

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

ANALÝZA A ROZPOZNÁNÍ LOGOPEDICKÝCH VAD V
ŘEČOVÉM PROJEVU

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

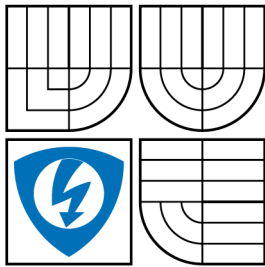
AUTOR PRÁCE
AUTHOR

JAN DIVIŠ

BRNO 2008



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND
COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

ANALÝZA A ROZPOZNÁNÍ LOGOPEDICKÝCH VAD V
ŘEČOVÉM PROJEVU
TITLE OF STUDENT'S THESIS

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

JAN DIVIŠ

VEDOUcí PRÁCE
SUPERVISOR

PROF. ING. ZDENĚK SMÉKAL, CSC.

BRNO 2008

ABSTRAKT

Bakalářská práce se zabývá logopedickou vadou dyslalií a její charakteristikou. Dále je popsán proces vytváření a reprezentace řeči. V práci jsou představeny základy zpracování a analýzy řečového signálu (LPC, keprální, MFCC). V programu MATLAB jsou provedena zobrazení charakteristik řeči a výpočty LPC, keprálních a melovských koeficientů. Práce je především zaměřena na nesprávné vyslovování hlásek „r“ a „ř“.

KLÍČOVÁ SLOVA

logopedie, dyslalie, analýza řeči, LPC, MFCC, keprum

ABSTRACT

This bachelor's thesis deals with logopaedia mistake called dyslalie and its characteristics. I described the process creation and representation of speech. There are presented bases of processing and analyses speech signal (LPC, cepstral, MFCC). I presented characteristics of speech and calculation of LPC, cepstral and Mel-frequency cepstral coefficients in the programme MATLAB. The bachelor's thesis includes problems of incorrect pronunciation sound „r“ and „ř“.

KEYWORDS

logopaedia, dyslalie, speech analysis, LPC, MFCC, cepstrum

DIVIŠ, J. *Analýza a rozpoznání logopedických vad v řečovém projevu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2008. 44 s.
Vedoucí bakalářské práce prof. Ing. Zdeněk Smékal, CSc.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Analýza a rozpoznání logopedických vad v řečovém Projevu“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne

.....

(podpis autora)

Děkuji vedoucímu bakalářské práce prof. Ing. Zdeňkovi Smékalovi, CSc. za vedení a konzultace při zpracování, které mi v průběhu psaní poskytl.

OBSAH

1 Úvod	10
2 Logopedie	11
2.1 Význam řeči	11
2.2 Poruchy artikulace	11
3 Dyslalie	12
3.1 Terminologie	12
3.2 Symptomatologie	12
3.3 Klasifikace z vývojového hlediska	12
3.4 Etiologie	13
3.5 Klasifikace podle etiologie	13
3.6 Výskyt	13
3.7 Dědičnost	13
3.8 Vliv prostředí	14
3.9 Poruchy zrakového a sluchového vnímání	14
3.10 Zásady krátkodobého cvičení	14
3.11 Substituční metoda	14
3.12 Zásady při substituční metodě	14
3.13 Individuální přístup	15
4 Proces vytváření řeči	16
4.1 Hlasové ústrojí	16
4.2 Artikulační ústrojí	17
4.2.1 Vytváření tónové struktury	17
4.2.2 Vytváření šumové složky	17
4.2.3 Samohlásky	17
4.2.4 Artikulační vlastnosti samohlásek	18
4.2.5 Souhlásky	18
4.2.6 Artikulační vlastnosti souhlásek	19
4.2.7 Souhlásky „r“ a „ř“	19
4.3 Reprezentace řeči	20
4.3.1 Fonetická úroveň	20
4.3.2 Akustická úroveň	20
4.4 Model vytváření řeči	21

5	Analýza řečového signálu	24
5.1	Pulsní kódová modulace (PCM)	24
5.2	Fourierova transformace	24
5.3	Lineární predikční analýza	25
5.4	Kepstrální analýza	26
5.5	Melovské kepstrální koeficienty	27
5.6	Analýza výslovnosti hlásek „r“ a „ř“	31
5.6.1	Hláska „r“	31
5.6.2	Hláska „ř“	31
5.6.3	Porovnání metod	32
6	Závěr	33
	Literatura	34
	Seznam symbolů, veličin a zkratk	35
	Seznam příloh	36
A	Přílohy	37
A.1	Výpočet LPC pro hlásky „ř,r“	37
A.2	Výpočet vyhlazené spektrální obálky pomocí kepstrálních koeficientů pro hlásky „ř,r“	38
A.3	Výpočet melovských kepstrálních koeficientů pro hlásky „ř,r“	39

SEZNAM OBRÁZKŮ

4.1	Hlasové ústrojí člověka. [4]	16
4.2	Vokalický trojúhelník	18
4.3	Časový průběh slova „tříška“.	20
4.4	Spektrum slova „tříška“.	21
4.5	Širokopásmový a úzkopásmový spektrogram slova „tříška“.	22
4.6	Model vytváření řeči. [4]	23
5.1	Model vytváření řeči s lineárním filtrem	25
5.2	Signálové spektrum pro hlásky „ř,r,l,p“ určené z koeficientů LPC (řád prediktoru $Q = 20$)	26
5.3	Postup kepstrální analýzy	27
5.4	Postup kepstrální analýzy pro hlásku „ř“	28
5.5	Signálové spektrum pro hlásky „ř,r,l,p“ určené z kepstrálních koeficientů	29
5.6	Nelineární úprava kmitočtové osy a výsledné rozmístění filtrů.	29
5.7	Postup výpočtu melovských kepstrálních koeficientů.	30
5.8	Melovské kepstrální koeficienty MFCC pro hlásky „ř,r,l,p“.	31
5.9	Porovnání správné výslovnosti s artikulační vadou pro hlásky „r“ a „ř“.	32

SEZNAM TABULEK

4.1	Typické hodnoty základního tónu řeči. [4]	17
4.2	Rozdělení souhlásek. [4]	19
5.1	Typické hodnoty počtu filtrů. [4]	30

1 ÚVOD

Téma bakalářské práce Analýza a rozpoznání logopedických vad v řečovém projevu je velice zajímavé, ale také obsáhlé, proto jsem se rozhodl zaměřit na logopedickou vadu dyslalií. Z důvodu nejčastějšího procentuálního výskytu se budu zabývat problematikou nesprávné výslovnosti hlásek „r“ a „ř“.

Na začátku bakalářské práce popisuji dyslalií a její charakteristiku. Dále pokračuji procesem vytváření řeči a následně její reprezentací. Poté nastíním základy zpracování řečového signálu a provedu analýzy zpracování řeči - lineárně predikční (LPC), kepstrální a výpočet melovských kepstrálních koeficientů (MFCC). Mým cílem je pomocí výše zmíněných analýz vyhodnotit nesprávné vyslovování hlásek „r“ a „ř“ v programu MATLAB.

2 LOGOPEDIE

Obor speciální pedagogiky, který se zabývá výchovou a vzděláním jedinců s poruchami řeči.

2.1 Význam řeči

- Řeč je nástrojem komunikace, plné funkce dorozumívání (kontakt).
- Nástrojem poznání (poznatků, vědomostí, citů, ...).
- Prostředkem výchovy.
- Obecná schopnost řeči se realizuje prostřednictvím konkrétního jazyka (mateřský jazyk), má význam i emocionální, protože se stává součástí vlastenectví, hodnot domova.
- Řeč je nositelem estetických kvalit, má blízko k umění (zejména literatuře).
- Kvalita řeči je jedním z předpokladů k úspěšnosti člověka ve společnosti. Jedinec trpící vadou řeči může mít sníženou sebedůvěru, komplexy méněcennosti.

2.2 Poruchy artikulace

Mezi poruchy artikulace, které se vyskytují u dětí předškolního věku, řadíme dyslalii a vývojovou dysartrii. Můžeme konstatovat, že se jedná o postižení, u něhož se vyskytuje odlišná etiologie, přičemž symptom, kterým je narušená artikulace, je společný.

3 DYSLALIE

3.1 Terminologie

Pojem dyslalie je termín, který zahrnuje několik specifických stránek narušené řečové produkce. Odborným termínem dyslalie se označuje několik jejich spojitých úrovní. Hlavní, jež dyslalii charakterizuje, je úroveň fonetická, vztahující se k užívání jednotlivých hlásek a představující analytickou stránku řeči. Fonologická úroveň se týká používání elementárních řečových zvuků ve spojeních do slabik a slov, případně vět, které jsou základem další, lingvistické úrovně řeči.

Porucha řeči na fonetické úrovni se projevuje např. vynecháváním hlásek (delecí), jejich zaměňováním či nahrazováním (substitucí) až nepřesným vyslovováním v místě artikulace (distorzí).

Na fonologické úrovni se poruchy projevují v plynulé řeči, když jsou jednotlivé hlásky ovlivňovány sousedstvím okolních hlásek, následujících nebo předcházejících, kupříkladu na začátku, uprostřed či na konci slova, či v důsledku suprasegmentálních znaků řeči, jako jsou pauzy v řeči, důrazy, přízvuky, melodie a rytmus ve větách nebo v tempu řeči.

Dyslalie v nejširším slova smyslu spočívá v neschopnosti nebo poruše používání zvukových vzorů řeči v procesu komunikace podle řečových zvyklostí a norem příslušného jazyka, což je i příčinou toho, že jde o nejrozšířenější vadu řeči v lidské společnosti.

3.2 Symptomatologie

Při hláskové dyslalii dítě hlásku:

- vynechá - tzv. mogilalie (tříška - tíška)
- nahrazuje jinou hláskou, tzv. paralalie (tříška - tlíška)
- tvoří chybně, pak se poruchy označují příponou -ismus, například: rotacismus, lambdacismus

3.3 Klasifikace z vývojového hlediska

- Fyziologická (vývojová), která je do sedmi let věku dítěte přirozeným mluvním projevem.
- Patologická - kdy dítě z mnoha příčin není schopné osvojit si tvoření některých hlásek či jejich skupin.

3.4 Etiologie

Dyslalie je nejčastější poruchou komunikačních schopností u dětí. V roce 1951 - 1952 se přibližně uvádí u dětí obecné školy 24% dyslalie. V roce 1973 uvádí Sovák [2] na základě šetření Olchavové u dětí 1. stupně základní školy 42,8% dyslalií. V roce 1989 uvádí Kutálková [2] u dětí pražských škol 38% dyslalií. V roce 1999 uvádí Lechta [3] u dětí 1. tříd 36% dyslalií.

3.5 Klasifikace podle etiologie

Funkční dyslalie vzniká u dětí s nedostatečnou vjemovou (senzorickou) nebo motorickou schopností. Při diagnostice nacházíme genetické dispozice, nesprávný mluvní vzor a jen minimální orgánové poškození během těhotenství nebo porodu.

Organická dyslalie vzniká v důsledkem narušení dostředivých (impresivní dyslalie) a odstředivých nervových drah (expresivní dyslalie), při narušení centra řeči (centrální dyslalie), dále při anatomických a inervačních odchylkách mluvidel (labiální, dentální, palatální, lingvální a nazální dyslalie). Také poruchy sluchu mohou podstatně ovlivnit vývoj výslovnosti. Závisí na rozsahu a druhu sluchového postižení. Nepřesné slyšení způsobuje nesprávnou realizaci hlásek.

3.6 Výskyt

Výskyt dyslalie u chlapců je poměrně vysoký, tvoří až dvě třetiny dětí přijatých do logopedické péče. U dětí s mentální retardací bývá zpožděný fyziologický vývoj mluvené řeči, spolupůsobí i organické odchylky mluvidel, sociální prostředí atd. Toto však není vždy pravidlem, protože dyslalie se vyskytuje také u dětí s nadprůměrným IQ, zvláště, je-li tempo jejich řeči velmi rychlé. Zpravidla pak jde o rotacismus. Mezi nejčastější příčiny patří vlivy dědičnosti, pohlaví dítěte, poruchy sluchu a zraku, poruchy CNS, riziková těhotenství, poškození při porodu a vlivy prostředí.

3.7 Dědičnost

Názory na vliv dědičnosti jsou různé. Dokazuje se výskyt narušených komunikačních schopností u příslušníků rodiny, nejčastěji u otce. Lechta [3] (1990) uvádí tento dědičný vliv jako nespecifickou dědičnost, projevující se artikulační neobratností, případně snížením schopnosti nebo vrozené řečové slabosti, která pak způsobuje vadnou výslovnost.

3.8 Vliv prostředí

Nesprávný řečový vzor v rodině, nepřiměřený řečový vzor - mazlivá řeč, vícejazyčné prostředí, chyby ve výchovném přístupu (tresty, výsměch, citové strádání, zanedbávající výchova v rodině) mohou vést nejprve k vadnému tvoření výslovnosti a následně k její fixaci.

3.9 Poruchy zrakového a sluchového vnímání

Výslovnost je vážně narušena tehdy, má-li dítě narušené sluchové vnímání. Nelze podceňovat převodní nedoslýchavosti, jako následek častého onemocnění horních cest dýchacích nebo následek zvětšení nosní vegetace. Nedoslýchavost způsobuje, že dítě správně nerozlišuje hlavně sykavky a má také narušeny modulační faktory řeči. Porucha zrakového vnímání může být rovněž příčinou vadné výslovnosti pro nepřesné vnímání artikulačních pohybů hlavně v raném věku.

3.10 Zásady krátkodobého cvičení

Úprava poruch a vad výslovnosti je pro každé dítě individuálně volena tak, aby postihla jeho vývojová specifika. U malých dětí je nízká schopnost udržet pozornost, proto cvičíme maximálně po dobu 3–4 minut, s intenzivním sluchovým a zrakovým vzorem, za využití dostupných pomůcek - zrcadla, indikátoru, počítačového vzoru, kresby atd. Cvičení v časových intervalech dále opakujeme a rozvíjíme po jednotlivých krocích. Dobu cvičení volíme podle individuálních možností dítěte.

3.11 Substituční metoda

Substituční metoda je velmi důležitá, protože vytváří nový sluchový a artikulační vzor. Vede dítě ke správnému tvoření hlásky a k vyhasínání vadných návyků. Výchova fonemického sluchu je proces obtížný, zvláště u dětí s opožděným vývojem řeči a u dětí s mentální retardací.

3.12 Zásady při substituční metodě

Děti mají snahu tvořit požadované hlásky ve cvičných slovech hlasitě a přehnaně, nevnímají pak jejich tvoření a znění. Proto raději cvičíme tiše a s minimální artikulační silou. Nedílnou součástí logopedické praxe jsou také zásady podporující proces učení.

3.13 Individuální přístup

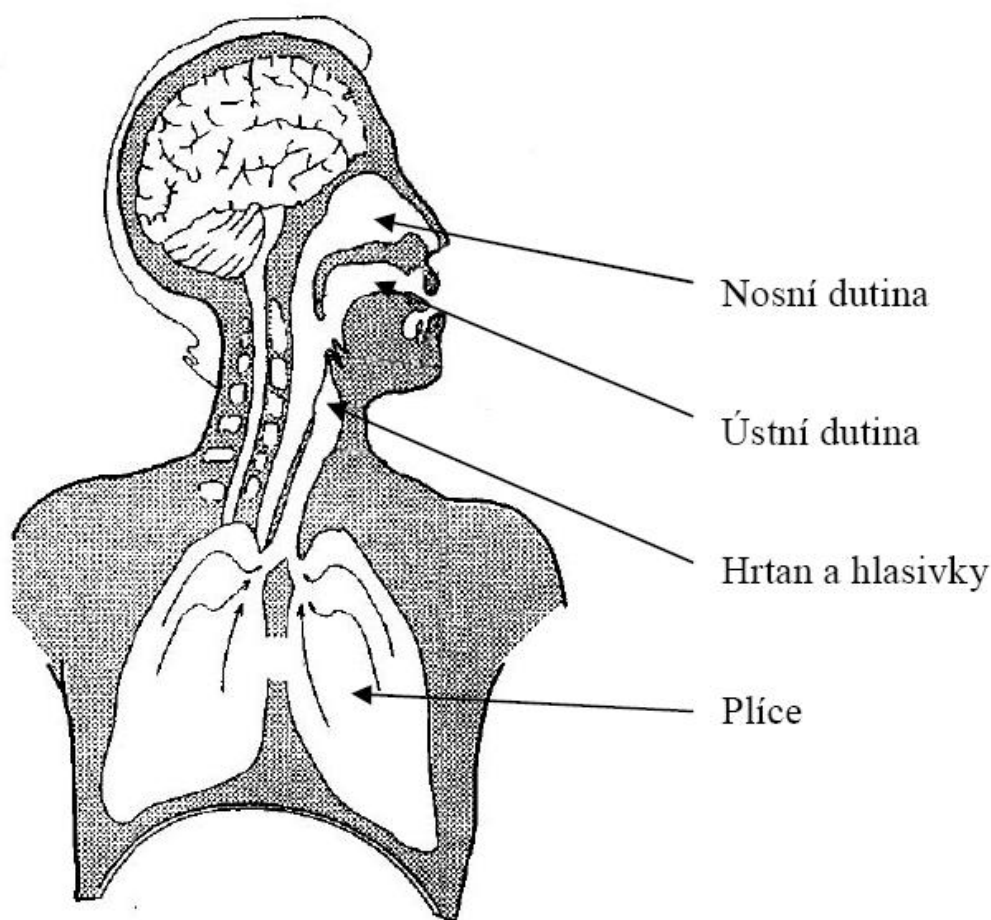
Zásada vývojovosti - při jejím nerespektování dochází, zřejmě z neznalosti vývoje hlásek, k závažným chybám při stanovení postupů úpravy a volby metod především v dětském věku. Včasné usměrnění artikulace hlásek zmenšuje riziko vytváření vadných artikulačních stereotypů. Vývojová řada hlásek je odrazem vývoje CNS a myelinizace nervových drah, které od napodobivého žvatlání až do 6–7. roku věku dítěte prodělávají velké změny. To umožňuje dítěti osvojovat si stále náročnější hlásky. Jako pomůcka k tomu, jak hodnotit fyziologický vývoj artikulace a kdy případné artikulační problémy řešit.

Zásada názornosti - k úspěšné nápravě je třeba využít multisenzoriální přístup - zapojení zraku, sluchu, hmatu a vnímání změny pohybu. Dítěti poskytneme co nejpřesnější informace o artikulačním postavení a sluchovém vjemu.

4 PROCES VYTVÁŘENÍ ŘEČI

4.1 Hlasové ústrojí

Pro vytváření řeči existuje v lidském těle několik řečových orgánů, které z hlediska tvorby řeči tvoří hlasový trakt. Ten je stručně znázorněn na obr. 4.1



Obr. 4.1: Hlasové ústrojí člověka. [4]

Zdrojem energie pro řeč jsou plíce. Nejdůležitější část hlasového ústrojí tvoří hlasivky, které se nacházejí v hrtanové dutině. Průchodem vzduchu se hlasivky rozkmitají. V důsledku kmitání hlasivek se vzduchový proud „rozdrobí“ tak, že se víceméně pravidelně střídá vždy proud hustšího a řídkého vzduchu. Vzniká vzduchová vlna, kterou vnímáme jako zvuk. Tento periodický proud vzduchových pulsů bývá označován jako základní hlasivkový tón. Frekvence kmitání hlasivek se označuje T_0 a nazývá se frekvence základního hlasivkového tónu. Převrácenou hodnotu $T_0 = 1/F_0$

nazýváme perioda základního hlasivkového tónu. Základní hlasivkový tón má rozsah asi 60–400 Hz. Frekvence kmitání je u žen, mužů a dětí různá viz. tabulka 4.1

Tabulka

Muži	80-160 Hz	Průměr - 132 Hz
Ženy	150-300 Hz	Průměr - 223 Hz
Děti	200-600 Hz	Průměr - 382 Hz

Tab. 4.1: Typické hodnoty základního tónu řeči. [4]

4.2 Artikulační ústrojí

Jeho význam spočívá v tom, že umožňuje vytvářet velké množství různých zvuků, které charakterizují mluvený jazyk. Skládá se z dutiny hrdelní, ústní a nosní. Mezi nejvýznamnější artikulátory patří jazyk, potom rty, měkké patro, zuby, tvrdé patro nebo čelisti.

4.2.1 Vytváření tónové struktury

Při průchodu základního hlasivkového tónu nadhrtanovými dutinami dochází vlivem rezonance uvnitř hrdelní a ústní dutiny ke změně rozložení akustické energie. Ta se soustředí kolem určitých frekvencí, kterým říkáme **formantové frekvence**. Oblasti zesílení akustické energie se nazývají **formanty** a značí se F_1, F_2, \dots . Pokud se do procesu zapojí i nosní dutina, dochází k potlačení některých frekvenčních oblastí, tzv. **antiformantů**. Vzniká složený zvuk (hlas) a tvoří podstatu znělých hlásek (především samohlásek).

4.2.2 Vytváření šumové složky

Artikulátory ovlivňují průchod vzduchu nadhrtanovými dutinami. Svým pohybem mohou průchod na různých místech zúžit, uzavřít nebo uvolnit. Výdechový proud vzduchu se pak prodírá přes vytvořené překážky a vzniká šum různého druhu. Šum tvoří základ částí řeči, kterým říkáme souhlásky.

4.2.3 Samohlásky

Čeština rozlišuje 10 samohlásek: /a/, /e/, /i/, /o/, /u/, /á/, /é/, /í/, /ó/, /ú/. Liší se od sebe jak nastavením hlasového traktu při jejich artikulaci, tak délkou. V češtině se odpovídající krátké a dlouhé samohlásky liší především trváním, jejich artikulace a

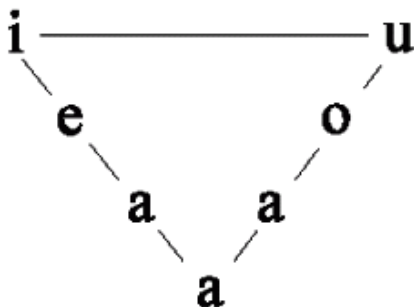
spektrální vlastnosti jsou v zásadě shodné. Samohlásky tvoří jádro slabiky a jsou nositelkami tónové kvality a estetického vyznění řeči. Akustické signály samohlásek jsou charakteristické kvaziperiodickým průběhem s periodou základního hlasivkového tónu a vyšší amplitudou.

Zvláštní postavení v českém jazyce mají dvojhlásky /ou/, /au/, /eu/, přičemž poslední dvě se objevují pouze ve slovní zásobě slov přejatých a bývají zřetelně pocíťovány jako příznak cizosti slova.

4.2.4 Artikulační vlastnosti samohlásek

Při artikulaci samohlásek se využívá fonačního postavení hlasivek - hlasivky kmitají a výdechový proud vzduchu volně prochází z hrtanu hrdelní a ústní dutinou a nakonec je vyzařován rty do okolí. Proud vzduchu se v nadhrtanových dutinách do cesty nestaví žádná překážka. Z řečových orgánů mají kromě hlasivek při artikulaci českých samohlásek největší význam jazyk, rty a měkké patro.

Poloha a pohyb jazyka je nejvýznamější artikulační charakteristikou českých samohlásek. Při pohybu jazyka horizontálním směrem rozlišujeme samohlásky přední, střední, zadní. Při pohybu jazyka vertikálním směrem rozlišujeme samohlásky vysoké, středové, nízké. Rozdělení českých samohlásek se tradičně zobrazuje ve formě vokálního trojúhelníku 4.2.



Obr. 4.2: Vokální trojúhelník

4.2.5 Souhlásky

Čeština má 27 souhláskových fonémů. Zatímco samohlásky jsou důležité zejména pro estetické vyznění řeči, přesná výslovnost souhlásek je základní oporou srozumitelnosti řeči. Akustické signály souhlásek jsou příznačné přítomností šumu a v porovnání s akustickými signály samohlásek i menší velikostí. Shluky jednotlivých souhlásek lze proto poměrně jednoduše rozeznat od samohlásek. Problémem však již může být jak rozlišit od sebe navzájem signály jednotlivých souhlásek.

4.2.6 Artikulační vlastnosti souhlásek

Při artikulaci souhlásek dochází ke „ztěžování“ průchodu výdechového proudu vzduchu hlasovým ústrojím. Ztíženým průchodu vzduchu je možné v zásadě docílit dvěma způsoby: vytvořením překážky nebo zúžením hlasového traktu v určitém místě hlasového ústrojí. Podle místa a způsobu vytvoření překážky, popř. zúžení, vznikají různé souhlásky, jejichž podstatou je šum. Při artikulaci českých souhlásek jsou důležité především čtyři artikulační charakteristiky: způsob artikulace, místo artikulace, postavení měkkého patra a činnost hlasivek. Pomocí těchto základních charakteristik je možné vytvářet a diferencovat jednotlivé souhlásky. Dělení souhlásek je v tabulce 4.2.6.

Tabulka

Souhlásky	Závěrové	Úžinové	Polozávěrové
Párové znělé	p t ť k	s š f ch	c č
Párové neznělé	b d ď g	z ž v h	dz dž
Nepárové znělé	m n ň	l j r ř	

Tab. 4.2: Rozdělení souhlásek. [4]

4.2.7 Souhlásky „r“ a „ř“

Tyto souhlásky vznikají tak, že se artikulační orgány na některém místě těsně přiblíží, ale úplně se neuzavřou a ponechají mezi sebou úžinu. Když se výdechový proud prodírá touto úžinou, vzniká třecí šum, který tvoří podstatu zvuku. Nazývají se také **třené souhlásky** (frikativy). Variantu úžinové artikulace představuje způsob, při němž se určitá pohyblivá část mluvidel během artikulace rozkmitá. Výsledný zvuk může mít do značné míry tónový charakter, jako je tomu u „r“, či u nich může převládat šumová složka jako u „ř“. Takto tvořeným hláskám říkáme **frikativy kmitavé** (vibranty).

Sonorní frikativa „r“ je velmi podobná souhláskám. Časové průběhy vlny jsou periodické, její amplituda však bývá menší, neboť při artikulaci se musí vypořádat se zúžením hlasového traktu. Také spektra této souhlásky se velmi podobají samohláskám, pouze bývá o několik decibelů slabší. Stejně jako u samohlásek je energie soustředěna v okolí nižších formantů. Šumová složka spektra bývá velmi slabá a ve spektrogramech není obvykle ani zřetelná. Spektrum je navíc modulováno periodickou frekvencí přibližně 20 Hz, která vzniká vibrací špičky jazyka při artikulaci.

U **šumové frikativy** „ř“ jsou zřejmé oba typy buzení hlasového traktu - periodické hlasivkové pulsy i frikativní šum vytvářený v místě zúžení hlasového traktu.

Díky své znělé složce obsahují většinou slabou formantovou strukturu. Obvykle dochází k modulaci šumové složky periodickou frekvencí přibližně rovnou 40–60 Hz, která vzniká vibrací špičky jazyka podobně jako u „r“. V případě „ř“ však tato periodičnost není tak výrazná, zvláště v pozici na začátku slova.

4.3 Re prezentace řeči

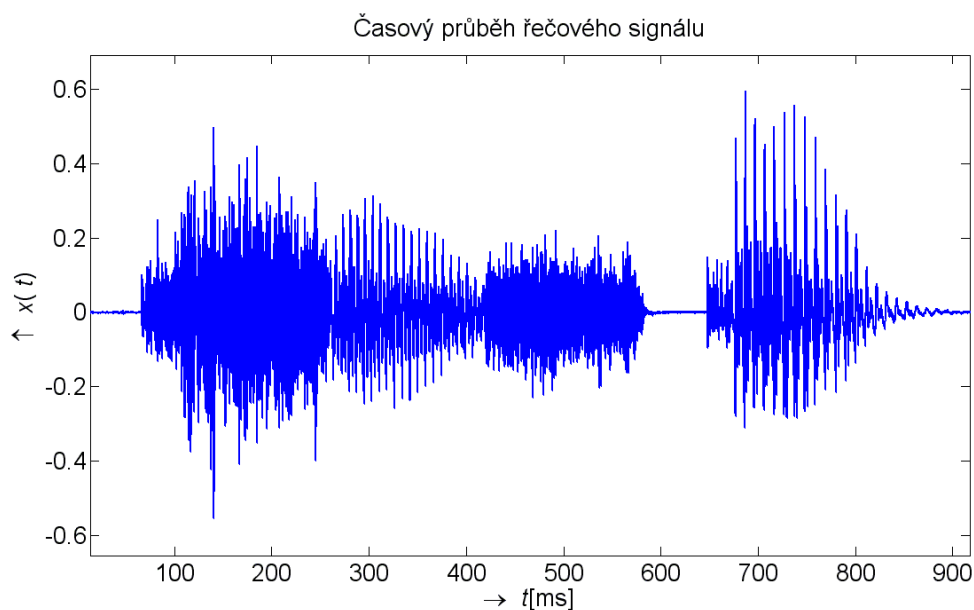
4.3.1 Fonetická úroveň

Na fonetické úrovni se zkoumá především způsob, jakým artikulační ústrojí člověka vytváří jednotlivé zvuky řeči, tedy studujeme činnost řečových orgánů při řeči, způsob tvoření, artikulace, charakter výsledných zvuků a také jejich sluchové hodnocení.

4.3.2 Akustická úroveň

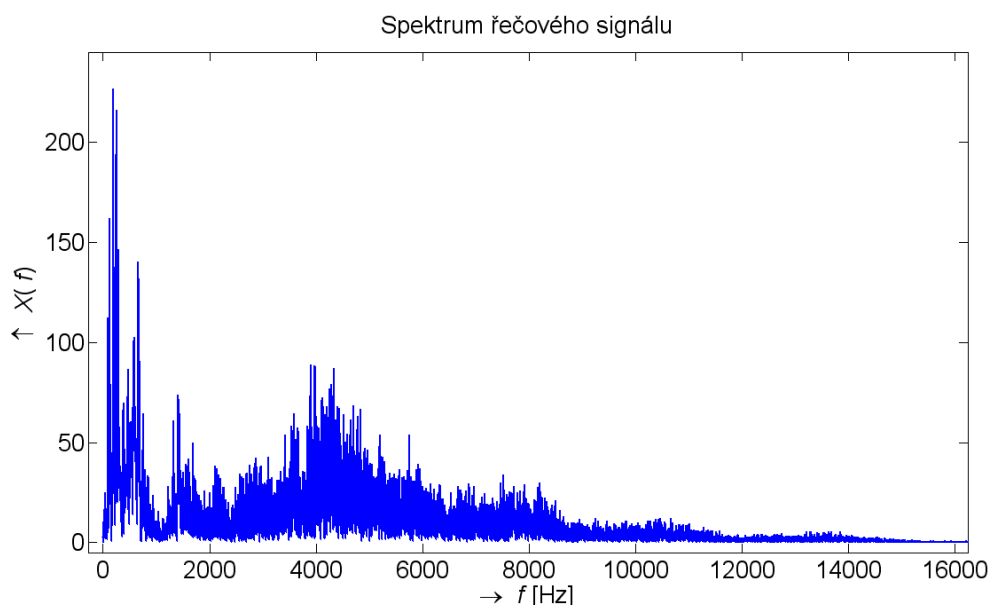
Akustický signál je vlnění v elastickém prostředí vznikající pohyby struktur hlasového ústrojí. Pro snímání použijeme mikrofon, který řeč převede na elektrický signál. Tento elektrický signál digitalizujeme a budeme reprezentovat jako posloupnost řečových vzorků viz. modulace PCM 5.1. Akustická forma má oproti fonetické formě obrovskou nadbytečnost, nese i informace o osobnosti řečníka, dále barvu hlasu, náladu, prostředí atd.

V časové oblasti se řečový signál reprezentuje pomocí časového průběhu vlny, který zobrazuje vývoj amplitudy v čase viz. obr. 4.3



Obr. 4.3: Časový průběh slova „tříška“.

Frekvenční oblast je vhodná k popisu spektrálních vlastností řeči. K výpočtu spektrálních charakteristik řeči se používá Fourierova transformace.



Obr. 4.4: Spektrum slova „tříska“.

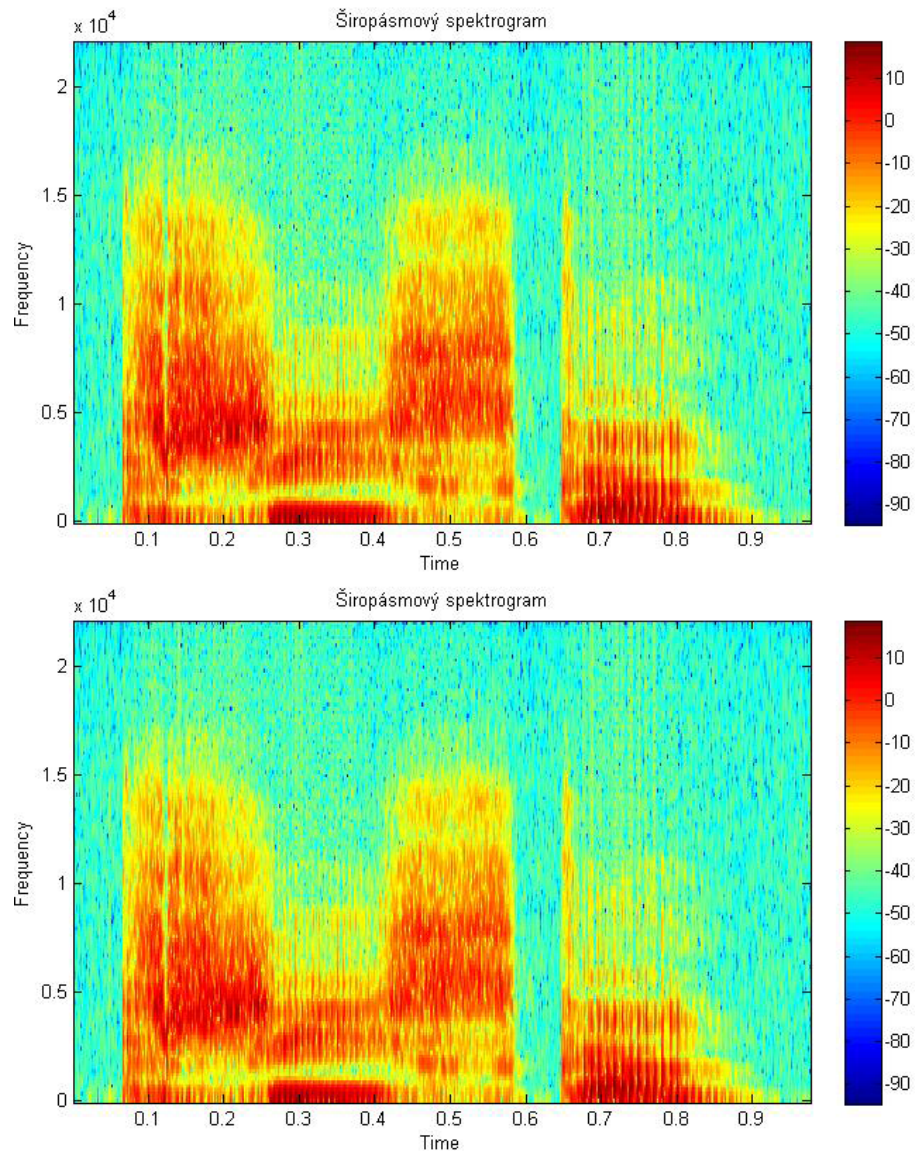
Spektrogramy zobrazují časové a spektrální vlastnosti řeči. Představují časově frekvenční reprezentaci řečového signálu - zachycují vývoj modulového spektra řeči v čase pomocí různých úrovní barev v dvourozměrném grafu. Podle toho, v jaké oblasti chceme zachytit detaily, se používají dva typy spektrogramů. Širokopásmové a úzkopásmové.

Širokopásmové spektrogramy používají při výpočtu krátká váhová okénka asi 5–10 ms. Zobrazují detaily v časové oblasti, ale nejsou schopny zobrazit jemné frekvenční detaily. Můžeme sledovat jednotlivé hlásky lišící se barvou tónů (témbrem), která je dána vlivem spolupodílejících se vedlejších frekvencí - formantů, které vznikají rezonancí dutin (ústní, hltanová, nosní). Širokopásmové spektrogramy jsou základem pro spektrální analýzu signálů.

Úzkopásmové využívají delší váhové okénko kolem 30 ms. Zobrazují detaily ve frekvenční oblasti. Lze z nich vyčíst jednotlivé harmonické frekvence odpovídající frekvenci základního hlasivkového tónu.

4.4 Model vytváření řeči

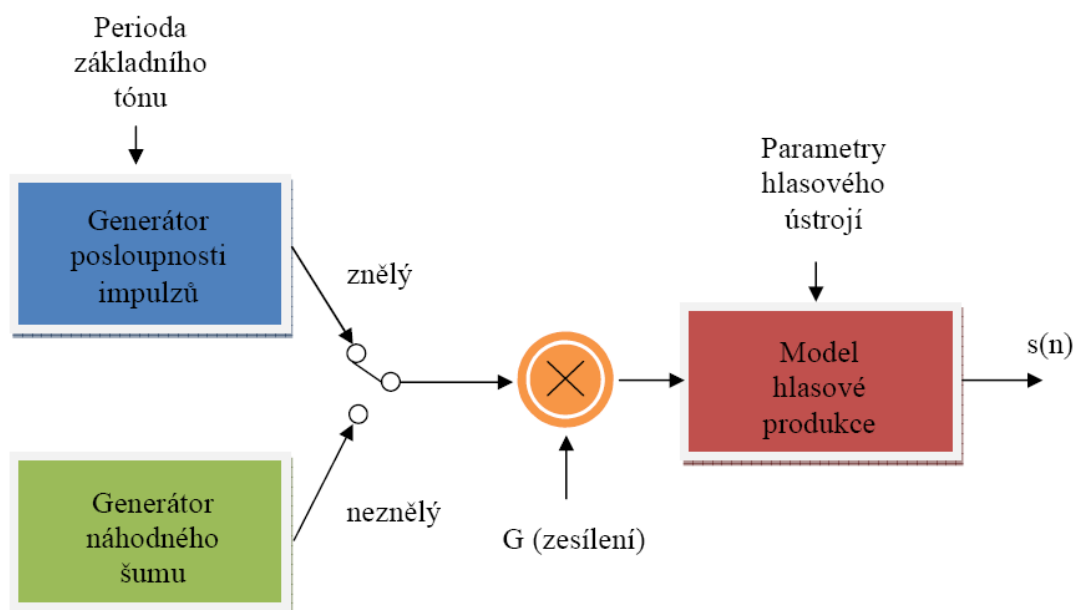
Při počítačovém zpracování řeči bývá obvykle užitečné jistým způsobem modelovat procesy produkce i vnímání řeči a řeč následně reprezentovat pomocí parametrů zvoleného modelu. Cílem modelování produkce řeči je nalézt matematické vztahy,



Obr. 4.5: Šírokopásmový a úzkopásmový spektrogram slova „tříska“.

které by mohly být využity pro reprezentaci akustických fyzikálních dějů spojených s touto produkcí, a to s požadavkem minimální složitosti a maximální přesnosti modelu. Z výpočetního hlediska by bylo ideální, kdyby takový model byl lineární a časově neměnný. Lidská řeč je ale souvislý časově proměnný proces, který vykazuje i některé nelineární charakteristiky vyplývající z toho, že hlasivky nejsou při mluvení úplně odpojeny od hlasového traktu. Nelze se proto divit, že zatím nebyl předložen ani fyziology, ani akustiky či specialisty na zpracování signálů jediný univerzální model, který by respektoval všechny uvedené požadavky. Na 4.6 je zobrazen zjednodušený model, který se používá například při analýze LPC.

Uvážíme-li, že pro mnohé zvuky řeči je možné počítat s tím, že typ buzení a vlast-



Obr. 4.6: Model vytváření řeči. [4]

nosti hlasového traktu zůstávají téměř konstantní v časovém úseku 10–30 ms. Odpovídající model produkce se tedy může skládat z lineárního modelu hlasového traktu s pomalu měnícími parametry, který je buzen vhodným budícím signálem. Signál buzení přitom představuje buď periodický sled pulsů pro znělou řeč, či „šumový“ signál pro neznělou řeč.

Model hlasové produkce by měl především respektovat vlastnosti dutiny hrdelní, ústní a pro nosovky i dutiny nosní. Parametry modelu závisí na tvaru, průřezu a délce těchto dutin a jsou většinou pro každého člověka i při vyslovování téhož zvuku rozdílné a významě se mění i u téhož řečníka při vyslovování různých zvuků. To je dáno kromě jiného například změnou objemu dutiny hrdelní a ústní, pohybem jazyka a podobně.

Uvedený model produkce řeči se přednostně využívá k získání širokého spektra reprezentací řečového signálu, například jednoduché kódování tvaru vlny nebo odhad parametrů modelu.

5 ANALÝZA ŘEČOVÉHO SIGNÁLU

5.1 Pulsní kódová modulace (PCM)

Analogové kmity pro další zpracování je zapotřebí převést do číslicového tvaru tak, aby byl reprezentován posloupností číselných údajů. Tento proces, pulsni kódová modulace nebo též digitalizace, zahrnuje provedení dvou kroků, a to vzorkování a kvantizace s kódováním. Má-li řeč být plně reprezentována číslicovým signálem, musí být splněn Nyquistův-Shannonův-Kotelnikovův teorém, že vzorkovací kmitočet F_s musí být minimálně dvakrát větší než je nejvyšší kmitočet F_m obsažený ve spektru signálu. Vzorkovací kmitočet zvukových karet PC se používá $F_s = 44.1$ kHz pro HIFI kvalitu. Každý vzorek je kvantován jednou m možných kvantovacích hladin a k jejich vyjádření potřebujeme $N = \log_2 m$ bitů. Zvukové karty PC pracují obvykle s $N = 16$ bitů. Výsledná informační rychlost c je vzorkovací kmitočet F_s násobený počtem N bitů.

$$c = F_s * N = 44100 * 16 = 705 \text{ kbit/s.} \quad (5.1)$$

5.2 Fourierova transformace

Fourierova transformace je vyjádření časově závislého signálu pomocí harmonických signálů. Slouží pro převod signálů z časové oblasti do oblasti frekvenční. Metodou PCM byly získány vzorky $s(t)$ řečového signálu. Fourierova transformace $S(\omega)$ je pak definována vztahem:

$$S(\omega) = \int_{-\infty}^{\infty} s(t)e^{-i\omega t} dt. \quad (5.2)$$

Funkci $S(\omega)$ říkáme spektrální funkce, nebo krátce spektrum. Funkce je komplexní a nedá se vypočítat (nekonečná). V praxi ji však dokážeme pouze odhadnout pomocí diskrétní Fourierovy Transformace 5.3.

$$S(k) = \sum_{n=0}^{N-1} s(n)e^{-i\omega n k/N}, k = 0, 1, 2, \dots, N - 1. \quad (5.3)$$

Výpočet DFT podle definičního vztahu vyžaduje N^2 komplexních součinů a N^2 komplexních součtů. Toto množství operací výrazně snižuje možnost aplikace DFT na výpočty v reálném čase. Existuje však efektivní algoritmus výpočtu DFT, nazývaný rychlá Fourierova transformace (FFT - Fast Fourier Transform).

Získané koeficienty se využívají zejména ve spektrálních analyzátoch řeči nebo jsou prvním stupněm zpracování různých parametrizačních technik v systémech rozpoznávání řeči.

5.3 Lineární predikční analýza

Lineární prediktivní kódování (LPC) je jednou z nejefektivnějších metod analýzy akustického signálu. Je to metoda, která se snaží opět na krátkodobém základu odhadnout přímo z řečového signálu parametry modelu vytváření řeči. Atraktivnost metody spočívá v její schopnosti zabezpečit odhad uvedených parametrů při relativně přijatelné výpočetní zátěži.

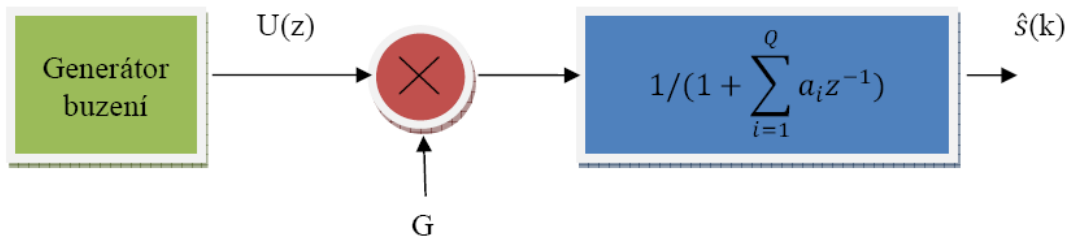
Model vytváření řeči se skládá ze systému a časově proměnným přenosem a generátoru budících funkcí, který budí systém při vytváření znělých zvuků posloupností impulsů a při vytváření neznělých zvuků náhodným šumem. Princip metody je založen na předpokladu, že k -tý vzorek signálu $s(k)$ lze popsat lineární kombinací Q předchozích vzorků a buzení $u(k)$, kde G je koeficient zesílení a Q je řád modelu.

$$s(k) = - \sum_{i=1}^Q a_i s(k-i) + Gu(k). \quad (5.4)$$

Přenosovou funkci modelu $H(z)$ lze zapsat ve tvaru:

$$H(z) = \frac{G}{1 + \sum_{i=1}^Q a_i z^{-i}} = \frac{1}{A(z)}. \quad (5.5)$$

Kde polynom $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_Q z^{-Q}$ a má řád Q . Pro vzorkovací kmitočet 8 kHz se často volí řád $Q = 10$ (pro vyšší kmitočty až 20). Proces modelování můžeme znázornit na 5.1



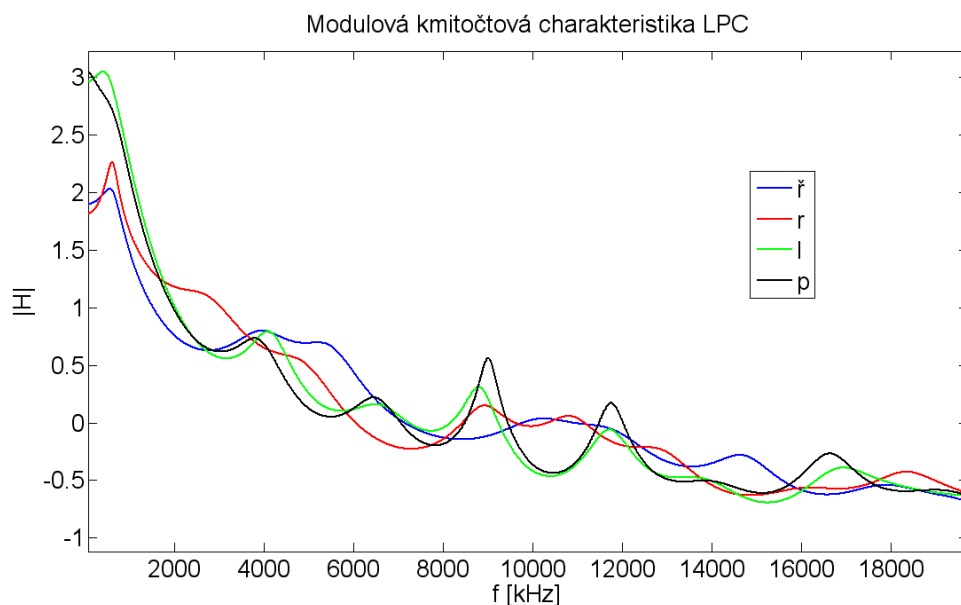
Obr. 5.1: Model vytváření řeči s lineárním filtrem

Signál $s(k)$ je dán lineární kombinací několika předchozích vzorků, považujeme jej za předpověď skutečného vzorku $s(k)$:

$$\hat{s}(k) = - \sum_{i=1}^Q a_i s(k-i). \quad (5.6)$$

Sledovanými parametry jsou zde koeficienty a_i číslicového filtru a koeficient zesílení G . Při neznámém členu Gu vzniká chyba predikce $e(k)$ mezi skutečnou $s(k)$ a předpovězenou $\hat{s}(k)$.

$$e(k) = s(k) - \hat{s}(k). \quad (5.7)$$

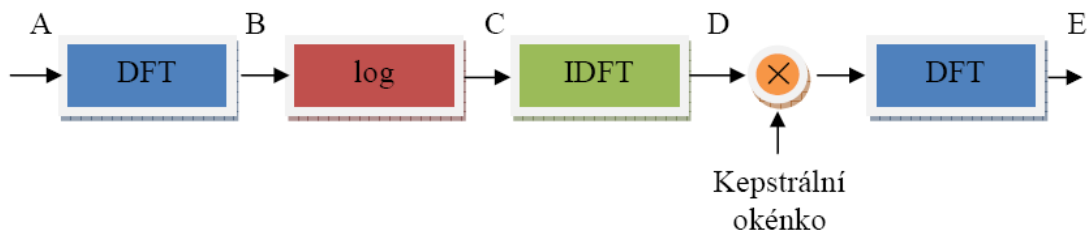


Obr. 5.2: Signálové spektrum pro hlásky „ř,r,l,p“ určené z koeficientů LPC (řád prediktoru $Q = 20$)

5.4 Kepstrální analýza

Řečové kmity mohou být modelovány na krátkodobém základě jako odezva lineárního systému na buzení sestávající ze sledu pulsů pro znělou a šumu pro neznělou řeč. Proces kepstrální analýzy je znázorněn na 5.3.

Signál A je přiveden na vstup bloku DFT a výstupní signál B, který je Fourierovou transformací buzení a impulzní odezvy hlasového ústrojí, přichází na blok log. Jeho výstup C je sumou logaritmu transformace buzení a transformace impulzní odezvy hlasového ústrojí. Pro názornost můžeme sledovat výsledky analýzy pro hlásku ř na 5.4. Křivka C se stává z pomalu se měnících složek způsobených přenosy hlasového ústrojí a rychle se měnící periodických složek způsobených periodickým buzením. Pomalu se měnící část logaritmu veličiny vytváří pomalu se měnící část kepra D a rychle se měnící periodická složka logaritmu veličiny se projevuje velkými špičkami, které jsou umístěny v časových okamžicích úměrných



Obr. 5.3: Postup kepstrální analýzy

periodě základního tónu. Přenosová funkce hlasového ústrojí může být získána tím, že nejdříve odstraníme měnící se složky v signálu C lineární filtrací například vynásobením kepra tzv. kepstrálním okénkem $l(n)$, kde n_0 je vybráno tak, aby bylo menší než základní perioda hlasivkového tónu a pak provedeme DFT, čímž získáme spektrální obálku E.

$$l(n) = \begin{cases} 1 & \text{pro } n < n_0 \\ 0 & \text{pro } n \geq n_0 \end{cases}$$

5.5 Melovské kepstrální koeficienty

Zpracování pomocí melovských kepstrálních koeficientů je navrženo tak, aby do jisté míry respektovalo nelineární vlastnosti vnímání zvuků lidským uchem. Kepstrální koeficienty se snaží „kompenzovat“ zejména nelineární vnímání frekvencí, a to využitím banky trojúhelníkových pásmových filtrů s lineárním rozložením frekvencí, jež je definována vztahem 5.8.

$$f_m = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \quad (5.8)$$

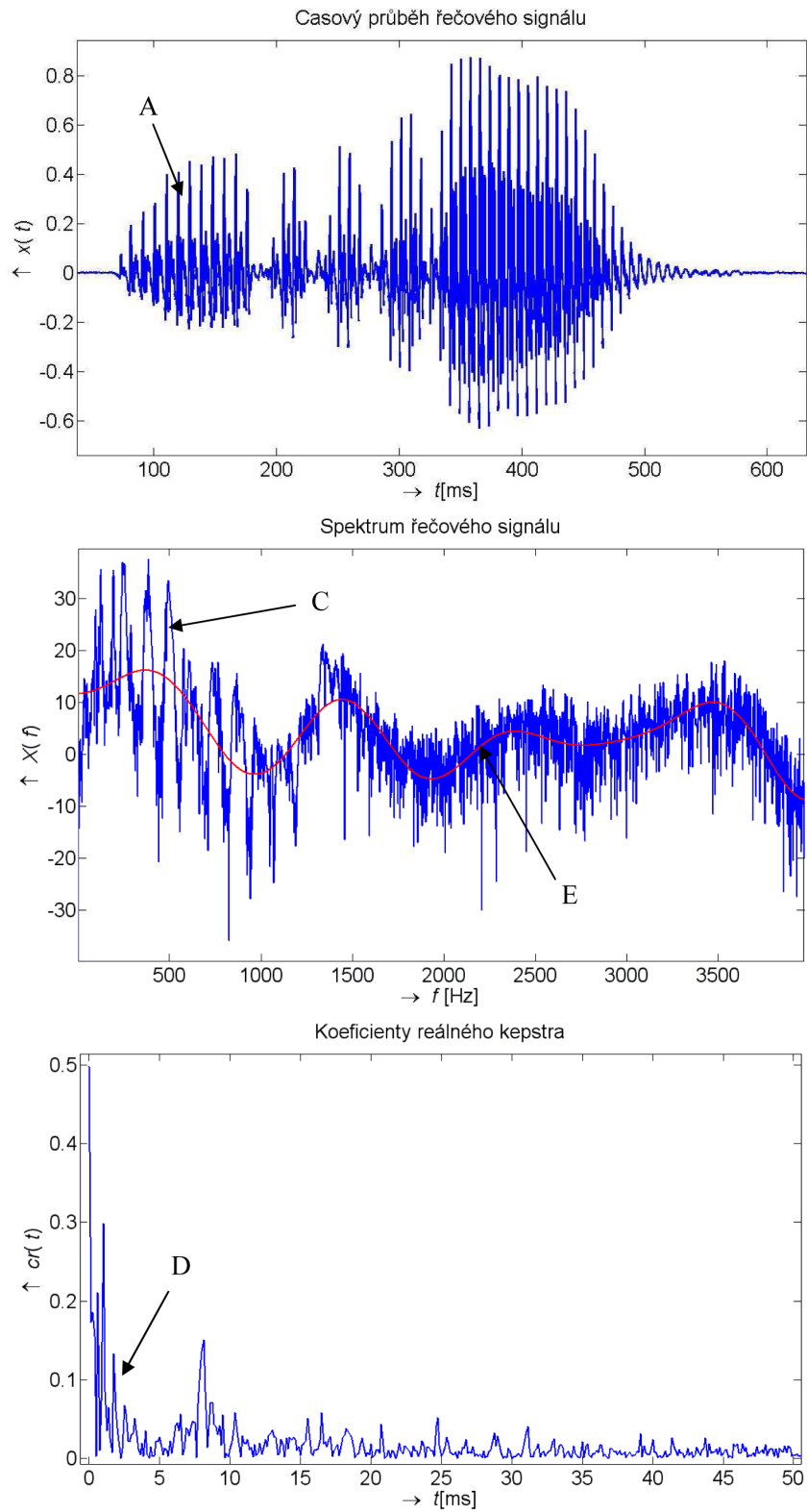
Frekvence v lineární škále je $f[Hz]$ a odpovídající frekvence v nelineární škále je $f_m[mel]$. Lineární rozmístění filtrů na melové ose má tedy za následek nelineární rozmístění na standardní kmitočtové ose 5.6.

Proces určení melovských kepstrálních koeficientů lze popsat následujícím postupem.

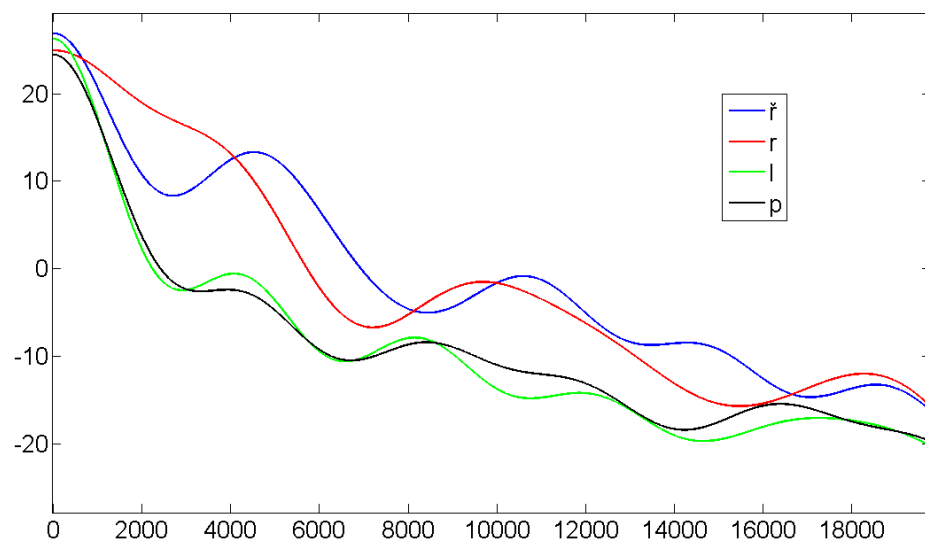
Na vstup systému jsou přiváděny vzorky řečového signálu $s(k)$. Je provedena preemfáze signálu a potom je aplikováno nejčastěji Hammingovo okénko.

V dalším bloku zpracování se pomocí FFT vypočte amplitudové spektrum $S(f)$ analyzovaného signálu.

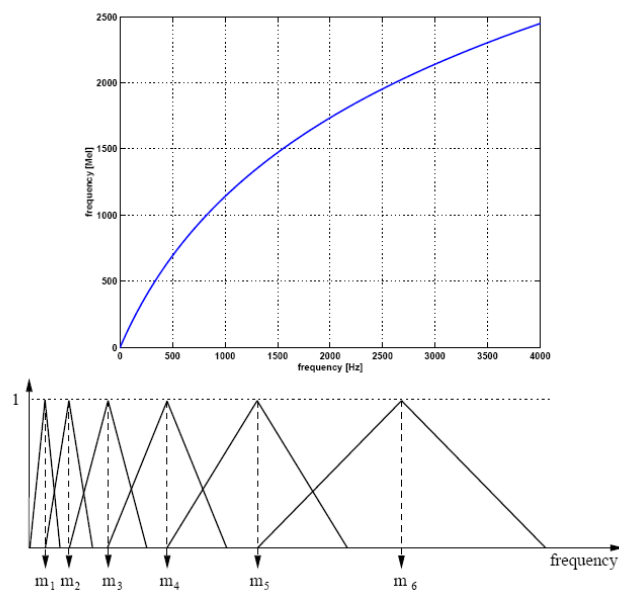
Klíčová část celého procesu zpracování je melovská filtrace. Výpočetní algoritmus této filtrace je realizován bankou trojúhelníkových filtrů podél frekvenční osy s měřítkem v melovské škále. Počet pásem je rozumné volit podle tabulky 5.5



Obr. 5.4: Postup keprální analýzy pro hlásku „i“



Obr. 5.5: Signálové spektrum pro hlásky „ř,r,l,p“ určené z kepstrálních koeficientů



Obr. 5.6: Nelineární úprava kmitočtové osy a výsledné rozmístění filtrů.

Tabulka

Frekvence vzorkování [Hz]	8000	16000	22000	44000
Přenašené pásmo [Hz]	0-4000	0-8000	0-11000	0- 22000
Přenašené pásmo [mel]	0-2146	0-2840	0-3174	0-3921
Počet pásem	15	20	22	27

Tab. 5.1: Typické hodnoty počtu filtrů. [4]

Další krok spočívá ve výpočtu logaritmu výstupů jednotlivých filtrů. Logaritmováním koeficientů se vkládá do procesu zpracování důležitý příznak kepstrální analýzy, který jednak příznivě omezí dynamiku signálu podobně jako to dělá lidské ucho a dále přinese i užitečné předpoklady pro případné další zpracování signálu.

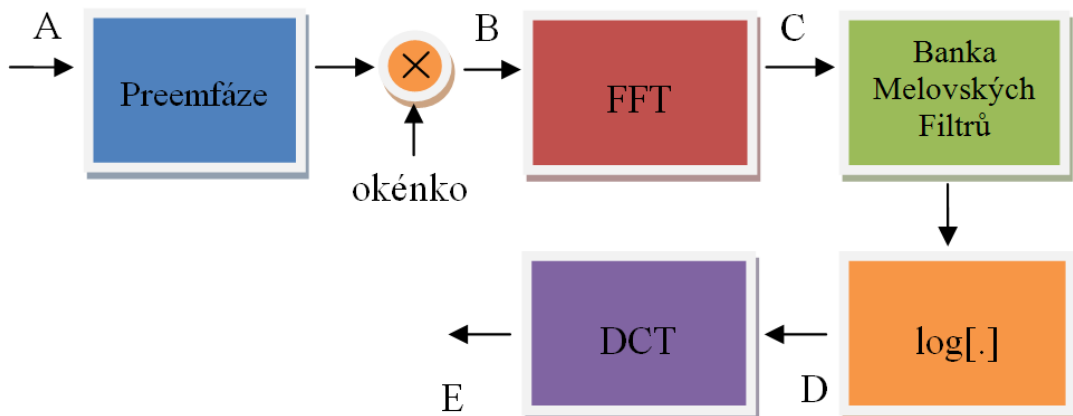
Posledním krokem při výpočtu melovských kepstrálních koeficientů MFCC je provedení diskrétní kosinové transformace DCT.

Celý proces je schematicky znázorněn na 5.7.

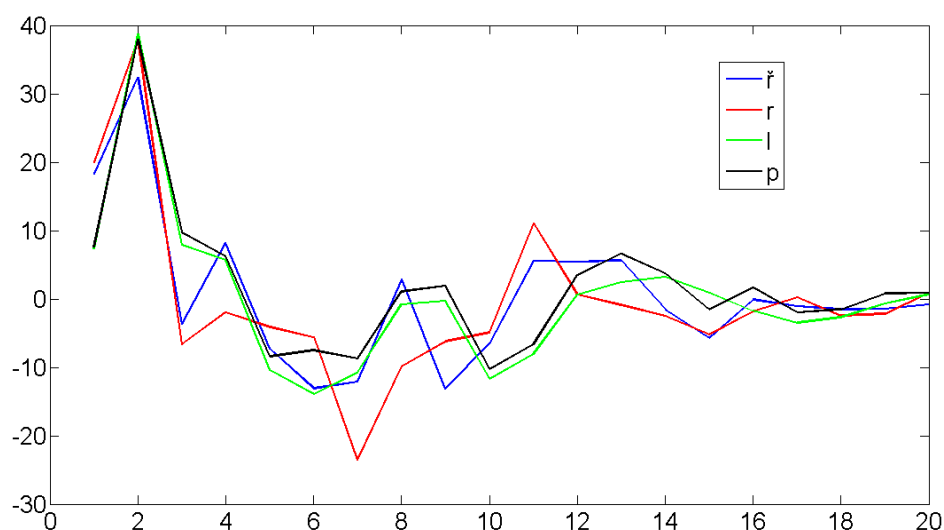
Protože kepstrální koeficienty s vyššími indexy nabývají standardně nižších hodnot, je z praktických důvodů užitečné provést liftering podle vztahu 5.9, abychom získali zobrazení parametrů s přibližně podobně velkými hodnotami.

$$c_{lift}(n) = [1 + \frac{L}{2} \sin(\pi n/L)]c(n). \quad (5.9)$$

L je váha lifteringu s typickou hodnotou L=22 [4].



Obr. 5.7: Postup výpočtu melovských kepstrálních koeficientů.



Obr. 5.8: Melovské keprální koeficienty MFCC pro hlásky „ř,r,l,p“.

5.6 Analýza výslovnosti hlásek „r“ a „ř“

Na obrázku 5.9 se můžeme podívat na výsledky analýzy všech zmíněných metod rozpoznávání řeči. Porovnávám zde správně vyslovené hlásky „r“ a „ř“ (vyznačeno vždy modrou barvou) s nesprávnou artikulací (vyznačeno červenou barvou). Hlásky jsou vždy vysloveny dvakrát z důvodu vyšší objektivity.

5.6.1 Hláška „r“

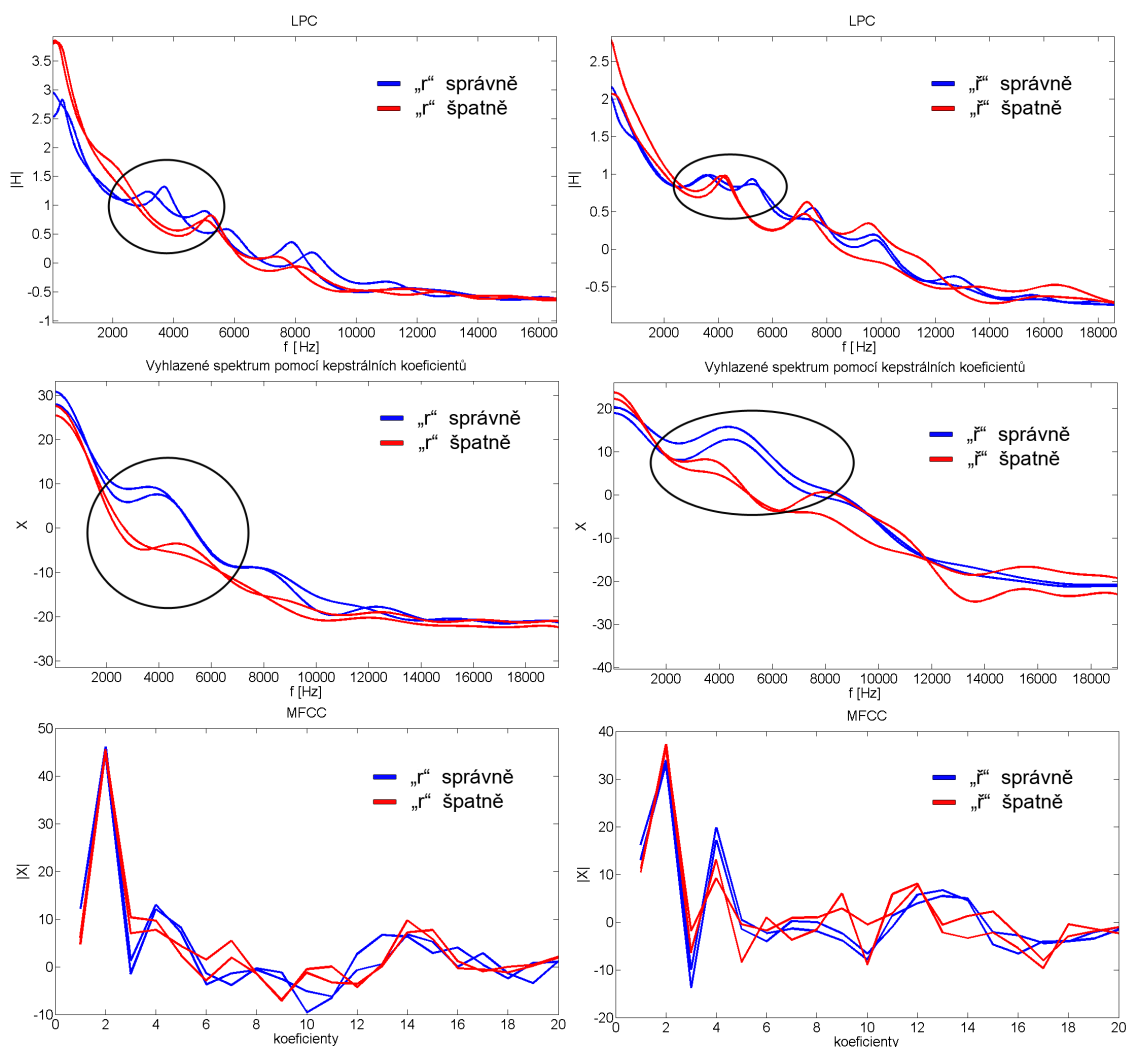
U LPC analýzy můžeme vidět tři formanty. U nesprávné výslovnosti dochází k vymizení prvního formantu oproti správné. Toto vymizení je dáno špatnou vibrací špičky jazyka při artikulaci. Ve vyhlazeném spektru pomocí keprálních koeficientů lze pozorovat pokles hodnoty velikosti prvního formantu a u výpočtu MFCC nastává výrazná změna hodnot koeficientů 3–13.

5.6.2 Hláška „ř“

LPC analýza ukazuje, že u této hlásky se objevuje více formantů a právě v oblasti mezi prvním a druhým je vidět vymizení jednoho formantu. Tento výsledek je opět způsoben špatnou vibrací špičky jazyka podobně jako u hlásky „r“. Ve vyhlazeném spektru pomocí keprálních koeficientů je také pokles hodnoty prvního formantu a koeficienty MFCC se liší od 5–14.

5.6.3 Porovnání metod

Při studii analýz a porovnání jejich výsledků jsem došel k závěru, že LPC analýza umožňuje dobře rozpoznávat formanty. Tyto hodnoty ale nelze efektivně porovnávat s jiným řečníkem. Podobný závěr je i u keprální analýzy, jen s rozdílem, že dochází k většímu rozdílu hodnot ve spektru. U výpočtu MFCC není zřetelně vidět formantová struktura. Pro více řečníků je tato metoda ale nejvhodněji využitelná, protože nerozlišuje do určité míry specifika řečníků. Tudíž by byla nejefektivnější z hlediska znělosti a nejlépe by šel vytvořit vzor promluvy, se kterou by se poté dala porovnávat správná artikulace.



Obr. 5.9: Porovnání správné výslovnosti s artikulací vadou pro hlásky „r“ a „ř“.

6 ZÁVĚR

V bakalářské práci Analýza a rozpoznání logopedických vad v řečovém projevu jsem popsal logopedickou vadu dyslalií a úvod do zpracování řečového signálu. Předvedl jsem základy zpracování řečového signálu. Dále jsem provedl analýzy zpracování řeči - lineárně predikční (LPC), kepstrální a výpočet melovských kepstrálních koeficientů (MFCC). Z důvodu velké náročnosti v programu MATLAB jsem se v závěru zaměřil pouze na vyhodnocení nesprávné výslovnosti hlásek „r“ a „ř“.

LITERATURA

- [1] ČERNOCKÝ, J. *Zpracování řečových signálů*. Brno: Skripta VUT, 2006. 183 s. (elektronické texty pro bakalářské studium).
- [2] KUTÁLKOVÁ, D. *Logopedická prevence, průvodce vývojem dětské řeči*. Praha: Portál, 1996. ISBN 8071786675.
- [3] LECHTA, V. a kol. *Diagnostika narušené komunikační schopnosti*. Praha: Portál, 2003. ISBN 8071788015
- [4] PSUTKA, J., MULLER, L., MATOUŠEK, J., RADOVÁ, V. *Mluvíme s počítačem česky*. Praha: ACADEMIA, 2006. 563 s. ISBN 80-20013-09-1.
- [5] SMÉKAL, Z. *Číslíkové zpracování signálu (MCSI)*. Brno: Skripta VUT, 2007. 183 s. (elektronické texty pro magisterské studium).
- [6] ZAPLATÍLEK, K., DOŇAŘ, B. *Matlab - začínáme se signály*. Ben, 2006. ISBN 8073002000.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

f_{vz} vzorkovací kmitočet

FFT Rychlá fourierova transformace

DFT Diskrétní fourierova transformace

DCT Diskrétní kosinova transformace

PCM Pulsní kódová modulace

LPC Lineární prediktivní kódování

MFCC Melovské keprální koeficienty

SEZNAM PŘÍLOH

A Přílohy	37
A.1 Výpočet LPC pro hlásky „ř,r“	37
A.2 Výpočet vyhlazené spektrální obálky pomocí keprálních koeficientů pro hlásky „ř,r“	38
A.3 Výpočet melovských keprálních koeficientů pro hlásky „ř,r“	39

A PŘÍLOHY

A.1 Výpočet LPC pro hlásky „ř,r“

```
% načtení hlásek .wav

[x1,fvz] = wavread('silvarr.wav');
[x2,fvz] = wavread('silvar.wav');

% výpočet LPC

N = length( x1);
f = (fvz/2).*(1:N/2)/(N/2);

a = lpc(x1,20);
Q = fft([a zeros(1,N-length(a))]);
lpcout = 20.*log10(abs(Q(1:N/2)));
Y1 = -lpcout./max(lpcout);

N2 = length( x2);
f2 = (fvz/2).*(1:N2/2)/(N2/2);

a2 = lpc(x2,20);
Q2 = fft([a2 zeros(1,N2-length(a2))]);
lpcout2 = 20.*log10(abs(Q2(1:N2/2)));
Y2 = -lpcout2./max(lpcout2);

% zobrazení

figure (1);
plot(f,Y1,'b');
xlabel('f [kHz]');
ylabel('|H|');
title('Modulová kmitočtová charakteristika LPC');
hold on
plot(f2,Y2,'r');
```

A.2 Výpočet vyhlazené spektrální obálky pomocí keprálních koeficientů pro hlásky „ř,r“

```
% načtení hlásek .wav

[x1, fvz, bits] = wavread('silvarr');
[x2, fvz, bits] = wavread('silvar');

% výpočet kepratra

N1 = length( x1);
t1 = (0:N1-1)/fvz;
X1 = fft( x1);
f1 = (0:N1-1)/N1*fvz;
cep1 = ifft( log( abs( X1)));
L1 = 13*fvz/fvz;
w1 = zeros( size( cep1));
w1([(1:L1),(N1-L1+2:N1)]) = ones( 2*L1-1,1);
cep1 = cep1.*w1;
tmp1 = fft( cep1);
tmp1 = real( tmp1);
Xr1 = exp( tmp1);

N2 = length( x2);
t2 = (0:N2-1)/fvz;
X2 = fft( x2);
f2 = (0:N2-1)/N2*fvz;
cep2 = ifft( log( abs( X2)));
L2 = 13*fvz/fvz;
w2 = zeros( size( cep2));
w2([(1:L2),(N2-L2+2:N2)]) = ones( 2*L2-1,1);
cep2 = cep2.*w2;
tmp2 = fft( cep2);
tmp2 = real( tmp2);
Xr2 = exp( tmp2);
```

```

% zobrazení

figure(1);
title( 'Spektrum řečového signálu');
xlabel( 'f [Hz]');
ylabel( 'X');

vel_1=length(f1)/2;
vel_2=length(f2)/2;

plot( f1(1:vel_1), 20*log10(Xr1(1:vel_1)), 'b');
hold on;
plot( f2(1:vel_2), 20*log10(Xr2(1:vel_2)), 'r');

```

A.3 Výpočet melovských kepstrálních koeficientů pro hlásky „ř,r“

```

% Načtení hlásek .wav

[x_1, fvz] = wavread('silvarr');
[x_2, fvz] = wavread('silvar');

% Načtení funkce pro výpočet Melovských koeficientů

mfcc_1 = mfcco(x_1, fvz);
mfcc_2 = mfcco(x_2, fvz);

% Srovnání délky vektorů

vel_1=length(mfcc_1);
vel_2=length(mfcc_2);

% Vykreslení

plot(1:vel_1,mfcc_1,'b')
hold on;
plot(1:vel_2,mfcc_2,'r')

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function mfcco = mfcc(sig, fsamp)

% Srovnání délek

nfft = length(sig) + mod(length(sig),2);

% Parametry

cepnum      = 20;           % Množství cepstralních koeficientů
liftercoe   = 22;           % Váha lifteringu
numchan     = 26;           % množství kanálů z MEL filtrační banky

% Inicializace MEL filtrační banky
fbank = initfiltb(length(sig), numchan, fsamp, nfft);
% Liftering
lifter = (1 + (liftercoe/2)*sin((pi/liftercoe)*(1:cepnum)) );
% FFT
fftm = abs(fft(sig));
% MEL filtrace
fb = fbank*fftm(1 : (nfft/2));
% log
logfb = log(fb);
% DCT
mfcco = dct(logfb);
% Úprava s kepstrem
mfcco = mfcco(1:cepnum);
mfcco = mfcco .* lifter';

function mels = mel(freq)

% Přepočítání Hz na Mel

mels = 2595 * log( 1 + (freq/700) );

function fbank = initfiltb(framelen,numchan,fsamp,nfft)

```

```

% Inicializace trojúhelníkových melfiltrů

fftfreqs = ((0:(nfft/2-1))/nfft)*fsamp;
melfft = mel(fftfreqs);
mel0 = 0;
mel1 = mel(fsamp/2);
melmid = ((1:numchan)/(numchan+1))*(mel1-mel0) + mel0;
fbank = zeros(numchan,nfft/2);
for k = 2:(nfft/2)
    chan = max([ 0 find(melfft(k)>melmid) ]);
    if(chan==0)
        fbank(1,k) = (melfft(k)-mel0)/(melmid(1)-mel0);
    elseif(chan==numchan)
        fbank(numchan,k) = (mel1-melfft(k))/(mel1-melmid(chan));
    else
        fbank(chan,k) = (melmid(chan+1)-melfft(k))
            /(melmid(chan+1)-melmid(chan));
        fbank(chan+1,k) = 1-fbank(chan,k);
    end
end

% Normalizace pro různou energii z každé banky filtrů
fbank = fbank ./ repmat(sum(fbank,2), 1, size(fbank,2));

```