

AKUSTICKÉ LISTY

České akustické společnosti
www.czakustika.cz

ročník 15, číslo 1

březen 2009

Obsah

Usnesení Valné hromady České akustické společnosti	3
Tisková zpráva k Mezinárodnímu dni proti hluku na Zelený čtvrtek 9. dubna 2009 <i>Jan Stěnička</i>	4
Úlohy a algoritmy hlasových technologií Tasks and Algorithms in Voice Technologies <i>Jan Uhlíř</i>	5
Analýza základní frekvence, amplitudového a frekvenčního kolísání hlasivek u Parkinsonovy nemoci Analysis of Fundamental Frequency and Amplitude Instability of the Vocal Cords in Parkinson's Disease <i>Jan Ruzs, Roman Čmejla a Hana Růžicková</i>	13
Pozvánka na seminář „Pružné uložení budov“	18

ČESKÁ AKUSTICKÁ SPOLEČNOST

Usnesení Valné hromady České akustické společnosti, konané dne 29. ledna 2009 v prostorech Fakulty elektrotechnické ČVUT

Valná hromada ČsAS bere na vědomí:

1. zprávu o činnosti Rady ČsAS;
2. zprávy o činnosti jednotlivých odborných skupin a o jejich dalším zaměření;
3. zprávu o přípravě 78. akustického semináře;
4. zprávu o výsledcích revize hospodaření společnosti;
5. výsledky voleb do Rady společnosti a výsledky voleb předsedů odborných skupin;
6. zprávu o usnášéníschopnosti Valné hromady, 51 / 101;
7. informace o plnění úkolů a poslání Akustických listů a vyzývá členy k zaslání příspěvků.

Pro funkční období roku 2009 byli v jednotlivých odborných skupinách zvoleni:

- A. Obecná, lineární a nelineární akustika
předseda – M. BEDNAŘÍK zástupce – M. ČERVENKA
- B. Ultrazvuk a akustické emise
předseda – J. PLOCEK zástupce – R. BÁLEK
- C. Hluk a vibrace
předseda – J. KOZÁK zástupce – V. STRNAD
- D. Prostorová, stavební a urbanistická akustika
předseda – M. MELLER zástupce – A. RYNDOVÁ
- E. Zpracování a záznam akustických signálů
předseda – T. SALAVA
- F. Psychoakustika, fyziologická akustika a akustika hudby a řeči
předseda – volba bude provedena dodatečně
- G. Elektroakustika
předseda – Z. KEŠNER zástupce – B. SÝKORA

Valná hromada ČsAS schvaluje:

1. zprávu o činnosti Rady za kalendářní rok 2008 a uděluje Radě absolutorium;
2. zprávu o hospodaření společnosti za kalendářní rok 2008;
3. výši členských příspěvků na rok 2009 (400 Kč pro členy, 100 Kč pro studenty a důchodce);
4. činnost Rady a odborných skupin v roce 2008.

Valná hromada ukládá nově zvolené Radě společnosti na kalendářní rok 2009:

1. pokračovat v odborné a organizační činnosti i v zahraničních kontaktech, v rozvíjení spolupráce s Českou matricí technickou, Slovenskou akustickou společností, společností Elektra, českou sekci AES, Českým centrem IEE, československou sekci IEEE, ICA, EAA a I-INCE;
2. věnovat pozornost pořádání odborných akcí a pravidelných seminářů odborných skupin;
3. nadále rozvíjet vydávání Akustických listů.

Valná hromada ukládá nově zvoleným předsedům odborných skupin na kalendářní rok 2009:

1. publikovat informace o připravovaných aktivitách skupin v Akustických listech a na webové stránce s předstihem tak, aby se zájemci mohli včas na akce přihlašovat.

Valná hromada doporučuje Radě ČsAS:

1. pravidelně se zabývat činností a plánem akcí odborných skupin;
2. pravidelně se zabývat plánem a zaměřením konaných akustických konferencí;
3. poskytovat možnost finančních výhod členům společnosti, např. nižšími sazbami vložného na akcích pořádaných společností.

Výsledky voleb do Rady České akustické společnosti:

Počet vydaných platných hlasovacích lístků:	51	
Odevzdáno platných hlasovacích lístků:	50	
předseda:	O. JIŘÍČEK	50
místopředseda:	V. KUNZL	50
sekretář:	M. BROTHÁNEK	50
hospodář:	O. KUDĚJOVÁ	50
revizní komise:	J. NOVÁK	49
	T. HELLMUTH	49
	J. PLOCEK	49

Návrh usnesení sestavila návrhová komise ve složení M. Meller a P. Urban.

Vážení kolegové,

jak se můžete dočíst v usnesení Valné hromady otiskném v tomto čísle Akustických listů, i v letošním roce zůstávají členské příspěvky nezměněny. Pro výdělečně činné členy tedy činí 400 Kč a pro důchodce a studenty 100 Kč. Pokud jste ještě nezaplatili, je nejlepším způsobem platba převodem z účtu na účet, tedy z Vašeho účtu v kterékoli bance na účet ČsAS, Komerční banka, a. s., Praha 6, číslo účtu 17838061/0100. Platbu můžete provést ze svého soukromého účtu, který nese název Vašeho příjmení. V případě, že platíte z jiného účtu, použijte prosím variabilní symbol ve tvaru XXX09, kde XXX je Vaše osobní číslo, které naleznete nad svým jménem na obálce. Dalším možným způsobem je složení hotovosti na přepážce Komerční banky. Na kterékoli pobočce Komerční banky můžete zaplatit své členské příspěvky složením hotovosti na účet číslo 17838061/0100. Variabilní symbol je shodný jako v předešlém případě. Finančně oboustranně nejméně výhodným způsobem je platba na poště složenkou typu A (zelená). I zde je variabilní symbol shodný jako v předešlých případech.

Marek Brothánek
sekretář společnosti

Tisková zpráva k Mezinárodnímu dni proti hluku na Zelený čtvrtek 9. dubna 2009

Jako každý rok, tak i letos na velikonoční Zelený čtvrtek, který je 9. dubna, ohlašuje Hudebně ekologické sdružení HUDEKOS při České hudební společnosti v návaznosti na mezinárodní aktivity

International Noise Awareness Day – Den uvědomění si hluku.

Celoroční snažení naší společnosti je zaměřeno na sdělování informací o působení hluku na populaci. Neustálé zvyšování tzv. akustického smogu jako součtu všech druhů hluku způsobuje nervozitu, snižuje imunitu vůči všem druhům nemocí a může vést k trvalému poškození zdraví. Kromě hluku z výroby a dopravy je to v naší epoše především jeden zbytečný hudební hluk. Je to hudba z rádií, rozhlasů, televize, reklam v obchodech apod., která je nám vnucována jako nechtěný poslech hudby na veřejnosti i v soukromí. Stále platí definice hluku: je to souhrn všech zvuků, které nás obtěžují. Ticho se stalo luxusem, o kterém nevědomí lidé ani nevědí.

Snižování hluku stojí velké peníze – kromě jednoho způsobu, který my doporučujeme. Hluk si musíme uvědomovat, a pokud nás reklama nezajímá nebo obtěžuje, tak můžeme hluk odstranit, a to ihned a zadarmo – vypnutím přístroje, odchodem z místa, změnou trasy.

Na rozdíl od zahraničí je u nás hluk zakotven v legislativě jednoho rezortu, Ministerstva zdravotnictví ČR. Je to zákon číslo 258/2000 Sb. „O ochraně veřejného zdraví“, ve znění pozdějších předpisů, kterým je zejména nařízení vlády číslo 148/2006, kde najdete hygienické limity hluku a jak s nimi nakládat. Dne 9. dubna 2009 chceme využít právo člověka podle Listiny základních práv a svobod na vypnutí, nebo alespoň snížení intenzity hluku hudby z reklam na veřejnosti.

Členové sdružení HUDEKOS – akustici, lékaři, učitelé, muzikanti, psychologové – jsou připraveni poskytnout podklady pro novináře, přednášet a zúčastnit se besed o hluku, pomáhat pedagogům, radit, jak se lze chránit před hlukem, dávat podněty a připomínkovat zákony a vyhlášky upravující životní prostředí a každý veřejný prostor. Chceme vyzvat ke spolupráci všechny, kdo s námi souhlasí, a těšíme se na spolupráci nejen v Den uvědomění si hluku. Uvědomit si hluk znamená psychicky se vyztužit proti němu, a proto vznikl International Noise Awareness Day. Připomínáme, že podle křesťanských tradic na Zelený čtvrtek odletěly zvony do Říma a bylo ticho. Snažme se v tento den 9. dubna 2009 o ticho a uijme si klid a krásu tohoto dne.

Za sdružení HUDEKOS

Jan Stěnička
www.hudekos.info

Úlohy a algoritmy hlasových technologií

Jan Uhlíř

ČVUT–FEL, Technická 2, 166 27 Praha 6

e-mail: uhliř@fel.cvut.cz

The article briefly describes digital technologies that are used in the process of speech communication. Digital voice processing is applied in telecommunications, as well as in man-machine communication.

1. Komunikace a její technické zázemí

Život v lidské společnosti je závislý na schopnosti jednotlivců komunikovat, vzájemně si sdělovat informace. Informace v abstraktní formě zpracovává lidský mozek a v okamžicích, kdy dospěje k potřebě informace sdílet s jinými lidmi, formuluje s využitím jazykových prostředků sdělení adresované jinému člověku nebo skupině lidí – dochází ke komunikaci. Prostředky vzájemné komunikace můžeme v současnosti charakterizovat dvěma cestami, hlasovou komunikací – artikulovanou lidskou řečí a písemnou komunikací. Obě tyto cesty jsou vázány na určitý (národní) jazyk. Komunikaci umožňují i další prostředky neverbální a mimojazykové, jako jsou neartikulované hlasové projevy a motorika. Můžeme sem zahrnout i kódování informace do vizuální podoby, jako jsou kouřové signály, znaková řeč neslyšících a indikátory na panelech přístrojů. Komunikace s přístroji a technologickými zařízeními je dosud většinou omezena na mechanické ovládací prvky – páky, tlačítka, klávesnice a obrazové a textové displeje apod. Je přirozenou snahou konstruktérů technických zařízení zlepšovat podmínky pro jejich řízení, resp. komunikaci s nimi. Hlasová komunikace tedy není již jen záležitostí sdělování mezi lidmi, ale bude se stále více stávat též součástí komunikace člověk – stroj.

Artikulace a percepce. Pro úvahy o technickém řešení systémů podporujících hlasové komunikace je potřebné poznat fyzikální podstatu vytváření (artikulace), přenosu a percepce řečového signálu.

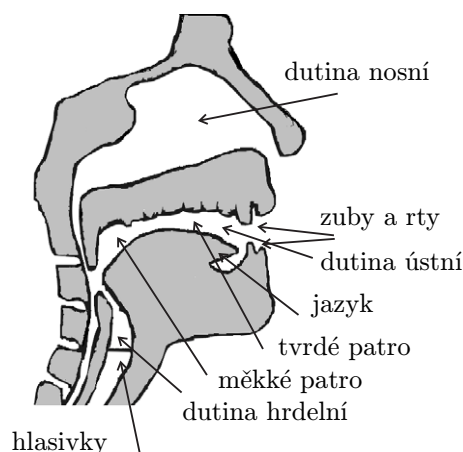
Řečový signál generuje artikulační aparát. Artikulace je založena na lidskou vůlí řízených změnách akustických vlastností hlasového traktu. V něm vzniká akustická vlna, která se od mluvčího šíří volným prostorem. Primárním zdrojem energie této akustické vlny je proud vzduchu vyhaněný z plic. Veškeré artikulované projevy v češtině jsou vytvářeny výdechovým proudem vzduchu (i když lze slyšitelné zvuky vytvářet i při nádechu). Proud vzduchu vyhaněný z plic vytváří slyšitelné artikulované zvuky dvěma základními principy.

Znělé úseky promluvy jsou vytvářeny tak, že proud vzduchu prochází sevřenými hlasivkami, které vibrují a tak vytvářejí sled impulsů, vstupujících do dutin hrdelní, ústní a nosní. Tyto dutiny se chovají jako rezonátory s různými vlastními rezonančními vlastnostmi (rezonanční frekvence, činitel jakosti), určenými jejich tvarem a veli-

kostí. Znělou promluvu tedy člověk artikuluje tím, že svalovou činností mění tvar uvedených rezonátorů.

Neznělé úseky promluvy vznikají bez účasti hlasivek. Výdechový proud vzduchu je v hlasovém traktu ovládán tak, že v určitých místech nastavené překážky vytvoří slyšitelné turbulence, jejichž akustický projev lze charakterizovat jako širokopásmový šum, který může být více či méně ovlivněn ve svých výsledných vlastnostech průchodem uvedenými dutinami. Jiné segmenty řeči vznikají přerušováním hlasivkami modulovaného, či nemodulovaného vzduchového proudu jazykem nebo rty.

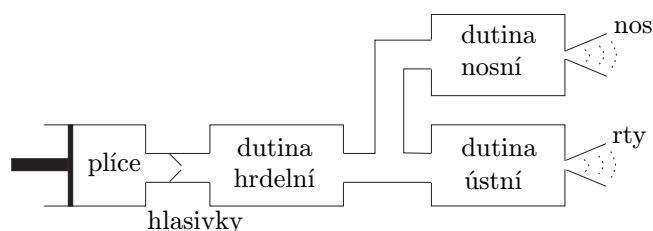
Průřez hlasovým ústrojím člověka ukazuje obr. 1.



Obrázek 1: Artikulační orgány

Na dalším obrázku je velmi zjednodušený akustický model artikulačního ústrojí. V průběhu artikulace se významně mění vlastnosti zobrazených dutin. Existuje rozsáhlá literatura věnovaná akustickému modelu artikulace, ve kterém se uplatňují nejrůznější geometrické charakteristiky uvedených rezonátorů.

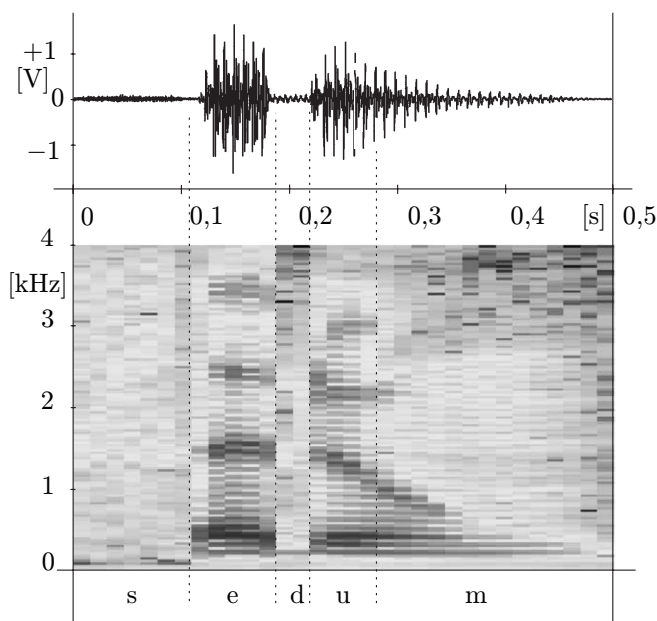
Hlasovým ústrojím modulovaný proud vzduchu vystupuje přes rty a nosem do vnějšího prostředí. Ve vzduchu se vytvoří zvuková vlna, kterou zachytí sluchové ústrojí posluchače. Neméně významné je, že mluvčí svým sluchem kontroluje vlastní artikulaci a na principu zpětné vazby ji koriguje do podoby, kterou považuje za odpovídající zamýšlenému sdělení. Má-li být sdělení přeneseno na velkou vzdálenost, nebo jakkoli technicky zpracováno, musí být akustická vlna převedena na elektrický signál mikrofonom.



Obrázek 2: Akustický model artikulačního aparátu

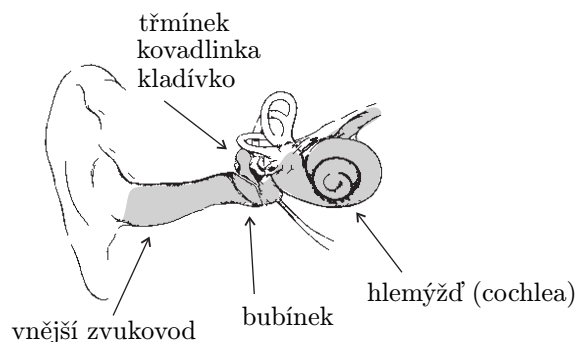
Časový průběh hlasového signálu odpovídá časovému průběhu změn akustického tlaku ve vzduchu a posléze i časovému průběhu výchylky bubínku ve sluchovém ústrojí. Aby elektrický signál z mikrofону mohl vytvořit slyšitelný zvuk, musí být zaveden do reproduktoru nebo sluchátka, které vytvoří odpovídající akustickou vlnu.

Věnujme se nejprve některým vlastnostem řečového signálu, který můžeme digitalizovat a analyzovat prostředky výpočetní techniky. Povšimněme si, že v časovém průběhu signálu snadno odhalíme důsledky výše popsaných mechanismů artikulace. Na obr. 3 je časový průběh signálu slova *sedum*. Počátek slova je tvořen neznělým šumovým úsekem, na který navazuje znělý úsek, dále je patrná exploze (uvolnění hlasivkami modulovaného proudu vzduchu jazykem původně opřeným o tvrdé patro) a koncový znělý úsek. Šumový úsek můžeme skutečně pozorovat jako náhodný signál s relativně malým rozkmitem. Naproti tomu ve znělém úseku jsou dobře patrné hlasivkové impulsy a odezvy tlumených rezonátorů. Signál je zjevně nestacionární a informaci nesou jeho měnící se okamžité vlastnosti.

Obrázek 3: Tvar vlny a krátkodobé spektrum slova *sedum*

Zpracování zvukového vjemu (percepce) je ve sluchovém ústrojí založeno na převodu mechanických pohybů

bubínku na podráždění nervových zakončení nervů vedoucích ze sluchového ústrojí do mozku (obr. 4).



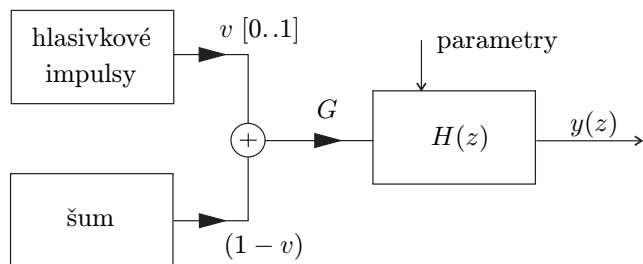
Obrázek 4: Ucho (popsány jsou jen části sloužící sluchu)

Převodní mechanismus výchylek bubínku je „zkonstruován“ z kladívka, kovádlínky a třmínku tak, že je sluchový orgán schopen zprostředkovat poslech ve velmi širokém dynamickém rozpětí hlasitosti. V pojmech z oblasti zpracování signálů jde o kompresi, příp. normalizaci amplitudy signálu. Nejpodstatnějším zjištěním však je, že mechanické kmity jsou analyzovány v hlemýždi, který má na obvodu zužujícího se profilu své dutiny ohromné množství nervových zakončení. Z toho lze usuzovat, že prvotní informace vedená do mozku je založena na vyhodnocení energie signálu ve frekvenčním spektru, protože stavba hlemýžďe (jeho zužující se profil, v němž vibruje kapalina) umožňuje rozlišit frekvenční složky a aktivovat v závislosti na rozkmitu v jeho jednotlivých místech příslušná nervová zakončení. Dostáváme tak z oblasti fyziologie sluchu argument pro to, abychom se při zpracování řečových signálů zabývali kromě časového průběhu signálu také časovým průběhem krátkodobého (v čase se měnícího – tekoucího) spektra. Proto je na obr. 3 uveden spektrogram, tj. graficky vyznačený časový průběh velikosti spektrálních složek signálu, ve kterém jsou amplitudy spektrálních složek vyznačeny stupněm šedé (čím tmavší pole, tím větší amplituda). Snadno odhalíme souvislosti spektrálního obrazu signálu s mechanismem artikulace jeho jednotlivých částí. Modelům percepce řeči a slyšitelných zvuků vůbec jsou věnovány rozsáhlé publikace, které se zabývají vystižením nejružnějších specifík sluchu. Od těchto modelů vede cesta k velmi efektivním algoritmům komprese zvukových signálů (např. kódování MP3).¹

Ukázali jsme, že artikulace představuje transformaci sdělení na řečový signál, který přenášenou informaci repre-

¹ Za určitý doklad toho, že lidský mozek pracuje se spektrální analýzou hlasového signálu, můžeme považovat skutečnost, že kochleární implantát (elektronické zařízení, které nahrazuje neslyšícím spojení z vnějšího ucha do hlemýžďe) je zkonstruován tak, že do hlemýžďe je zavedena řada elektrod končících v jeho různé hloubce. Jimi jsou stimulována nervová zakončení podél hlemýžďe. Stimulační elektrický signál je odvozen z hlasového signálu signálovým procesorem tak, že jednotlivé elektrody dostávají impulsy s amplitudou závislou na okamžitém rozložení energie ve spektru řečového signálu. Lidem s kochleárním implantátem je tak umožněno slyšet a komunikovat hlasem.

zentuje měnicím se rozložením amplitud složek krátkodobého spektra, a že příjemce takto zakódovanou informaci interpretuje právě s využitím spektrální analýzy. Toto zjištění dovoluje navrhnout pro syntézu hlasového signálu číslicový model artikule uvedené na obr. 5. Hlasivkový tón vytváří generátor impulsů, a to pro znělé hlásky. Neznělé hlásky jsou generovány z bílého šumu. Plynulý přechod mezi znělým a neznělým buzením, kdy se oba signály mísí, je reprezentován v čase se měnícím parametrem v . Hlasitost umělé promluvy určuje multiplikativní činitel G . Mění se spektrální složení výsledného zvuku určuje přenosová funkce průběžně „přeladovaného“ číslicového filtru $H(z)$. Struktura uvedeného filtru je předmětem rozsáhlých, dosud neuzavřených výzkumů. V dalším textu poznáme, že parametry tohoto filtru mohou být významné nejen pro věrnou resyntézu hlasu, ale že mohou sloužit i v systémech rozpoznávání promluv a mluvčích. Pro ozvučení by musel být na výstup filtru umístěn D/A převodník a reproduktor.



Obrázek 5: Číslicový model artikulačního aparátu

2. Řečový signál

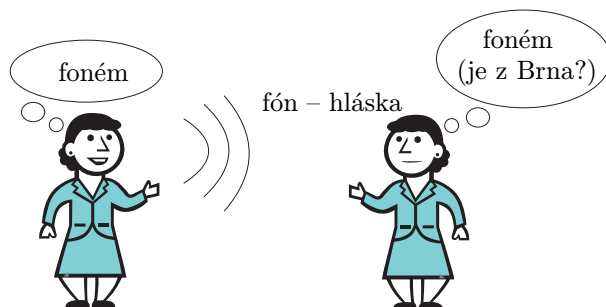
Vzájemná komunikace pomocí zvukových signálů je v přírodě běžná i mezi živočichy. Člověk zdokonalil tuto komunikaci tak, jak to vyžaduje a umožňuje jeho intelekt a potřeba vzájemného sdělování informací. Lidé navíc umějí zachytit a prezentovat informace pro své okolí psaným textem.² Oba způsoby sdělování mají společné to, že informace je nesena řetězem elementů, u psaného textu ji nesou zřetěžené grafémy – písmena, u mluvené řeči jsou zřetězeny fonémy. V našem stručném úvodu pomineme, že existují jazyky zachycující graficky celá slova. Řetězení nejrozličnějších kombinací elementů řeči dovoluje vytvořit řetězce s nekonečnými možnostmi variací, a to přesto, že inventář takových elementů je co do jejich počtu omezený. To, že

²Každý napsaný text je možno přečíst, příp. mu porozumět a vyslovit ho jako promluvu. Není tomu naopak, protože hlasový projev může mnohdy nést informace, které nelze, nebo jen s obtížemi lze zachytit psaným textem. Tuto „chudobu“ psaného textu dokážou překlenout umělci při interpretaci divadelních rolí, nebo při recitaci. Bezděčně ji překlenu v jednoduchých případech i každý, kdo hlasitě čte napsaný text. Rovněž se jí snaží překlenout konstruktivní systémů syntetizujících hlas z textu, a to tím, že text analyzují a podle lingvistických pravidel zvukovou podobu promluvy modifikují tak, aby projev stroje byl bližší projevu lidskému. Na druhé straně mohou rysy přirozené artikulované promluvy nezachytitelné písmem usnadnit činnost automatického rozpoznávající řeči.

jde o řetězení fonetických elementů řeči, ukazuje, že řečový signál je jednorozměrný a informaci, kterou přenáší, lze identifikovat jen analýzou jeho časového vývoje, v digitalizovaném signálu potom analýzou časové řady jeho vzorků.

2.1. Fonetické segmenty – hlásky

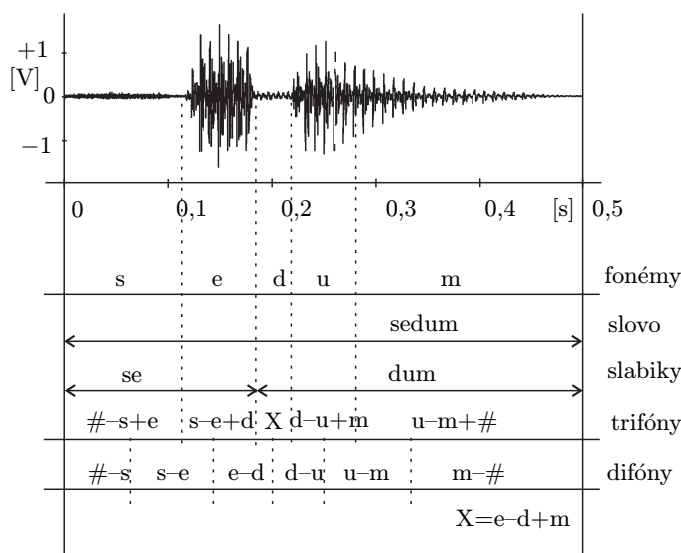
Již v úvodu tohoto odstavce musíme poznamenat, že zvukovou podobu řeči zachycují fonémy pouze formálně. Jejich konkrétní zvuková podoba vyžaduje pro svou reprezentaci vytvoření inventáře elementů, který zohlední skutečné chování fonémů v řetězci tvořícím konkrétní promluvu. Tuto odlišnost popisuje jazykovědný obor fonetika. [1] Fonémy jsou definovány jako elementy, které svou sekvencí identifikují zvukovou podobu každého slova. Inventář fonémů je blízký inventáři grafémů a nezohledňuje variabilitu artikulace fonémů v jejich různých kombinacích (sloves) a u různých mluvčích. Konkrétní zvukovou podobu vysloveného fonému označujeme jako hlásku nebo, s použitím anglického termínu, jako fón. Tento vztah abstraktního reprezentanta hlasového elementu řeči a jeho konkrétní artikulované podoby ukazuje obr. 6. Obrázek ukazuje, že v procesu percepce dochází k interpretaci hlásky (třeba velmi specificky vyslovené) fonémem, protože k porozumění potřebuje mozek pracovat se zobecněným inventářem řečových elementů (a specifiku výslovnosti použije pro jiný účel, např. pro hodnocení osobnosti mluvčího). V odborné literatuře a v diskusích odborníků se však velmi často setkáme se zaměňováním, nebo úplným pominutím tohoto terminologického odlišení pojmů a jako fonémy jsou mnohdy označovány hlásky.



Obrázek 6: Foném a fón

Vraťme se k obrázku 3, kde je v podtitulku uveden zápis slova *sedum* pomocí grafémů, tak jak to odpovídá umístění slova v psaném textu a jeho pravopisu. V obrázku je zachyceno zřetězení fonémů, které respektuje obvyklou fonetickou realizaci. Kdybychom však podobně zapsali jiná slova a zkoumali časový průběh signálu nebo jeho spektra, zjistili bychom, že téměř fonému *e* nebo *m* bude odpovídat jiný signál. Fónů, tj. konkrétních signálů odpovídajících fonémům, bude ohromné množství. To je pro rozpoznávání i syntézu řeči problém, který je nutno řešit hledáním řečových elementů, které mají v různých slovech určité opakující se rysy, avšak zachytí odchylky v realizaci

fonémů v různých kontextech. Takto vytvořený kompromis vede na inventář elementů, které jsou specifitější než fonémy, ale redukuje počet fónů. Některé z možností, používaných při volbě inventáře řečových elementů, ukazuje obr. 7.



Obrázek 7: Elementy řečového signálu

Nejpřesnější zachycení specifiky výslovnosti fonémů lze očekávat tam, kde je promluva zachycena v co nejdelším kontextu. Proto jsou rozpoznávače pevně stanovených promluv úspěšné a proto i syntetizéry řeči založené na reprodukci přirozených promluv považujeme za kvalitní. Ovšem všech možných promluv, které lze ve vzájemné komunikaci pronést, je nekonečně mnoho. Tudíž je reprezentace řečového signálu tímto způsobem omezena na velmi malý okruh aplikací. V obrázku jsme naznačili, že za element promluvy lze považovat slovo. Okruh aplikací, kde by taková reprezentace hlasového signálu mohla najít uplatnění, je významně větší a uvidíme, že jak v syntetizérech, tak v rozpoznávačích může být užitečná. Ještě univerzálnější by mohly být slabiky. Jejich problémem však je výskyt slov, u kterých se hranice slabik nedají určit jednoznačně. Kromě toho by byl inventář slabik ještě stále velmi rozsáhlý, pokud by měly být kdykoli použitelnými reprezentanty hlasových elementů. Proto se v praxi téměř nepoužívají. Za dostatečně univerzální, a co do rozsahu přijatelné, lze považovat inventáře, jejichž základem je difón nebo trifón. Prakticky významných difónů je nutno zavést cca 500 a trifónů několik desítek tisíc, což je reálné.

Poznamenejme, že obrázek ořezuje, proč se tak často setkáváme s pojmem „kontextově závislý foném“, namísto správnějšího označení „trifón“. Z hlediska výše zmíněné definice takový foném nemá smysl. V mysli netvoříme promluvu s úvahou, která by jakkoli stála na potřebě hledět na sousedství zřetězených fonémů, sousedství fonémů neidentifikuje slovo.

2.2. Charakteristiky hlásek

Zavedením kontextově závislých řečových elementů jsme podpořili možnost reprezentace zvukové stránky promluvy řetězem elementů z omezeného, i když v některých případech velmi rozsáhlého inventáře.

Charakteristikou hlásek se nejprve zabývala lingvistika ve své fonetické specializaci. Založila třídění hlásek na popisu jejich tvorby v artikulačním ústrojí. Jistě takovým popisem nelze vystihnout všechny varianty hlásek. Většinou jde o dělení nezávislé na kontextu, pokud kontext přímo neurčuje výraznou odchylku od nastavení artikulačních orgánů. Této charakteristice se velmi stručně věnujeme. Je užitečná i v technických aplikacích, protože odkaz na způsob artikulace je současně odkazem na určité typické akustické, a tedy i spektrální charakteristiky.

Tabulka 1 ukazuje jednoduché dělení českých samohlásek a souhlásek odkazující na mechanismus artikulace. V [1] se setkáme s tříděním hlásek, které ještě důkladněji, než náš zjednodušený systém, charakterizuje hlásky podle artikulace.

poloha jazyka	přední	střední	zadní
vysoká	i í		u ú
středová	e é		o ó
nízká		a á	
postavení rtů	nezaokrouhlené		zaokrouhlené

znělost/artikulace	závěrové	úžinové	polozávěrové
párové znělé	b d d' g	z ž v h ř	dz dž
párové neznlé	p t t' k	s š f ch ř	c č
nepárové znělé	m n ň	l r j	

Tabulka 1: České hlásky podle artikulace

V horní tabulce je ukázáno roztrídění samohlásek. Samohlásky jsou všechny znělé a při jejich tvorbě není zúčastněna turbulence ve výdechovém proudu vzduchu. Hlasivkové impulsy vstupují do rezonátorů, jejichž tvar určuje hlavně poloha jazyka. Výsledný spektrální obraz samohlásek navíc ovlivňuje nastavení rtů v okamžiku jejich vyslovení. Ukážeme později, že pro odlišení samohlásek jsou nejvýznamnější informací hodnoty tzv. formantů, resp. formantových kmitočtů, které jsou rezonančními kmitočty hlavních dutin vytvořených jazykem a dalšími orgány v hlasovém traktu. Hlasivkové impulsy budí dutinové rezonátory signálem s širokým spektrem, a ty reagují tlumenými kmitý, skládajícími se do výsledného „zvlnění“ proudu vzduchu vystupujícího mezi rty a částečně nosem.

Mezi samohlásky je řazen rovněž neutrální zvuk označovaný „schwa“, který se vyskytuje v různých jazycích. Jde o hlásku středovou, střední, nezaokrouhlenou. V češtině se nevyskytuje. Vyslovíme ji za souhláskou, když hlásku jednoduše písmena abecedy (souhlásky), např. b, c, d (nikoliv jejich názvy bé, cé, dé).

V dolní tabulce je zjednodušený systém souhlásek. Souhlásky mohou být znělé i neznělé. Jejich artikulace se liší podle toho, jak zasahují součástí ústní dutiny do proudu vzduchu z plic, a to buď plynulého (neznělé souhlásky), nebo zvlněného hlasivkovými impulsy (znělé souhlásky).

Tyto zásahy jsou v podstatě dvojí. První typ souhlásek se označuje jako závěrový, což znamená, že hláska je vytvořena dočasným uzávěrem nastaveným do cesty vycházejícímu proudu vzduchu a následným uvolněním tlaku, který se před překážkou vytvořil. Časový vývoj akustické podoby se skládá ze dvou částí – okluze (uzavřený stav) a exploze (uvolnění). Uzávěr průchodu vzduchu může vytvořit v ústech různě opřený jazyk nebo rty. Typ uzávěru při artikulaci závěrových souhlásek (retné, zuboretné, patrové) je dále dělí do menších skupin [1].

Druhý typ souhlásek je charakteristický souvislým průchodem vzduchu skrze hrtan a ústní a nosní dutinu. Artikulaci určující části úst pouze vytvářejí zúžení, ve kterých dochází ke tření vzduchu o překážku a následné turbulenci. Proto hlásky úžinové. Místo, ve kterém je zúžení vytvořeno, určuje, jak velkou částí artikulačního ústrojí bude turbulentní proud vzduchu procházet. Bude-li zúžení až v blízkosti rtů, nezasáhne selektivní účinky dutin do spektra signálu a úžinová souhláska bude mít málo zvlněné širokopásmové spektrum. V [1] nalezneme další dělení úžinových souhlásek (opět např. patrové, zuboretné, ...).

Kombinací obou mechanismů se vyznačují hlásky polo-závěrové, kde za zmínku stojí znělé hlásky označené *dz* a *dž*. Jedná se o popis specifické artikulace úžinového *ž* a *z* v kombinaci se závěrovým *d*, např. ve slově *džbán* (cítime, že *ž* ve slově *žába* je „jiné“ než ve slově *džbán*, *džber* apod.). Takových specifických typů hlásek lze při podrobnějším zkoumání najít ještě celou řadu (dvojhásky, znělé a neznělé *ř* apod.) [1].

3. Systémy hlasových technologií

3.1. Analogové a digitální přenosové prostředky

Přenos mezi mluvčím a posluchačem, bez podpory technickými prostředky, lze jistě považovat za přenos analogový. Průběh akustického tlaku ve vzduchu je spojitý jak v čase, tak v okamžitých hodnotách. Elektrický signál z mikrofonu je obrazem časového průběhu akustického tlaku (někdy jeho derivace), tedy je rovněž analogový. Sluchátko i reproduktor jsou rovněž zařízení určená pro zpracování analogového signálu.³

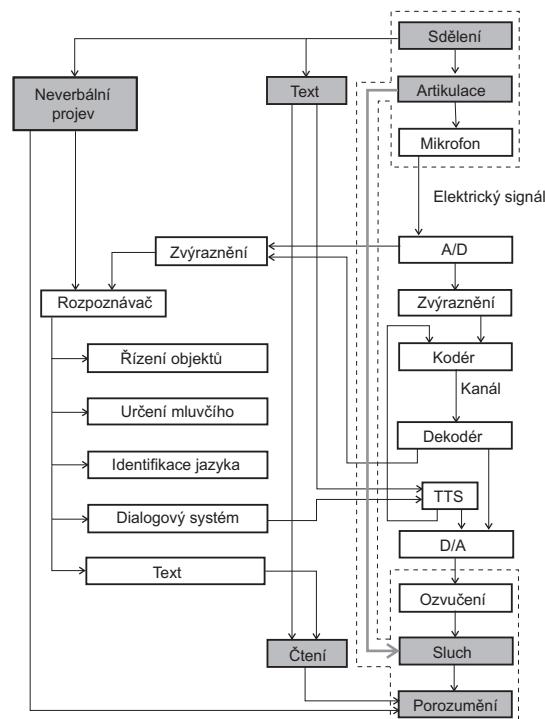
Pro analogové signály je významným parametrem přenosového média šířka frekvenčního pásma, ve kterém je signál bez zkreslení přenesen. Lidský sluch je schopen slyšet zvuky ve frekvenčním intervalu 20 Hz až 16 kHz. Již v počátcích telefonie bylo zjištěno, že pro srozumitelnou komunikaci hlasem stačí frekvenční pásmo 300 Hz až

3,4 kHz. Přenos s takovou šíří frekvenčního pásma dává na přijímací straně po ozvučení tzv. telefonní kvalitu. V oblasti hlasových technologií se setkáme s požadavky na šíří pásma počínaje tímto intervalem, až po požadavky odpovídající přenosu všech slyšitelných zvuků.

Analogové systémy hlasových technologií jsou v současné době na okraji zájmu technického vývoje. Hlasové technologie stojí na číslicové technice a využití počítačů. Požadavky na analogové telefonní spoje však vedly v minulosti k ohromnému rozvoji teorie frekvenčních filtrů a korektorů a rozvoji teorie vedení. Výsledky tohoto výzkumu se nyní uplatňují při zpracování analogového signálu před digitalizací a při úpravách signálu vytvořeného D/A převodníkem, než je zaveden do reproduktoru. Kromě toho jsou znalosti z oblasti analogových systémů v řadě aplikací východiskem pro vývoj systémů číslicových. S konfrontací analogového a digitálního „světa“ se setkáme v oboru hlasových technologií velmi často.

Úplný řetězec, který zprostředkovává sdělení informace mezi dvěma lidmi hlasem, ukazuje obr. 8. V obrázku jsou naznačeny všechny funkční bloky umožňující komunikaci na dálku (v čárkovaném rámci analogovou cestu „vzduchem“) a „odbočky“, umožňující hlasovou komunikaci se stroji.

V obr. 8 jsme analogovou cestu elektrického signálu neuvodili, protože se jí nebudeme dále zabývat. Věnujeme se možnostem digitálních systémů.



Obrázek 8: Hlasové technologie

V obrázku jsou šedě vyplněny bloky, které patří komunikačním nástrojům (fyziologickým) člověka. Ostatní bloky představují technické prostředky hlasových komunikací.

³V současné době existují mikrofony schopné přímé digitalizace okamžitých hodnot vstupní veličiny. Rovněž jsou známy konstrukce ozvučovacího zařízení, jejichž vstupem je binární kód přímo určující výchylku membrány. Jím se zde nebudeme zabývat.

Digitální přenos. Na uvedeném obrázku si nejprve povšimneme bloků, které umožňují digitální přenos řečového signálu a zabezpečují některé funkce, které s digitálním přenosem souvisejí.

Již jsme uvedli, že signál vytvoří mikrofon, za kterým je obvykle zapojen zesilovač. Již na tomto místě nás musejí zajímat nejen vlastnosti mikrofonu, ale i frekvenční vlastnosti a nelineární zkreslení zesilovače. V praxi se můžeme setkat s tím, že analogový signál je nejen zesilován, ale i komprimován ve svém rozkmitu. V algoritmech dalšího digitálního zpracování se taková nelineární transformace musí brát v úvahu. Frekvenční vlastnosti zesilovače, příp. s ním spojeného frekvenčního filtru, jsou významné z hlediska následujícího vzorkování. Pro vzorkování musí být dodržen vzorkovací teorém [2]. V systémech, ve kterých se digitalizuje analogový signál frekvenčně omezený do telefonního pásma, je běžné použití vzorkovací frekvence 8 kHz. V systémech, kde je požadován kvalitnější hlasový signál (pro přenos, syntézu i rozpoznávání), zachycující identitu mluvčího, barvu hlasu, jeho intonační rysy, apod., používáme analogového kanálu s širším frekvenčním pásmem a tomu odpovídající vyšší vzorkovací frekvenci.

Vzorky hlasového signálu jsou vedeny do analogově-číslicového převodníku, který signál kvantuje a digitalizuje. Proces digitalizace vede k narušení signálu kvantizačním šumem [2]. Při rekonstrukci hlasového signálu na přijímací straně je kvantizační šum činitelem zhoršujícím jeho srozumitelnost i subjektivní akceptovatelnost. Kvantizace a digitalizace jsou procesy, které mapují spojitý interval maximálního rozkmitu analogového signálu na interval diskretních číselných hodnot, jejichž počet závisí na počtu bitů, kterými je digitalizovaný signál reprezentován. Ve většině aplikací je to tak, že konstantní podinterval (kvantovací krok) q analogové veličiny reprezentuje krok, se kterým se mění číselná hodnota o jedničku. Takový kvantizér označujeme jako lineární a kvantizační šum je pro náhodný signál rovněž náhodným signálem s rovnoměrným rozdělením. V telekomunikačních systémech se používá kvantizér, který nemá pro celý číselný interval konstantní hodnotu q . Pro malé analogové hodnoty je q menší a pro velké amplitudy je q větší. Kvantizér označujeme jako nelineární. Lze ukázat [3], že pro řečový signál lze navrhnout nelineární kvantizér tak, že výkon kvantizačního šumu je výrazně menší než při kvantizaci kvantizérem lineárním. Navíc je charakter kvantizačního šumu pozměněn tak, že je po rekonstrukci digitalizovaného signálu méně vnímán posluchačem. Proto je v přenosových telekomunikačních systémech nelineární kvantizér běžně používán. Tuto skutečnost je nutno mít na paměti tehdy, kdy používáme jakýkoli algoritmus lineárního zpracování digitálního signálu. Na vstup číslicového filtru, korektoru, spektrálního analyzátoru, apod., musejí být zaváděny vždy číselné hodnoty odpovídající výstupu lineárního kvantizéru. Pokud je v systému nelineární kvantizér, musí být data transformována na hodnoty odpovídající lineární kvantizaci.

Telekomunikační technika využívá pro standardní digitální přenos osmibitových čísel z výstupu nelineárního kvantizéru. K tomuto řešení se dospělo prověřením srozumitelnosti a akceptovatelnosti rekonstruovaného signálu. Bitový tok tedy je 64 kbit/s. Pro potřeby zpracování takového signálu je nezbytné použít transformace do dvanáctibitového kódu tak, aby zůstala zachována kvalita. V řadě aplikací hlasových technologií se však setkáme se systémy, které používají lineární kvantizéry a vícebitové kódování.

Uvedené zpracování řečového signálu je považováno za bezztrátové, což znamená, že lze získat z digitálního signálu v D/A převodníku signál se stejným časovým průběhem, jaký měl signál původní, jen s tím, že je narušen kvantizačním šumem. Při dodržení vzorkovacího teorému k žádné další degradaci nedojde. V moderních telekomunikacích se používají algoritmy, které umožňují výrazně snížit bitový tok potřebný pro přenos hlasu (např. v systémech mobilní komunikace). Tam pak dochází k dalšímu zpracování digitalizovaného signálu a přenosu jiných dat, než jsou čísla nesoucí informaci o okamžitých hodnotách signálu. K přenosu hlasového signálu získáváme informaci v podobě tzv. parametrické reprezentace. Při takové kompresi informace je kladen důraz na srozumitelnost rekonstruovaného signálu na přijímací straně a nikoli na přesnost reprezentace jednotlivých vzorků časového průběhu, takže nemusí být vždy rekonstruovaný signál přesným obrazem signálu vstupního – elementární informace o okamžitých hodnotách je „ztracena“. Takovou transformaci a reprezentaci signálu označujeme jako ztrátovou a může do signálu vnést specifické zkreslení.

S ohledem na další text poznamenejme, že v literatuře nacházíme pod pojmem kodér a kódování dvě úrovně zpracování signálu získaného z A/D převodníku. V prvním případě se jedná o způsob reprezentace jednotlivých vzorků hlasového signálu např. binárním kódem se znaménkem a absolutní hodnotou [4], nebo u ztrátového zpracování o reprezentaci binárně kódovaných koeficientů lineární predikce. Jde tedy o způsob vyjádření čísel, která potřebujeme pro reprezentaci signálu. V druhém případě se jedná o vytvoření binárních kódů (většinou z výše zmíněných dat), které jsou vhodné pro přenos telekomunikačními přenosovými prostředky. V jednom každém z těchto procesů vytvoření digitálních dat se mohou uplatnit požadavky, které vyplývají z potřeb toho druhého.

Digitalizovaný hlasový signál je obrazem akustického tlaku, který zachytil mikrofon. Není tedy reprezentantem ideální artikulace, ale nese řadu vad, které komunikaci v konkrétních podmínkách provázejí. Za neodstranitelné považujeme vady způsobené samotným mluvčím při artikulaci (nemoc, stres, opilost). Posluchač je obvykle schopen porozumět takto narušené promluvě, avšak v automatických rozpoznávacích řeči může degradovaná řeč způsobit těžko překonatelný problém. Jiné narušení hlasového signálu může pocházet z prostředí, ve kterém je mikrofon umístěn. Patří sem dozvuky a ozvěny v uzavřených prostorech a hluky a rušivé zvuky, např. v jedoucím automobilu, v kanceláři, ve výrobním provozu, na hlučné ulici a ko-

nečně hlasy jiných mluvčích. V těchto případech jde o narušení signálu, pro která jsou vyvíjeny algoritmy umožňující jejich redukci, či úplné odstranění. Proto jsme v obr. 8 zařadili blok označený „Zvýraznění“.

Jako kodér jsme v našem obrázku označili blok, který upravuje digitální data do podoby, která je vhodná pro digitální přenos. Úkolem dekodéru je přivést do bloku rekonstruujícího hlasový signál potřebná digitální data. Degradace hlasového signálu digitálním přenosovým systémem je velmi specifická (ztracené pakety, zpoždění, ozvěny ap.). Algoritmy hlasových technologií, především rozpoznávačů řeči, ale i prostých ozvučovacích procedur, které s takovými degradacemi počítají a jejich vliv redukuje nebo odstraňují, představují zvláštní oblast výzkumu.

Blok D/A převodu není v našem obrázku chápán jen jako elektronická součástka převádějící binární kód na analogovou úroveň napětí, ale zahrnuje všechny algoritmy vedoucí z číselné reprezentace signálu k jeho analogové podobě, tedy i převod z parametrické reprezentace. Výstupem tohoto bloku je většinou schodovitý časový průběh napětí, jehož obálka má tvar analogového časového průběhu hlasového signálu. Pro použití ve sluchátku nebo reproduktoru musí být průběh napětí korigován vhodným analogovým filtrem [2].

V následujícím textu ukážeme, jak jsou do řetězce sdělování řečových informací začleňovány automatické elektronické systémy.

3.2. Syntéza a rozpoznávání řeči

V systému komunikací lze oprávněně očekávat, že mluvčím nemusí být jen člověk, ale že lidský hlas bude produkovat i stroj, a to na základě textu, který mu předepíšeme.

Za jednu z nejstarších automatických hlasových služeb lze považovat telefonní informaci o přesném čase. Hlas živého mluvčího byl uložen v podobě optického záznamu na skleněných kotoučích, které se otáčely přesně stanovenou rychlostí a reprodukován ze signálu na výstupu fotoelektrického snímače. Záznam celé informace jistě nemusel mluvčí nahrávat celých čtyřicet hodin. Hlášení byla skládána z opakujících se úseků, a to podle požadované výstupní informace. Systém byl zjevně analogový. Tento princip je v digitální technologii aplikován i v současných hlasových informačních systémech. Lze ho použít všude tam, kde hlasová informace vyžaduje pouze omezený slovník a omezenou stavbu produkovaných vět, přičemž pojem „omezený“ je velmi relativní a závisí na velikosti paměti, kterou máme k dispozici. Z výše zmiňovaných technologií však může sehrát významnou roli jakýkoli algoritmus umožňující kompresi digitálních dat. Pokud se stupeň komprese dat stane hlavním požadavkem na konstrukci konkrétního zařízení, může se algoritmus komprese stát činitelem rozhodujícím o srozumitelnosti a kvalitě produkované řeči. Takto pracující systém nelze doslovně považovat za hlasový syntetizér, protože reprodukuje záznam hlasu konkrétního mluvčího. Dojem „syntetičnosti“ hlasu u tohoto systému vznikne tehdy, jsou-li na sebe navázána

slova, která nebyla jako sousední vyslovena a zaznamenána, tedy není-li respektována koartikulace mezi slovy, nebo tehdy, kdy zaznamenaná slova nerespektují intonaci odpovídající jejich pozici ve větě určitého typu (pokles hlasu na konci oznamovací věty, intonace věty tázací). Uvedený princip nepřináší z hlediska hlasových technologií zvláštní problémy.

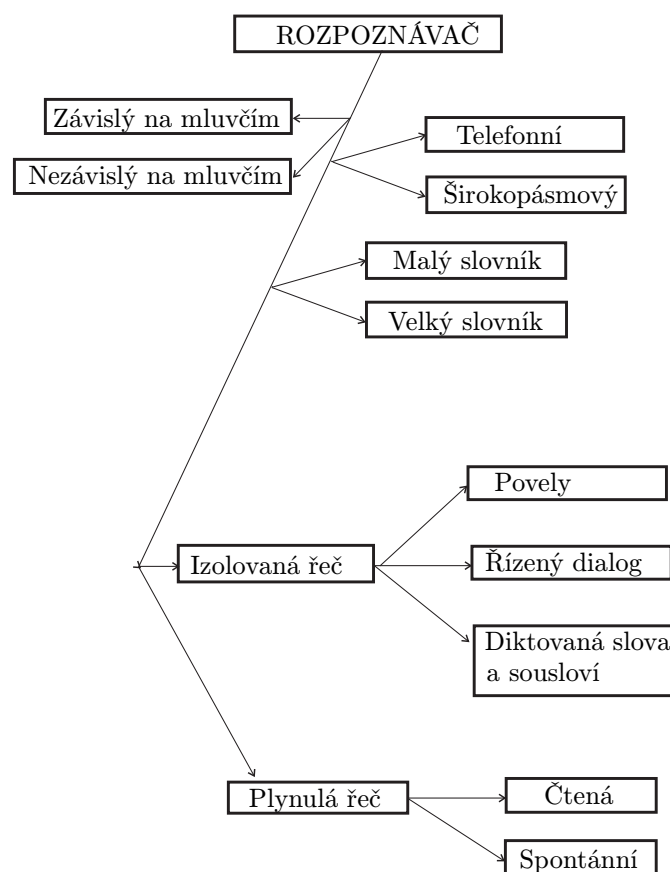
Počítač pracující se systémem řetězení slov, příp. opakujících se slovních spojení, může být podle našeho obr. 8 použit jak v systému pracujícím v telekomunikační síti, tak v systému lokálním. V telekomunikační síti tento princip využívají telefonní informační služby. Příkladem lokálního systému může být „hovořící“ palubní deska automobilu, hlášení nádražního rozhlasu apod. Významné je i použití v interaktivních dialogových systémech, kde dialog může být multimodální, tj. např. uživatelský vstup je řešen klávesnicí (mobilního telefonu, počítače), nebo dotykovou obrazovkou a stroj odpovídá hlasem. Jak jsme naznačili v obrázku, jsou konstruovány i systémy, ve kterých je dialog veden ze strany uživatele i ze strany stroje přirozenou řečí.

Mezi syntetizéry můžeme zařadit i systémy, které pracují s promluvou či jejím záznamem a určitým způsobem promluvu modifikují. Hovoříme pak o resyntéze přirozené promluvy, a to např. s cílem odstranit z promluvy charakteristické rysy řeči určitého mluvčího, nebo s cílem výrazně zrychlit tok řeči – je známo, že nevidomí jsou schopni porozumět i několikanásobně zrychlené řeči ze záznamu pořizovaného normálně hovořícím mluvčím. Pokusy s resyntézou promluv, které byly zaznamenány a zpracovány algoritmy parametrizace a komprese dat, mohou být velmi užitečným východiskem pro vývoj systémů TTS. Skutečné systémy TTS ozvučují zcela libovolný text, který nebyl při konstrukci systému znám.

Jednou z nejsledovanějších oblastí hlasových technologií je algoritmizace a implementace rozpoznávačů řeči. Obrázek 8 ukazuje, že jejich umístění v komunikačním prostředí může být lokální (v počítači na našem stole, v automobilu, na informačním sloupu na ulici ap.). Velmi časté však je jejich umístění na konci telekomunikačního spoje (v telefonní ústředně, na serveru v bance, v turistickém informačním centru nebo v rezervační kanceláři jízdenek a letenek). Na obrázku vidíme, že kromě rozpoznávání obsahu promluvy může být požadováno rozpoznání (identifikace nebo verifikace) mluvčího, nebo rozpoznání jazyka, kterým mluvčí hovoří.

Nejaktuálnějším problémem je však konstrukce systému pro rozpoznávání obsahu promluvy. Na obrázku 9 je ukázána širší specifika, kterými se rozpoznávače vzájemně odlišují. Takto vymezené specifikace vznikaly s vývojem algoritmů a s možnostmi výpočetní techniky, do které byly algoritmy implementovány.

Hledisko založené na počtu mluvčích, pro které má rozpoznávač sloužit, velmi významně ovlivňuje náročnost konstrukce rozpoznávačů ve všech níže uvedených typech. Každý mluvčí má určitou charakteristickou výslovnost (ostatně té je využito v systémech rozpoznávání mluvčího) a rozpoznávač s ní může počítat – „naučí se naslouchat



Obrázek 9: Kategorie rozpoznávačů řeči

hlasu svého pána“. Pro rozpoznávače nezávislé na mluvčím musejí být nalezeny rysy promluv, které jsou velkému počtu mluvčích společné. Jejich nalezení vyžaduje velké množství hlasových záznamů, z kterých se společné charakteristiky hlásek extrahují a specifika jednotlivých mluvčích eliminují.

Umístění rozpoznávače na konci telekomunikačního spoje nebo v místě hlasového vstupu je dalším hlediskem, které se uplatní v konstrukci níže uvedených typů. Telefonní rozpoznávač se vyznačuje většinou neovlivnitelným vzorkovacím kmitočtem a s ním spojeným omezeným frekvenčním pásmem. Dále je v telefonním rozpoznávači dána úroveň kvantizačního šumu. Ke kvantizačnímu šumu se přidávají další rušivé složky pocházející z komunikačního kanálu. Naproti tomu lokální rozpoznávač může pracovat s kvalitněji digitalizovaným signálem, tj. s vyšším vzorkovacím kmitočtem a větším počtem bitů v datech reprezentujících hlasový signál.

Pokud uvažujeme o možnostech mluvčího při komunikaci se strojem, můžeme podmínky rozdělit podle toho, zda rozpoznávač pracuje s izolovanými slovy, pro která má ve své paměti uloženy vzory, nebo zda rozpoznávač bude muset vzory hledat tak, aby teprve složený do delší sekvence identifikovaly odpovídající promluvu. Takové skládání bude jistě snazší, když bude promluva mít pevnou stavbu respektující gramatiku jazyka, než když půjde o promluvu spontánní vznikající bez přípravy. Pod heslem „Čtené“ si představujeme nejen skutečné předčítání z napsané předlohy, ale i promluvu určenou pevnými pravidly, tak krátkou, že si ji mluvčí přesně zapamatuje a nezmění ji podle vlastní úvahy. Algoritmus rozpoznávání delších souvislých promluv bude jednodušší, když půjde o ustálená slovní spojení, např. diktování vícemístných čísel, ať desetinných, či např. čísel telefonních.

Každý rozpoznávač musí být vybaven slovníkem slov, která má rozpoznat. V systémech rozpoznávajících izolovaná slova z malého slovníku (několik desítek slov) lze mít v rozpoznávači uložena jako vzory celá slova zaznamenaná ve fázi trénování. V případě rozsáhlejších slovníků již nelze systém trénovat na všechna možná slova, ale slovník musí být vytvořen z foneticky přepsaných textů, které očekávaná slova obsahují a trénování musí probíhat na souborech promluv, které obsahují všechny fonetické elementy použité při přepisu reprezentativních textů do fonetické podoby.

Reference

- [1] PALKOVÁ, Z.: *Fonetika a fonologie češtiny.*, Univerzita Karlova, vydavatelství Karolinum, 1994.
- [2] UHLÍŘ, J. – SOVKA, P.: *Číslíkové zpracování signálů.*, Vydavatelství ČVUT Praha, 2002. 2. vydání.
- [3] JAYANT, N. S. – NOLL, P.: *Digital Coding of Waveforms: Principles and Applications to Speech and Video.*, Prentice-Hall, Englewood Cliffs, N. J., 1984.
- [4] SOVKA, P. – UHLÍŘ, J. – ČMEJLA, R.: *Úvod do číslíkového zpracování signálů.*, Vydavatelství ČVUT Praha - skriptum, 2003.
- [5] UHLÍŘ, J. – SOVKA, P. – POLLÁK, P. – HANŽL, V. – ČMEJLA, R.: *Technologie hlasových komunikací.*, Vydavatelství ČVUT Praha - monografie, 2007.

Analýza základní frekvence, amplitudového a frekvenčního kolísání hlasivek u Parkinsonovy nemoci

Jan Rusz^{a,b}, Roman Čmejla^a a Hana Růžičková^c

^aKatedra teorie obvodů, ČVUT–FEL, Technická 2, 166 27 Praha 6

^bUniverzita Karlova v Praze, Laboratoř biokybernetiky a počítačové podpory výuky,

Ústav patologické fyziologie 1. LF UK, U nemocnice 5, 128 53 Praha 2

^cNeurologická klinika 1. LF UK a VFN, Kateřinská 30, 128 21 Praha 2

e-mail: [ruszj1;cmejla]@fel.cvut.cz, hruzi@email.cz

This study examined the acoustic measures of fundamental frequency F_0 , frequency instability (*Jitter*), amplitude instability (*Shimmer*) and noise to harmonic ratio (*NHR*), which are the traditional parameters examined in the acoustic voice analyses and it was used in the Parkinson disease (PD) patients analyses. The acoustic voice characteristics of 14 patients with early stage PD were compared with data from 14 normal control subjects (CS). Voice recordings consisted of two read texts consisting of 80 and 136 words and a free monologue lasting for 90 seconds for F_0 variance measurement, one sentence consisting of 28 words and scale singing for F_0 range measurement. For the next F_0 range measurement, one announce, one imperative and one interrogative sentence were used. Next voice recordings consisted of at least 4 second prolongation of the vowel (i) and prolongation of the nasal (m) for *Jitter*, *Shimmer* and *NHR* measurements. For all the parameters, automatic scripts were developed or analyses software Praat was used. Voiced parts of texts were used only for frequency variability measurement. The voices of PD groups demonstrated lower F_0 range and monopitch. Present data also suggested that the PD patients' voices were characterized by excess *Jitter*, *Shimmer* and *NHR*. We introduce a new measure of dysphonia for the objective assessment of frequency instability with a standard linear whitening filter.

1. Úvod

Idiopatická Parkinsonova nemoc (PN) je progresivní neurodegenerativní onemocnění, které bylo poprvé popsáno Parkinsonem (1817) v jeho „Essay on the Shaking Palsy“. PN je charakterizována progresivní ztrátou dopaminergních neuronů, zejména v pars compacta substantiae nigrae [1], ale také v jiných částech mozku (Bhatnagar & Andy, 1995; Brodal, 1998; Lang & Lozano, 1998). Ztráta dopaminergních neuronů je asociována se značnými motorickými a nemotorickými deficity jako jsou klidový třes končetin, bradykineze, svalová rigidita, posturální poruchy a poruchy chůze. U velkého množství pacientů je obvyklým a jedním z prvních příznaků porušení řeči, které je charakterizováno jako hypokinetická dysartrie [2, 3].

Minulá studie ukázala [4], že u pacientů s výskytem PN dochází ke změnám v artikulační rychlosti, počtu pauz, intenzitě hlasitosti, a statisticky potvrdila, že u všech měření došlo k výraznému rozdílu mezi pacienty s výskytem PN a kontrolní skupinou (KS). PN pacienti dělají v čtených textech menší počet pauz, které jsou ovšem delší; mají sníženou artikulační rychlost a dochází u nich k většímu kolísání a klesání intenzity hlasitosti. Cílem této studie je analýza základní frekvence hlasivek F_0 (označována též jako fundamentální frekvence) pro čtené texty, monolog, zpívanou stupnici a věty, frekvenčního kolísání (*Jitter*), amplitudového kolísání (*Shimmer*), poměru šum-signal (*NHR*) a práce zároveň přináší novou metodu pro objektivní hodnocení nepravidelnosti kmitání hlasivek spjaté s dysartrickým profilem PN, který se využívá v klinické logopedické

praxi při diagnostice motorických řečových změn. Jedná se o vyšetření založené na subjektivním ohodnocení řady úkolů (týkající se respirace, fonace, faciální svalové činnosti, diadochokineze, reflexní činnosti, artikulace, srozumitelnosti četby a mluvy, tempa a prozodie řeči). Na základě bodového hodnocení jednotlivých úkolů je posléze zpracováván tzv. celkový dysartrický profil vyšetřované osoby [5]. Objektivní analýza hlasových nahrávek není důležitá pouze pro stanovení stupně dysfonie, ale také pro následné monitorování zlepšení stavů pacientů po nasazení léčby.

2. Metody

2.1. Pacienti

Pro tuto studii byla získána data od 14 českých PN pacientů ve věkovém rozmezí 52 až 82 let (průměr: 64,55; 12 mužů, 2 ženy), ve srovnání se 14 účastníky kontrolní skupiny KS (průměr: 56,43; 6 mužů, 8 žen). Všichni diagnostikovaní pacienti se nachází v rané fázi PN, kdy poškození hlasu není v mnoha případech rozeznatelné lidským sluchem, ale postihuje více než 80 % pacientů s výskytem PN [6]; jejich onemocnění nepřekročilo délku trvání 5 let, nemají žádnou historii neurologických, respiračních, laryngeálních, řečových nebo hlasových poruch či onemocnění a nebyla u nich zahájena léčba pomocí levodopy ani jiných anti-Parkinsonových léčiv. Členové KS nemají žádnou evidenci neurologických onemocnění a hlasových poruch.

2.2. Nahrávání promluv

Nahrávky pro všechny PN pacienty byly pořízeny v odhlučněné místnosti (< 45 dB) Neurologické kliniky 1. lékařské fakulty Univerzity Karlovy v Praze ve vysoké kvalitě s vzorkovací frekvencí 48 kHz. PN pacienti byli požádáni o:

- a) přečtení dvou referenčních textů o délce 80 (*Text 1*) a 136 (*Text 2*) slov;
- b) jednu volnou promluvu o délce 90 sekund (*Monolog*);
- c) zpívanou stupnici vzestupně a sestupně s využitím postupné produkce tónových intervalů hlásky (a) (*Stupnice*), přičemž všichni zúčastnění byli požádáni o svůj nejlepší výkon;
- d) udržovanou fonaci hlásky (i) po dobu nejméně čtyř sekund;
- e) udržovanou znělou bilabiální nazálu (m) po dobu nejméně čtyř sekund;
- f) vybranou větu o délce 28 slov (*Věta*);
- g) krátkou oznamovací (*Ozn*), tázací (*Taz*) a rozkazovací větu (*Roz*).

2.3. Akustické analýzy

Základní frekvence hlasu frekvence F_0

Základní frekvence hlasivek F_0 je fyzikální charakteristikou řečového signálu a odpovídá výšce hlasu tak, jak je vnímána posluchačem. Periodou základního hlasivkového tónu (též pitch perioda) nazýváme $T_0 = 1/F_0$. Výpočet F_0 je proveden v časové oblasti s využitím autokorelační funkce [7]

$$r_N(d) = \frac{1}{N} \sum_{i=0}^{N-d-1} x(i)x(i+d), \quad (1)$$

kde N je počet vzorků a d je délka okna. Algoritmus pracuje ve dvou iteracích, přičemž v první iteraci je spočítán práh znělosti a v druhé iteraci je tento práh aplikován pro výpočet F_0 . Skript je automaticky ošetřen, v případě vzniku chyby v označení periody jsou tyto špatně označené periody smazány. Pro ověření byly provedeny kontroly v softwaru Praat. První měření patří do oblasti prozodie a jedná se o analýzu melodie ve dvou referenčních čtených textech a jednom monologu, která je realizována jako směrodatná odchylka F_0 , tj. $F_{0\sigma}$ [8–13]. Dalším měřením spadajícím do oblasti fonace je měření rozsahu F_0 ($F_{0R} = \max\{F_0\} - \min\{F_0\}$) u zpívané stupnice [13] a čtené věty [13, 14]. Poslední měření je také rozsah F_{0R} z oblasti intonace, přičemž jde o vyslovování krátké věty se stejným zněním v kontextu oznamovacím, rozkazovacím, tázacím, kdy je převážně prokázáno, že PN pacienti nejsou schopni výrazné změny tónu u vyslovování tázací věty, tudíž produkují významně menší rozsah F_0 [12].

U hodnocení základní frekvence hlasu narážíme na podstatný problém, který se týká především pohlaví jednotlivců. Míra přirozené variace hlasu souvisí s průměrnou

F_0 jednotlivce; mluvčí s vyšší F_0 hlasu (obvykle žena) má mnohem vyšší vibrace a chvění hlasivek než s mluvčí s nižší F_0 hlasu (obvykle muž). Problém tedy nastává, je-li F_0 měřena v absolutní frekvenční stupnici (Hz), kde nastávají značné rozdíly při vzájemné klasifikaci mezi oběma pohlavími. Výsledky těchto pozorování vedou k myšlence, že mnohem relevantnější stupnice, na které lze stanovit variace F_0 hlasu, je stupnice logaritmická (tónová), protože má mnohem vhodnější předpoklady pro hodnocení charakteristik než stupnice absolutní frekvenční [14–16], a to nejen z hlediska potlačení rozdílu mezi pohlavím jednotlivců, ale také je očekáván mnohem lepší dopad při hodnocení PN, která je spojena s dysfonií [15, 16]. Pro převod z absolutní frekvenční do logaritmické půltónové stupnice [17] je použit vztah

$$F_{0p} = 12 \frac{\log\left(\frac{F_0}{60}\right)}{\log 2}, \quad [\text{půltón}]. \quad (2)$$

Jitter/Shimmer/NHR

Analýzy klinických hlasových nahrávek ve zvukové technice jsou obvykle doprovázeny použitím klasických metod založených na lineárních matematických principech zahrnujících frekvenční kolísání (*Jitter*), amplitudové kolísání (*Shimmer*) a poměr šum-harmonická (*NHR*) [13, 18, 19]. *Jitter*, *Shimmer* a *NHR* byly spočítány s využitím softwaru Praat, který je založen na klasických lineárních krátkodobých signálových technikách.

Shimmer je definován jako střední rozdíl mezi amplitudami hlasivkových pulzů. Pro tuto práci byl využit lokální *Shimmer*, který je průměrným dekadickým logaritmem podílů mezi amplitudami po sobě jdoucích period

$$\text{Shimmer} = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| 20 \log\left(\frac{A^{(i+1)}}{A^i}\right) \right|, \quad [\text{dB}]. \quad (3)$$

Jitter je definován jako střední rozdíl délek sousedních period, dělený střední délkou periody. Byl použit lokální absolutní *Jitter*, který je definován jako průměrný absolutní rozdíl mezi periodami jdoucími po sobě

$$\text{Jitter} = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| T_0^{(i)} - T_0^{(i+1)} \right|, \quad [\text{ms}]. \quad (4)$$

NHR představuje relativní poměr šumu a harmonických složek signálu (5). V případě, že se bude jednat o čistě periodický signál, činitel *NHR* bude nabývat nuly. S přibývajícím šumem roste hodnota *NHR* k jedné.

$$\text{NHR} = \frac{E_{\text{noise}}}{E_{\text{harm}}}, \quad [-]. \quad (5)$$

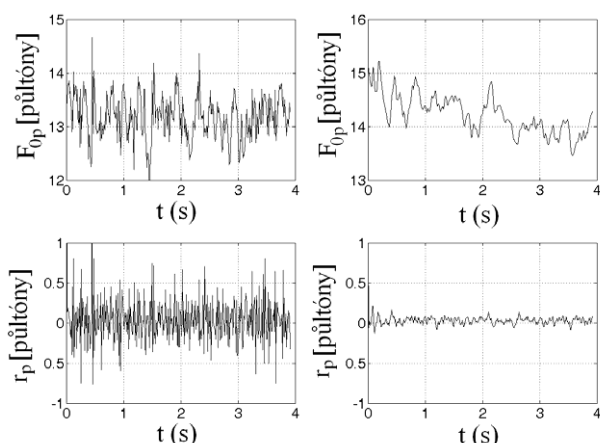
Konstanta N je počet vzorků signálu, A je maximální hodnota amplitudy signálu, T_0 je perioda signálu a E_{harm} je energie harmonické složky signálu. Přesnější popis algoritmů lze nalézt v [20, 21]. Parametry jsou počítány u fonace hlásky (i) a nazály (m) po dobu nejméně čtyř sekund.

Hodnocení nepravidelnosti kmitání hlasivek s využitím bělicího filtru

Objektivní analýza hlasu bývá tradičně limitována technikami založenými na lineárních matematických principech (*Jitter*, *Shimmer*, *NHR*). Tyto techniky kooperují s důležitým faktem, a to, že je v signálu zahrnuta určitá periodičita, tzn. signál obsahuje opakující se znaky, které mohou být rozděleny do časových segmentů a individuálně analyzovány. V důsledku některých dysfonií, jako je např. PN, dochází k nepravidelnosti kmitání hlasivek a pacientův hlas často nevykazuje známky periodicity. Tento fakt vedl k vytvoření nového parametru pro objektivní hodnocení analýz s využitím bělicího filtru [16]. Analýza je provedena na hláse (i) a nazále (m) v délce nejméně čtyř sekund. Pro automatickou implementaci algoritmu je v prvním kroku spočítána sekvence pitch periody a převedena do půltónové logaritmické stupnice F_{0p} . V dalším kroku jsou získány koeficienty lineární predikce LPC (linear predictive coefficients) [22], které jsou následně použity jako bělicí filtr

$$r_{p(i)} = F_{0p} - \sum_{j=0}^{P-1} a_{(i)} F_{0p(i-j)}, \quad (6)$$

kde P je řád prediktoru (délka signálu – 1) a a jsou koeficienty predikce $a = [1 \ a(2) \ \dots \ a(P+1)]$. Ty jsou pak použity jako FIR filtr a výsledkem je relativní půltónová variační sekvence r_p , z níž se pak jako výsledný parametr určí směrodatná odchylka ($r_{p\sigma}$), která nám určuje pravidelnost F_0 tím způsobem, že čím je větší $r_{p\sigma}$, tím méně je pravidelná F_0 jednotlivce.



Obrázek 1: Detaily pitch periody v půltónech F_{0p} (horní část); detaily relativní půltónové variační sekvence r_p (dolní část); levá část znázorňuje pacienta s výskytem PN, pravá část účastníka KS

Tabulka 1: Výsledky analýzy rozsahu základní frekvence hlasu v různých promluvách

	PN		KS	
	μ	σ	μ	σ
$F_{0\sigma}$ [půltón]				
Text 1**	1,48	0,47	2,35	0,60
Text 2**	1,73	0,66	2,59	0,46
Monolog**	1,36	0,30	2,47	0,67
F_{0R} [půltón]				
Stupnice*	12,08	2,71	15,58	3,72
Věta	10,69	1,70	12,15	4,06
Ozn	6,90	2,21	7,39	2,43
Roz**	7,49	1,93	11,29	3,12
Taz**	8,24	2,31	12,13	2,94

μ – střední hodnota; σ – směrodatná odchylka;

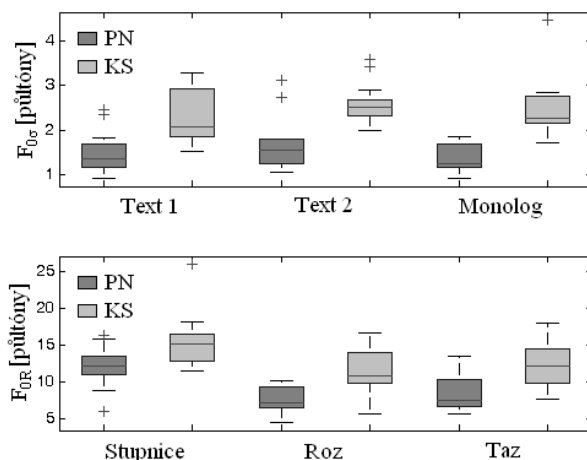
* $p < 0,01$;

** $p < 0,001$.

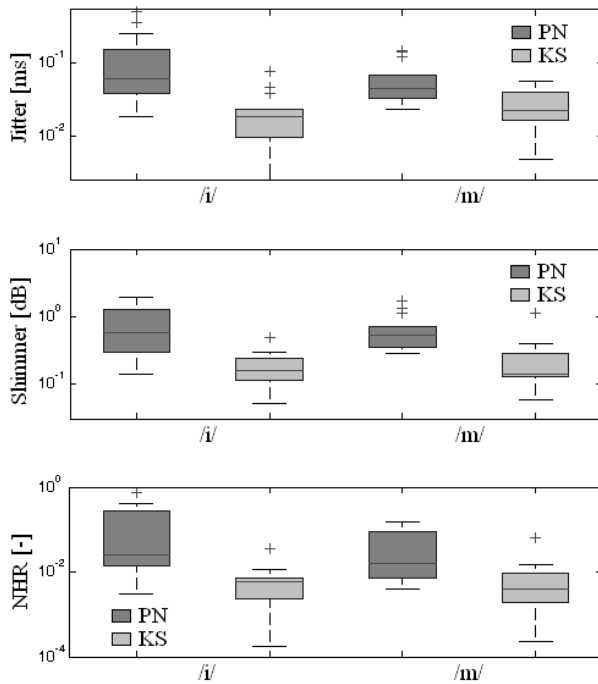
3. Výsledky

3.1. Základní frekvence hlasu F_0

Pro všechna měření je stanovena střední hodnota a směrodatná odchylka a výsledky byly ověřeny statistickými t -testy. Z výsledků (Tabulka 1) lze konstatovat, že u pacientů s výskytem PN dochází k výraznému snížení melodičnosti jak při čtení textů, tak i při samostatné promluvě. U těchto měření melodie $F_{0\sigma}$ jsou výsledky na nejvyšší hladině významnosti. Při měření rozsahu F_{0R} bylo potvrzeno, že se vznikem PN dochází u pacientů ke snížení jejich hlasového rozsahu. To je prokázáno nejen zpíváním stupnice, ale mnohem větší rozdíly nastávají u rozkazovací a tázací věty, kdy PN pacienti nejsou schopni vytvořit u těchto vět patřičný důraz. U krátké nebo dlouhé oznamovací věty nedochází k závažným rozdílům. Všechna statisticky významná měření jsou vynesena v grafu (Obrázek 2).



Obrázek 2: Měření rozsahu základní frekvence hlasu F_0 pacientů s výskytem PN a KS



Obrázek 3: Měření *Jitter/Shimmer/NHR* u hlásek s prodlouženou fonací pacientů s výskytem PN a KS

3.2. *Jitter/Shimmer/NHR*

Dle výsledků (tab. 2, obr. 3) je možné konstatovat, že parametry, které posuzují amplitudovou a frekvenční nepravidelnost a poměr šum-signal, jsou statisticky významné pro hodnocení PN. U *Jitteru* došlo k lepším výsledkům pro nazálu (m), zatímco u *Shimmeru* pro hlásku (i), a *Shimmer* je statisticky nejvýznamnějším parametrem pro tuto sérii měření. Statistickou hladinu významnosti $p < 0,05$ u obou měření dosáhl i parametr *NHR*.

Tabulka 2: Výsledky analýzy *Jitter/Shimmer/NHR* u hlásek s prodlouženou fonací

	PN		KS	
	μ	σ	μ	σ
<i>Jitter</i> [ms]				
(i)*	0,13	0,14	0,02	0,02
(m)**	0,06	0,04	0,03	0,02
<i>Shimmer</i> [dB]				
(i)***	0,82	0,59	0,18	0,11
(m)**	0,65	0,42	0,25	0,27
<i>NHR</i> [-]				
(i)*	0,150	0,220	0,007	0,009
(m)*	0,041	0,049	0,010	0,016

* $p < 0,05$;

** $p < 0,01$;

*** $p < 0,001$.

3.3. Nepravidelnosti kmitání hlasivek $r_{p\sigma}$

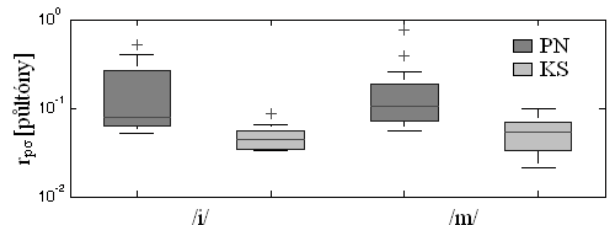
Výsledky hodnocení frekvenční nepravidelnosti kmitání hlasivek novou metodou $r_{p\sigma}$ dosáhly u hlásky (i) vyšší statistické významnosti než hodnocení s využitím *Jitteru* (tab. 3, obr. 4). Obě měření jsou statisticky významná. Pro srovnání a ověření hodnocení je provedena vzájemná korelace r s *Jitterem*, který rovněž hodnotí frekvenční nepravidelnost. Pro hlásku (i) dosáhl korelační koeficient $r = 0,93$ ($p < 0,001$) a pro nazály (m) $r = 0,8$ ($p < 0,001$). Pro prokázání, že se nedá dosáhnout stejných výsledků klasickým měřením směrodatné odchylky F_{0p} přímou metodou z hlásky (i) a nazálu (m), bylo taktéž provedeno pro obě měření hodnocení této směrodatné odchylky F_{0p} , to však nedosáhlo žádné statistické významnosti.

Tabulka 3: Výsledky analýzy nepravidelnosti kmitání hlasivek u hlásek s prodlouženou fonací $r_{p\sigma}$

	PN		KS	
	μ	σ	μ	σ
$r_{p\sigma}$ [přítón]				
(i)**	0,17	0,15	0,05	0,02
(m)*	0,18	0,19	0,06	0,03

* $p < 0,05$;

** $p < 0,01$.



Obrázek 4: Měření nepravidelnosti kmitání hlasivek $r_{p\sigma}$ pacientů s výskytem PN a KS

4. Diskuze

Předchozí zahraniční studie prokázaly, že u pacientů s výskytem PN dochází ke změně základní frekvence hlasu F_0 . Jedním z typických vznikajících projevů je ztráta melodičnosti ve čtených či samostatných promluvách [7–12], jejichž významné statistické výsledky jsou ověřeny současně v této studii. Mezi další změny, které se v práci prokázaly, patří zmenšení rozsahu při zpívané stupnici [12]. Při čtení oznamovací věty nedochází k rozdílu mezi skupinami PN a KS, avšak v literatuře [11] se udává, že u PN pacientů dochází k nižšímu rozsahu u tázací věty, způsobenému monotónností hlasu a následně špatným vnímáním důrazu. V této studii bylo prokázáno, že ke statisticky stejnému významu dochází nejen u vyslovování tázací, ale i u rozkazovací věty.

Mezi další měření pro hodnocení PN patří určení střední hodnoty F_0 při čtení textu, monologu a obvykle pro hlásku (i). Jedná se o měření, ve kterém nelze omezit závislost na pohlaví ani po převedení do pŕltónové logaritmické stupnice. Prŕmĕrná střední hodnota základní hlasivkové frekvence žen a mužŕ se zásadně liší. Pro tuto studii nebylo možné obstarání většihŕ množství dat ženskĕho pohlaví v počátečním stadiu PN, avšak předešlé studie dostatečně potvrdily [7–12], že u tohoto měření nedochází ke statisticky významným rozdílŕm mezi skupinami PN a KS.

Dalšími parametry, u kterých byly již dříve prokázány změny v souvislosti s výskytem PN jsou *Jitter/Shimmer/NHR* [12, 17, 18], přičemž práce prokázala statistickou významnost ve všech těchto měřeních, kdy ve studii [12] je prokázán statistický význam pouze u *Jitteru*.

Pro novou metodu hodnocení frekvenční nepravidelnosti $r_{p\sigma}$ byla provedena vzájemná korelace se standardní metodou využívající *Jitter*. Z výsledků tohoto ovĕření lze tvrdit, že oba parametry hodnotící frekvenční kolísání mají vysokou vzájemnou korelaci, a že obĕma metodami lze objektivně posuzovat nepravidelnost kmitání hlasivek. Hodnocením s využitím parametru $r_{p\sigma}$ je dosaženo podobných výsledků, ale zásadní výhodou této metody pro posouzení frekvenční nepravidelnosti s využitím bělicíhŕ filtru je však na rozdíl od *Jitteru* spolehlivý algoritmus i při hodnocení neharmonických signálŕ, které jsou ŕzce spojeny právě s dysartrickĕm profilem PN pacientŕ.

5. Závĕr

Výsledky této studie mohou být použity jako základ pro pochopení zmĕn vznikajících v řeči pacientŕ již v rané fázi PN a dále jako základ pro její možnou klasifikaci. Studie ukázala, že u PN pacientŕ vznikají řečové dysartrie, které se týkají základní frekvence hlasivek a různé hlasové nepravidelnosti. S výskytem PN dochází postupně k větší monotónnosti hlasu, což se projevuje neschopností přidat důraz při vyslovování otázky či rozkazu a celkovým snížením variace hlasu při čtení textu nebo samostatné promluvě. Zároveň ze studie vyplývá, že u pacientŕ s výskytem PN dochází při vyslovování samohlásky nebo znĕlé nazály k vyšší frekvenční a amplitudové nepravidelnosti. Práce zároveň přináší nový postup pro měření frekvenční nepravidelnosti vhodné zvláště pro dysartrické profily hlasu.

Podĕkování

Tato práce je podporována z výzkumných zámĕrŕ „Transdisciplinární výzkum v oblasti biomedicínskĕho inženýrství“ (MSM 6840770012), „Neuropsychiatrickĕ aspekty neurodegenerativních onemocnění“ (MSM 0021620849) a (MSM 0021620806), grantŕ „Analýza a modelování biologických a řečových signálŕ“ (GAĀR 102/08/H008) a Rozpoznávání mluvené řeči v reálných podmínkách (GAĀR 102/08/0707), projektŕ „Non-motorické projevy Parkin-

sonovy nemoci“ (IGA MZĀR NR9220) a (MSM 2C06031), a firmou Creative Connection, s.r.o.

Autoři ĕekují MUDr. Janĕ Picmausovĕ, MUDr. Veronice Majerovĕ, MUDr. Jiřimu Klempířovi, Ph.D. a doc. MUDr. Janu Rothovi, CSc. za poskytnutí souboru klinických dat a prof. MUDr. Evženu Růžickovi, DrSc. za koncepci a vedení klinické studie a za připomínky k rukopisu.

Reference

- [1] Hornykiewicz, O.: *Biochemical aspects of Parkinson's disease*, Neurology, 1998, 51(Suppl 2):S2–S9.
- [2] Canter, G. J.: *Speech characteristic of patients with Parkinson's disease. I. Intensity, pitch, and duration*, J. Speech Hear Disord, 1963, 28:221–229.
- [3] Darley, F. L., Aronson, A. E., Brown, J. R.: *Differential diagnostic patterns of dysarthria*, J. Speech Hear Res, 1969, 12:426–496.
- [4] Ruzs, J., Čmejla, R.: *Analýza rychlosti řeči a intenzity u Parkinsonovy nemoci*, Akustické listy, 2008, 14(2–4):13–16.
- [5] Hedánek, J., Roubíĕková, J.: *Dysartrický profil: test 3F*, 1997, Filozofická fakulta, UK Praha, 71s.
- [6] Pereira, J. C., Schelp, A. O., Montagnoli, A. N., et al.: *Residual signal auto-correlation to evaluate speech in Parkinson's disease patients*, Arq Neuropsiquiatr, 2006, 64(4):912–915.
- [7] Hess, W.: *Pitch determination of Speech signals*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983, p. 352.
- [8] Goberman, A. M.: *Correlation between acoustic speech characteristics and non-speech motor tasks in Parkinson's disease*, Med. Sci. Monit, 2005, 11(3):CR109–116.
- [9] Goberman, A. M., Elmer, L. W.: *Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease*, Journal of Communication Disorders, 2005, 38:215–230.
- [10] Goberman, A. M., Blomgren, M.: *Fundamental Frequency Change During Offset and Onset of Voicing in Individuals with Parkinson Disease*, Journal of Voice, 2008, 22(2):178–191.
- [11] Goberman, A. M., Coelho, C., Robb, M.: *Phonatory characteristics of Parkinsonian speech before and after morning medication: the ON and OFF states*, Journal of Communication Disorders, 2002, 35:217–239.
- [12] Goberman, A. M., Coelho, C.: *Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy*, NeuroRehabilitation, 2002, 237–246.

- [13] Holmes, R. J., Oates, J. M., Phyland, D. J., Hughes, A. J.: *Voice characteristics in the progression of Parkinson's disease*, Journal of Communication Disorders, 2000, 35(3):407–418.
- [14] Mori, H., Kobayashi, Y., Kasuya, H., et al.: *Evaluation of fundamental frequency (F_0) characteristics of speech in dysarthrias: A comparative study*, Acoust. Sci. & Tech., 2005, 26(6):540–543.
- [15] Moore, B. C. J.: *An introduction to the psychology of hearing*, 5th ed. Amsterdam; Boston: Academic Press, 2003.
- [16] Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., Ramig, L. O.: *Suitability of dysphonia measurement for telemonitoring of Parkinson's disease*, IEEE Transaction on Biomedical Engineering, 2008, (in press).
- [17] Miller, M.: *The Complete Idiot's Guide to Music Theory*, 2nd ed. Indianapolis: Alpha, 2005, p.19.
- [18] Mourao, L. F., Aguiar, P. M. C., Ferraz, F. A. P., et al.: *Acoustic voice assesment in Parkinson's disease patients submitted to posteroventral pallidotomy*, Arq Neuropsiquiatr, 2005, 63(1):20–25.
- [19] Costello, D. A. E., Little, M. A., Harries, M. L.: *Objective dysphonia quantification in vocal cord paralysis: novel nonlinear and classical methods*, 2008, Journal of Voice, (in review).
- [20] Boersma, P., Weenink, D.: *Praat, a system for doing phonetics by computer*, Glot International, 2001, 5(9/10):341–345.
- [21] Boersma, P.: *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound*, In: Proc. the Institute of Phonetic Sciences, 1993, Vol. 17, pp. 97–110.
- [22] Psutka, J.: *Komunikace s počítačem mluvenou řečí*, 1. ed. Academia Praha, 1995.

Pozvánka na seminář „Pružné uložení budov“

Název semináře: Pružné uložení budov
Datum konání: 26. května 2009, 9.00–14.00
Místo: Praha, Sál Hospodářské komory, Freyova ulice
Přihlášky: Studio D akustika, s. r. o.
U Sirkárny 467/2a
370 04 České Budějovice
akustikad@akustikad.com
Vložené: Kč 1000,– za osobu – zahrnuje sborník, oběd a občerstvení

Program:

- Pružné uložení budov v České republice, vize – Ing. Jan Stěnička
- Zdroje vibrací a strukturálního hluku. Očekávané problémy v interiéru budov a fyzikální modely předpokládaných hodnot vibrací a strukturálního hluku. Měření snížení vibrací a strukturálního hluku v budovách – Dr. Wildör (Drážďany – Genest und Partner Ing.-GmbH)
- Výsledky zkoušek pružných podložek vhodných jako vibroizolace do základů budov – Ing. Zuzana Fišarová (VUT Brno)
- Ochrana před strukturálním hlukem a vibracemi z metra, železnice a tramvají – Ing. Jana Dolejší (Český rozhlas Praha)
- Využití SYLOMERu pro pružné uložení budov – Ing. Zindler (Getzner Werkstoffe GmbH, Mnichov)
- Konstrukce – Návrh – Kontrola jakosti

Akustické listy: ročník 15, číslo 1 březen 2009

Vydavatel: Česká akustická společnost, Technická 2, 166 27 Praha 6

Počet stran: 20 Počet výtisků: 200

Redakční rada: M. Brothánek, O. Jiříček, J. Kozák, R. Čmejla, J. Volín

Jazyková úprava: R. Svobodová

Uzávěrka příštího čísla Akustických listů je 29. května 2009.

ISSN: 1212-4702

Vytisklo: Nakladatelství ČVUT, výroba

© ČsAS
NEPRODEJNÉ!