabik.

Pro vyhodnocení těchto měření byl navržen nový algoritmus, který je založen na detekci lokálních maxim reprezentujících jednotlivé slabiky. V prvním kroku je řečový signál normalizovaný

do intervalu $[-1, 1]$ a následně je spočtena integrální obálka s integrační konstantou nastavenou na 0.997. Po provedení obousměrné digitální filtrace klouzavým průměrem získáváme vyhlazený sinusoidální signál, který opět normalizujeme do intervalu $[-1, 1]$. V posledním kroku vyhodnotíme lokální maxima s využitím Signal Processing Toolboxu, který je součástí programového prostředí Matlab. DDK rychlost je počítána jako počet těchto lokálních maxim za sekundu a DDK pravidelnost jako jejich variance. Obrázek 1 ilustruje ukázkou algoritmu na vybraném řečovém signálu.



Obrázek 1: Ukázkou řešení DDK úlohy na zvoleném řečovém signálu. Šedé čáry zobrazují nalezená lokální maxima.

Důraz

Jak bylo zmíněno již v úvodu, snížení hlasitosti a melodie v promluvách může být jedním z nejčastějších příznaků zhoršení řeči u PN. Modifikace čteného textu o zvýšení intonačního důrazu na předem stanovených slovech může být problémovým vokalickým testem pro pacienty s výskytem PN a tak může být zároveň vhodnou úlohou pro zvýraznění rozdílů mezi pacienty s PN a zdravými mluvčími KS. V této úloze byla měřena variace relativní intenzity hlasitosti (E_{SD}) klouzavými průměry v okně o délce L vzorků

$$E[n] = 10 \log(x^2[n] + \dots + x^2[n-L+1]) / L, \quad (1)$$

a variace fundamentální frekvence v púltónech (F_0_{SD}) s využitím algoritmu [3].

Emoce

Další z nestandardních vokalických úloh vhodných pro hodnocení zhoršené řeči u PN je měření emocí. Účastníci jsou požádáni o co nejvěrnější napodobení emocí. Měření se skládá z deseti po sobě jdoucích vět v emocionálním kontextu: vzrušení, smutek, nervozita, znudění, zloba, hořkost,

strach, zklamání, překvapení a radost. Tyto po sobě jdoucí věty se mohou výrazně lišit v základní frekvenci hlasivek, ale také v prvních dvou formantových frekvencích, kde by se měla emoční intonace více projevit.

Variace fundamentální frekvence je měřena v půltónech (F0 SD) s využitím algoritmu [3]. Pro automatické vyhodnocení prvních dvou formantových frekvencí F1 a F2 bylo použito robustního trasování formantů s relativně vysokou odolností vůči šumu [13]. Tento algoritmus je založen na trasování formantů pomocí adaptivní banky filtrů s využitím preemfáze a Hilbertovy transformace. Více o algoritmu se lze také dočíst v [5]. Variace obou formantů (F1 SD, F2 SD) je měřena v půltónech.

Rytmus

Snížená schopnost vnímání rytmu může být dalším z deficitů PN. Pro jeho vyhodnocení bylo navrženo měření, kde byli účastníci požádáni o zopakování osmi veršů v co nejpodobnějším rytmu dle puštěné referenční nahrávky pořízené logopedem. Cílem měření pak je efektivně porovnat podobnost rytmu jednotlivce a nahrané reference. Pro porovnání je vhodné použít jeden z algoritmů sloužících pro rozpoznávání řeči. Avšak jedním z problémů při rozpoznávání řeči je, že přesné časování jednotlivých slov a trvání celé promluvy nebudou stejné. Vhodné řešení může být použití programovací techniky známé jako Dynamic Time Warping (DTW), aneb dynamické borcení času [14]. Tato metoda je používána k urovnání rozdílů v časování mezi řečovými úseky a referenčním vzorem. Základní princip DTW metody umožňuje rozložit tyto jednotlivé řečové úseky do prostoru jako 2D matici Euklidovských vzdáleností mezi dvěma body $d = (x - y)^2$. V této matici následně hledá mezi sousedními rohy optimální cestu, která maximalizuje lokální podobnost jednotlivých úseků promluvy. Délka (cena průchodu) této cesty je dobrým ukazatelem pro vyhodnocení podobnosti mezi dvěma promluvami. Nalezení této cesty je většinou řešeno metodou dynamického programování.

Ačkoliv byl DTW používán v mnoha doménách, může produkovat chyby. Jedním z případů je, když se algoritmus snaží vysvětlit variabilitu v ose y tím, že ji přiřadí k ose x . To může vést k celkově špatnému narovnání signálu z důvodu, že algoritmus přiřadí jeden bod z první časové sekvence k většímu úseku druhé časové sekvence. Další problémy u DTW mohou nastat v případě, když se průběhy částečně liší v ose y . Algoritmus jednoduše selhává při hledání zřejmých příznaků ve dvou sekvencích proto, že příznak v jedné sekvenci má mírně vyšší nebo nižší hodnotu než v sekvenci druhé, a výsledné přiřazení pak nemusí být správné. Oba problémy jsou vyřešeny modifikací algoritmu na Derivative Dynamic Time Warping (DDTW) [15]. U tohoto algoritmu počítáme opět Euklidovskou vzdálenost, ale ne z y hodnot bodů, ale z derivací v těchto bodech. Tím eliminujeme problém s odlišnostmi v bodech y , protože pracujeme přímo s průběhem funkce, což je to, co nás ve skutečnosti zajímá.

Měření rytmu je počítáno s využitím následujícího algoritmu:

a) Derivace v bodě vypočteme odhadem dle

$$D[x_i] = \frac{(x_i - x_{i-1}) + ((x_{i+1} - x_{i-1})/2)}{2}. \quad (2)$$

b) Vypočítáme krátkodobou Fourierovou transformaci (STFT) derivační sekvence D pro obě nahrávky s 25% procentním překryvem okna.

c) Sestavíme matici Euklidovských vzdáleností mezi jednotlivými body obou STFT.

d) Použijeme dynamické programování k nalezení nejlevnější cesty $\gamma[x, y]$ mezi opačnými rohy matice

$$\gamma[x, y] = d(x, y) + \min \left\{ \begin{array}{l} d(\gamma[x-1][y-1]) \\ d(\gamma[x-1][y]) \\ d(\gamma[x], [y-1]) \end{array} \right\}. \quad (3)$$

e) Použijeme celkovou cenu cesty normalizovanou celkovou cenou matice jako klasifikátor.

2.4 Statistika

Pro získání statisticky významných rozdílů mezi skupinami PN a KS bylo využito dvourozměrného neparametrického Wilcoxonova testu.

3 Výsledky

Pro všechna měření byla stanovena střední hodnota, směrodatná odchylka a statistická hladina významnosti p (viz tabulka 1). Obrázek 2 zobrazuje krabicové diagramy pro všechna statisticky významná měření.

Tabulka 1: VÝSLEDKY MĚŘENÍ PRO SKUPINY PN A KS SE STŘEDNÍ HODNOTOU (MEAN), SMĚRODATNOU ODCHYLKOU (SD) A STATISTICKOU VÝZNAMNOSTÍ (P).

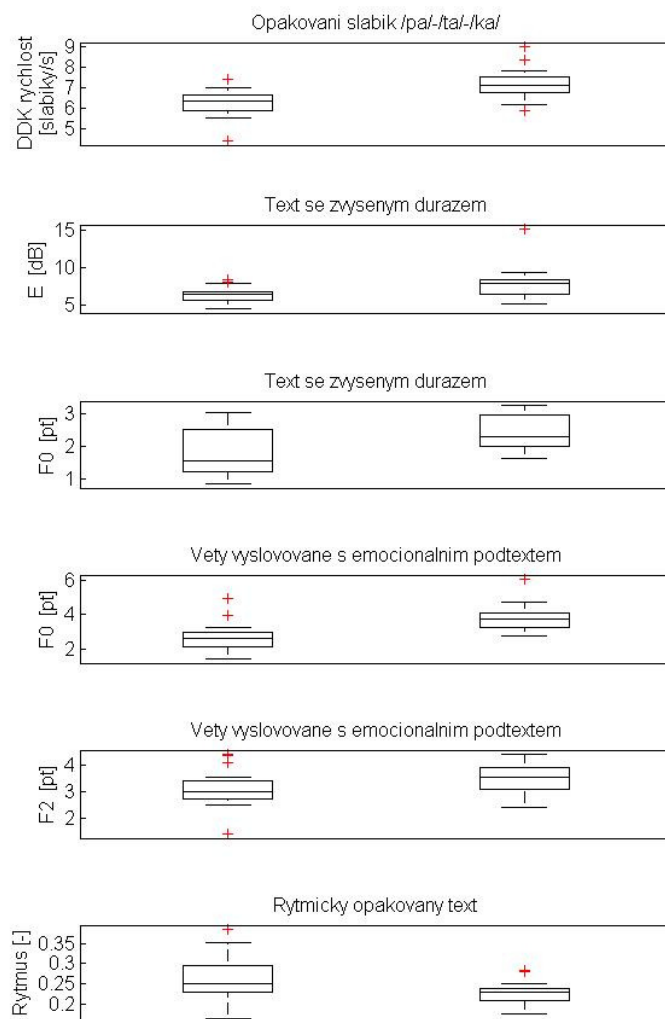
Měření	PN		KS		Statistika
	Mean	(SD)	Mean	(SD)	
<i>/pa-ta-ka/</i>					
DDK rychlost	6.22	0.63	7.16	0.73	$p < 0.001$
DDK pravidelnost	0.54	0.58	0.67	0.36	$p = 0.49$
<i>Text s důrazem</i>					
E SD	6.28	1.01	7.62	1.88	$p < 0.01$
F0 SD	1.70	0.69	2.35	0.59	$p < 0.01$
<i>Emoční věty</i>					
F0 SD	2.38	0.59	3.65	0.82	$p < 0.001$
F1 SD	3.10	0.65	2.94	0.42	$p = 0.43$
F2 SD	3.11	0.64	3.47	0.51	$p < 0.05$
<i>Rytmičkový text</i>					
Rytmus	2.65	0.55	2.27	0.28	$p < 0.01$

3.1 DDK úloha

Z výsledků lze konstatovat, že mezi oběma skupinami dochází ke statisticky významným rozdílům pouze u DDK rychlosti. U pacientů v raném stádiu PN nedochází ke sníženému vnímání pravidelnosti při rychlé artikulaci, ale může již docházet ke snížené schopnosti artikulace.

3.2 Důraz

Obě měření zvýšené intonace důrazu statisticky potvrdily rozdíl mezi skupinami PN a KS. Obrázek 3 vlevo znázorňuje, že při vzájemném poměru základní hlasivkové frekvence a intenzity hlasitosti ve většině případů dochází k oddělení pacientů s výskytem PN a zdravých mluvčích KS. Z toho vyplývá, že lidé s výskytem PN mají problémy mluvit zároveň melodicky i hlasitě, což může být způsobeno nejen sníženou motorikou, ale také zhoršeným vnímáním řeči.



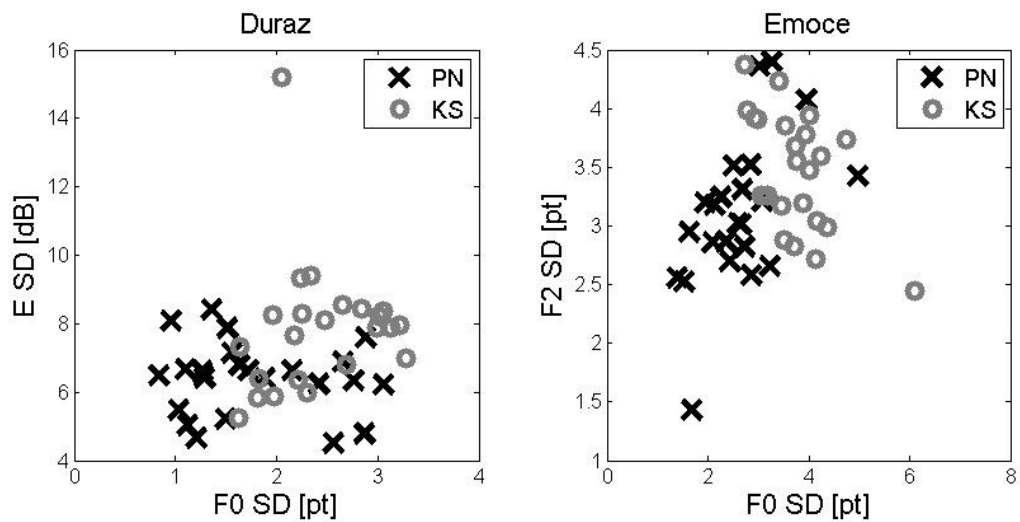
Obrázek 2: Krabicové diagramy pro statisticky významná měření u pacientů s výskytem PN a KS.

3.3 Emoce

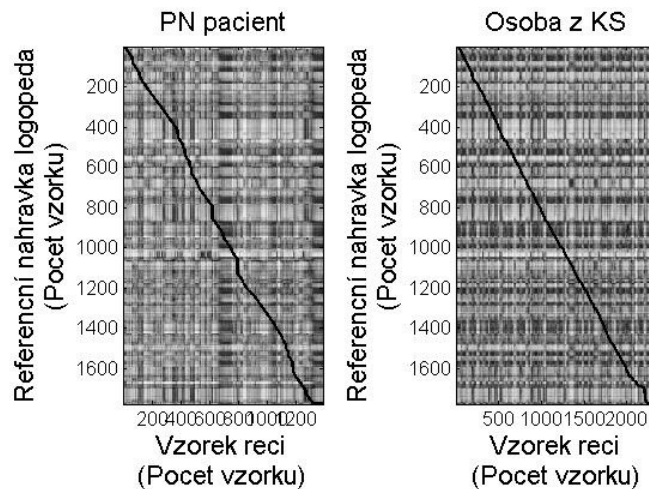
Statisticky významné rozdíly u vět vyslovovaných s emocionálním podtextem se ukázaly pouze v měřeních základní hlasivkové frekvence F0 SD a druhého formantu F2 SD. Významnější rozdíly mezi skupinami PN a KS u druhého formantu F2 mohou být způsobeny zhoršenými možnostmi pohybu jazyka, což je často asociovaným problémem u výskytu PN. Na obrázku 3 vpravo je ilustrována vzájemná závislost obou měření. Ve většině případů opět dochází k oddělení pacientů s výskytem PN a zdravých mluvčích KS.

3.4 Rytmus

U pacientů s výskytem PD se také projevila snížená schopnost napodobení příslušného rytmu. Obrázek 4 zobrazuje ukázkou vyhodnocení optimální cesty mezi referenční nahrávkou logopeda a jednoho účastníka z obou skupin. Měření dosáhlo statistické významnosti.



Obrázek 3: 2D zobrazení výsledků F0 SD a E SD textu se zvýšeným důrazem na předem stanovených slovech (vlevo); 2D zobrazení výsledků F0 SD a F2 SD vět vyslovovaných s emocionálním kontextem (vpravo).



Obrázek 4: Detaily měření rytmu s využitím DDTW algoritmu, použité při měření schopnosti opakovat příslušný text s určitým rytmem. Obrázky zobrazují matice Euklidovských vzdáleností mezi STFT dvou promluv s vyznačenou optimálně nalezenou cestou (černé čáry). Čím více je čára lineární, tím více se podařilo účastníkovi udržet stejný rytmus. Vertikální osy zobrazují nahrávku logopeda, horizontální osy zobrazují promluvu pacienta s výskytem PN (vlevo) a účastníka KS (vpravo).

4 Závěr

V této studii bylo analyzováno 23 českých pacientů v raném stádiu PN. V první části práce byl navržen algoritmus pro stanovení artikulační rychlosti a pravidelnosti. V navržené metodě využívající tento algoritmus bylo statisticky významné pouze měření artikulační rychlosti, což může být způsobeno zhoršenou motorikou řeči. Dále byla změřena základní hlasivková frekvence a intenzita hlasitosti u čteného textu se zvýšeným důrazem na předem stanovených slovech. Obě tyto měření dosáhly statistické významnosti. U vět vyslovovaných s emocionálním podtextem byla statisticky významná měření základní hlasivkové frekvence a druhého formantu. Z toho se dá usuzovat, že lidé s výskytem PN mají špatné vnímání melodičnosti řeči a zároveň zhoršené možnosti pohybu jazyka. Poslední část této práce se zabývala měřením schopnosti rytmického opakování textu. Pro jeho vyhodnocení byl navržen nový algoritmus na bázi DDTW. Toto měření bylo statisticky významné a potvrdilo, že u PN může docházet také ke zhoršené vnímavosti rytmu.

Výsledky studie, navržené vokalické testy a nové algoritmy mohou být použity k rozšíření objektivního hodnocení dysartrií nejen u PN, ale i u jiných neurodegenerativních onemocnění, které způsobují poruchu řeči. Díky jejich robustnosti mohou být také použity jako zpětná vazba při terapii řeči.

Poděkování

Tato práce je podporována primárně z grantů SGS 10/180/OHK3/2T/13 Hodnocení poruch hlasu a řeči, GAČR 102/08/H008 Analýza a modelování biomedicínských a řečových signálů, a dále z výzkumných záměrů MSM 6840770012 Transdisciplinární výzkum v oblasti biomedicínského inženýrství II, MSM 0021620849 Neuropsychiatrické aspekty neurodegenerativních onemocnění a MSM 0021620806 Molekulární biologie a patologie buňky za normy a u vybraných klinických závažných patologických procesů.

Autoři děkují MUDr. Janě Picmausové, MUDr. Veronice Majerové, MUDr. Jiřímu Klempířovi, PhD a Doc. MUDr. Janu Rothovi, CSc. za poskytnutí souboru klinických dat a prof. MUDr. Evženu Růžičkovi, DrSc. za koncepci a vedení klinické studie a za připomínky k rukopisu.

References

- [1] Canter, GJ.: Speech characteristic of patients with Parkinson's disease. I. Intensity, pitch, and duration, *J Speech Hear Disord*, 1963, 28:221-229.
- [2] Darley, FL., Aronson, AE., Brown, JR.: Differential diagnostic patterns of dysarthria, *J Speech Hear Res*, 1969, 12:426-496.
- [3] Rusz, J., Čmejla, R., Růžičková, H.: Analýza základní frekvence, amplitudového a frekvenčního kolísání hlasivek u Parkinsonovy nemoci, *Akustické listy*, 2009, 15(1):13-18.
- [4] Čmejla, R., Rusz, J., Růžičková, H.: Oslabení okluzí v promluvách pacientů s Parkinsonovou chorobou, *Akustické listy*, 2009, 15(3):12-18.
- [5] Rusz, J., Čmejla, R., Růžičková, H.: Formantová analýza a nové metody pro hodnocení míry artikulace u Parkinsonovy nemoci, *Akustické listy*, 2008, 15(2):3-8.
- [6] Rusz, J., Čmejla, R.: Analýza rychlosti řeči a intenzity u Parkinsonovy nemoci, *Akustické listy*, 2008, 14(2-4):13-16.
- [7] Mobes, J., Joppich, G., Stiebritz, F., Dengler, R., Schroder, Ch.: Emotional speech in Parkinson's disease, *Mov. Disord.*, 2008, 23:824-829.
- [8] Goberman, AM., Coelho, C.: Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy, *NeuroRehabilitation*, 2002, 237-246.
- [9] Rosen, KM., Kent, RD., Delaney, AL.: Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers, *Journal of Speech, Language, and Hearing Research*, 2006, 49:395-4.
- [10] D'Alatri, L., Paludetti, G., Contarino MF., et al.: Effects of Bilateral Subthalamic Nucleus Stimulation and Medication on Parkinsonian Speech Impairment, *J Voice*, 2006.
- [11] Tjaden, K., Watling, E.: Characteristics of Diadochokinesis in Multiple Sclerosis and Parkinson's disease. *Folia Phoniatica et Logopaedica*, 2003, 55:241-259.
- [12] Midi, I., Dogan, M., Koseoglu, M., et al.: Voice abnormalities and their relation with motor dysfunction in Parkinson's disease, *Acta Neurol Scand*, 2008, 117:26-34.

[13] Mustafa, K., Bruce, IC.: Robust formant tracking for continuous speech with speaker variability, IEEE Transactions on Speech and Audio Processing, 2006.

[14] Turetsky, R., Ellis, D.: Ground-Truth Transcriptions of Real Music from Force-Aligned MIDI Syntheses, 4th International Symposium on Music Information Retrieval ISMIR-03, Baltimore, 2003, 135-141.

[15] Keogh, EJ., Pazzani, MJ.: Derivative Dynamic Time Warping, First SIAM International Conference on Data Mining, Chivago, IL, 2001.

Jan Rusz

České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
Univerzita Karlova v Praze, Laboratoř biokybernetiky a počítačové podpory výuky,
Ústav patologické fyziologie 1. LF UK
e-mail: ruszjan@fel.cvut.cz

Roman Čmejla

České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra teorie obvodů
e-mail: cmejla@fel.cvut.cz

Hana Růžičková

Univerzita Karlova v Praze, Neurologická klinika 1. LF UK a VFN
e-mail: hruci@seznam.cz