

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2021

BcA. Petr Kupka



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV RADIOELEKTRONIKY

DEPARTMENT OF RADIO ELECTRONICS

DETEKCE ALKOHOLU V ŘEČOVÉM SIGNÁLU

DETECTION OF ALCOHOL IN SPEECH SIGNAL

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

BcA. Petr Kupka

VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. Milan Sigmund, CSc.

BRNO 2021

Diplomová práce

magisterský navazující studijní program **Elektronika a komunikační technologie**

Ústav radioelektroniky

Student: BcA. Petr Kupka

ID: 89136

Ročník: 2

Akademický rok: 2020/21

NÁZEV TÉMATU:

Detekce alkoholu v řečovém signálu

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s problematikou zpracování řečových signálů a vlivu alkoholu na řečové parametry. Zpracujte rešerši významných publikací a patentů pojednávajících o vlivu alkoholu na parametry řečových signálů. Vybrané algoritmy na detekci alkoholu v kmitočtové a cepstrální oblasti naprogramujte a ověřte na reálném řečovém signálu z existující databáze. Určete nutnou minimální úroveň alkoholu pro správnou činnost detekčních algoritmů.

Vytvořte databázi profesionálně předstírané a skutečné alkoholické řeči s referenčními údaji a poslechovými testy. Pomocí vytvořených algoritmů určete, které fonetické jednotky jsou nejvíce citlivé na alkohol. Algoritmy optimalizujte z pohledu fonetické struktury, kvality a délky testovaných řečových signálů. Vytvořte ucelený program na detekci alkoholu v řečovém signálu. Spolehlivost detekce statisticky vyhodnoťte.

DOPORUČENÁ LITERATURA:

[1] PSUTKA, J., MÜLLER, Z., MATOUŠEK, J., RADOVÁ, V. Mluvíme s počítačem česky. Praha: Academia, 2006.

[2] SIGMUND, M. Rozpoznávání řečových signálů. Skriptum FEKT VUT v Brně. Brno: MJ servis, 2007.

Termín zadání: 8.2.2021

Termín odevzdání: 20.5.2021

Vedoucí práce: prof. Ing. Milan Sigmund, CSc.

prof. Ing. Tomáš Kratochvíl, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

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

Abstrakt

Diplomová práce Detekce alkoholu v řečovém signálu nejdříve popisuje vliv alkoholu na lidský organismus. Další část se zabývá způsobem, jakým se dají získat parametry, kterými se popisuje řečový signál. Stručně pohlíží na dřívější práce a patenty zabývající se detekcí alkoholu v řečovém signálu. V poslední části se zabývá popisem získání hlasové databáze alkoholické řeči. Popsána analýza je implementována do aplikace, která umožňuje vykreslit změny parametrů hlasu při intoxikaci.

Klíčová slova

formanty, základní hlasový tón, LPC, kepstrum, ALC

Abstract

The diploma thesis Detection of Alcohol in Speech Signal first describes the effect of alcohol on the human body. The second part deals with ways to obtain parameters that describe the speech signal. The third part provides a brief overview of previous case studies and patents focused on the detection of alcohol in the speech signal. The fourth part presents the collected own database of voice recordings and developed software application for the analysis of intoxicated speech. The final part describes the measured changes in speech signal parameters that indicate alcohol intoxication.

Keywords

formants, pitch, LPC, cepstrum, ALC

Bibliografická citace

KUPKA, Petr. *Detekce alkoholu v řečovém signálu [online]*. Brno, 2021 [cit. 2021-05-09]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/133607>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky. Vedoucí práce Milan Sigmund.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta:	<i>Petr Kupka</i>
VUT ID studenta:	<i>89136</i>
Typ práce:	<i>Diplomová práce</i>
Akademický rok:	<i>2020/21</i>
Téma závěrečné práce:	<i>Detekce alkoholu v řečovém signálu</i>

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné 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 závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné 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 části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne: 19. květen 2021

podpis autora

Poděkování

Děkuji vedoucímu diplomové práce prof. Ing. Milanu Sigmundovi, CSc. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce, mé ženě Sabině Kupkové za podporu a mé dceři Anně Kupkové za smysl mého konání.

V Brně dne: 19. květen 2021

podpis autora

Obsah

SEZNAM SYMBOLŮ A ZKRATEK	2
SEZNAM OBRÁZKŮ	3
SEZNAM TABULEK	5
ÚVOD	6
1. ALKOHOL A JEHO VLIV [1]	7
1.1 ÚČINKY KONCENTRACE ALKOHOLU V KRVI	8
1.2 DOBA ÚČINKU ALKOHOLU.....	9
1.3 STANOVENÍ ALKOHOLU V ORGANISMU [3].....	9
1.4 UŽITÁ METODA ZJIŠTĚNÍ ALKOHOLU V ORGANISMU.....	10
2. ZPRACOVÁNÍ ŘEČOVÝCH SIGNÁLŮ	11
2.1 VYTVÁŘENÍ ŘEČI ČLOVĚKEM [4]	11
2.2 FONETICKÝ POPIS ČESKÉHO JAZYKA [4].....	13
2.3 ZPRACOVÁNÍ ŘEČOVÉHO SIGNÁLU [6]	14
2.4 PARAMETRY ŘEČOVÉHO SIGNÁLU	15
3. VÝZNAMNÉ PUBLIKACE	27
3.1 ŘEČOVÉ KORPUSY.....	27
3.2 PARAMETRY V ČASOVÉ OBLASTI	28
3.3 PARAMETRY VE FREKVENČNÍ OBLASTI	31
3.4 KMITOČET ZÁKLADNÍHO TÓNU ŘEČI.....	32
3.5 PATENTY	34
4. IMPLEMENTACE	36
4.1 TEST ALKOHOLU	36
4.2 POUŽITÁ TECHNIKA PRO ZÍSKÁNÍ DATABÁZE	37
4.3 DATABÁZE NAHRÁVEK	37
4.4 PROGRAM ALCTEST.....	42
5. VÝSLEDKY NAMĚŘENÝCH HODNOT	51
5.1 ENERGIE NAHRÁVKY.....	51
5.2 ZÁKLADNÍ HLASOVÝ TÓN F_0	52
5.3 FORMANTOVÉ PŘÍZNAKY	56
5.1 KEPSTRÁLNÍ KOEFICIENTY	59
5.2 JITTER A SHIMMER.....	63
5.3 POČET PRŮCHODŮ NULOU	68
5.4 PŘEDSTÍRANÁ OPILÁ ŘEČ	69
6. ZÁVĚR	76
LITERATURA	78

SEZNAM SYMBOLŮ A ZKRATEK

Zkratky:

ACM	Audio Compression Manager
ADH	alkoholdehydrogenáza
AKF	autokorelační funkce
ALC	Alcohol Language Corpus
AMDF	Average Magnitude Difference Function
BAC	Blood Alcohol Concentration
BAL	Blood Alcohol Level
DFT	diskrétní Fourierova transformace
FEKT	Fakulta elektrotechniky a komunikačních technologií
GNU	General Public License
IDFT	inverzní diskrétní Fourierova transformace
LPC	linear predictive coding
MATLAB	MATrix LABoratory
MAUS	Munich Automatic Segmentation
NTSB	National Transportation Safety Board
PCM	pulzní kódová modulace
SNR	odstup signálu od šumu
VUT	Vysoké učení technické v Brně
WAV	Waveform audio format

Symboly:

‰	promile	[g/kg]
F_0	základní hlasový tón	[Hz]
F_1, F_2, \dots, F_n	formanty	[Hz]
B_i	šířka formantů	[Hz]
M_1, M_2, \dots, M_m	antiformanty	[Hz]
E	energie signálu	[J]
SD	standardní odchylka	
f_m	medián	
Δf_{qq}	čtvrt kvantilový rozsah	

SEZNAM OBRÁZKŮ

1.1	Nehody zaviněné pod vlivem alkoholu a jejich následky [1]	7
2.1	Hlasový trakt člověka [4]	11
2.2	Hrtan a hlasivky [4]	12
2.3	Formanty ve spektru samohlásek [5]	13
2.4	Rozdělení fonémů do skupin podle energie řečového signálu [7]	16
2.5	Blokové schéma metody Center-Clipping [7]	17
2.6	Jednoduché parametry řeči založené na vlastnostech spektra [7]	17
2.7	Stacionární, spojité a přerušené spektrum různých fonémových skupin [7]	18
2.8	Vliv řádu prediktoru M na LPC spektrum [7]	23
2.9	Vliv normalizace dlouhodobým LPC spektrem na spektrum hlásky „a“ [7]	24
2.10	Systém k určení kepru řečového signálu [7]	25
2.11	Logaritmické spektrum a odpovídající kepru znělého segmentu [7]	25
2.12	Četnost průchodu nulou u znělých a neznělých úseků řeči. [7]	26
3.1	Trvání segmentů řeči ve frázi Exxon Valdez v různých časech nahrávání. a) Kumulativní doby trvání označující celkové prodloužení doby trvání. b) Doba trvání seskupené podle segmentů, které ukazují, u kterých segmentů se doba trvání zvýšila v době nehody [15]	30
3.2	a) Délka slova Valdez z pásek NTSB ve srovnání se stejným slovem vytvořeným v podobné pozici v televizním rozhovoru. b) Trvání fráze třináct a šestnáct ze záznamů pořízených třikrát v době nehody [15]	30
3.3	5 rytmických funkcí pro intoxikovaný (A) a neintoxikovaný (b) projev [19]	31
3.4	Výkonové spektrum a) /s/ a b) /ʃ/ [15]	32
3.5	Průměrná F_0 napříč frázemi Exxon Valdez, třináct a šestnáct a z jedné věty v televizním rozhovoru [15]	33
3.6	F_0 intoxikované (A) a střízlivé (B) řeči dělená pro muže(m) a ženy(f) [19]	34
4.1	Alkohol tester Dräger Alcotest 7410	36
4.2	Použitý mikrofon Audio-Technika PRO 37 a externí zvuková karta M-Audio M-track Solo.	37
4.3	Kmitočtová charakteristika mikrofonu Audio-Technika PRO 37 [37]	37
4.4	Ilustrační fotografie z průběhu nahrávání.	40
4.5	Aplikace ALCTest	42
4.6	LPC spektrum aplikace ALCTest	43
4.7	Základní hlasový tón aplikace ALCTest	44
4.8	Kepru aplikace ALCTest	44
4.9	Keprální koeficienty aplikace ALCTest	45
4.10	Rozložení keprálních koeficientů v aplikaci ALCTest	45
4.11	Spektrogram nahrávky aplikace ALCTest	46
4.12	Zobrazení špiček v nahrávce pro výpočet Jitteru a Shimmeru	46
4.13	Soubor data.mat a datastatic.mat s uloženými parametry	49
4.14	Alkoholová kalkulačka v aplikaci ALCTest	50
5.1	Vliv alkoholu u samohlásky Á na parametr Energie [J] u mužů a, a žen b,	52
5.2	Vliv alkoholu u samohlásky Á na parametr F_0 u mužů	53
5.3	Vliv alkoholu u samohlásky Á na parametr F_0 u žen.	53
5.4	Vliv alkoholu u samohlásky É na parametr F_0 u mužů.	53
5.5	Vliv alkoholu u samohlásky É na parametr F_0 u žen.	54
5.6	Vliv alkoholu u samohlásky Í na parametr F_0 u mužů.	54
5.7	Vliv alkoholu u samohlásky Í na parametr F_0 u žen.	54
5.8	Vliv alkoholu u samohlásky Ó na parametr F_0 u mužů.	55
5.9	Vliv alkoholu u samohlásky Ó na parametr F_0 u žen.	55
5.10	Vliv alkoholu u samohlásky Ú na parametr F_0 u mužů.	55

5.11	Vliv alkoholu u samohlásky Ú na parametr F_0 u žen.	56
5.12	Nejvýraznější změny formantových příznaků u mužů při intoxikaci.	57
5.13	Nejvýraznější změny formantových příznaků u žen při intoxikaci.	58
5.14	Nejvýraznější průměrné změny kepstrálních koeficientů u mužů při intoxikaci.	60
5.15	Nejvýraznější průměrné změny maxim kepstrálních koeficientů u mužů při intoxikaci.	61
5.16	Nejvýraznější průměrné změny kepstrálních koeficientů u žen při intoxikaci.	62
5.17	Nejvýraznější průměrné změny maxim kepstrálních koeficientů u žen při intoxikaci.	63
5.18	Nejvýraznější průměrné změny <i>Jitteru</i> u mužů při intoxikaci.	64
5.19	Nejvýraznější průměrné změny <i>Shimmeru</i> u mužů při intoxikaci.	65
5.20	Nejvýraznější průměrné změny <i>Jitteru</i> u žen při intoxikaci.	66
5.21	Nejvýraznější průměrné změny <i>Shimmeru</i> u žen při intoxikaci.	67
5.22	Nejvýraznější změna průměrného počtu průchodů nulou u samohlásky O u žen v závislosti na intoxikaci.	68
5.23	Změny u parametru <i>Energie</i> při předstírané intoxikaci u samohlásky Á.	69
5.24	Změny u parametru <i>Energie</i> při předstírané intoxikaci u čtené řeči.	69
5.25	Změny u parametru <i>Energie</i> při předstírané intoxikaci u volné řeči.	70
5.26	Změny u parametru F_0 při předstírané intoxikaci u samohlásky Á.	70
5.27	Změny u parametru F_0 při předstírané opilé čtené řeči.	71
5.28	Změny u parametru F_0 při předstírané opilé volné řeči.	71
5.29	Změny u parametru F_1 při předstírané intoxikaci u samohlásky Ú.	71
5.30	Změny u parametru F_1 při předstírané intoxikaci u samohlásky Í.	72
5.31	Změny u parametru F_2 při předstírané intoxikaci u samohlásky Ú.	72
5.32	Změny u parametru C_1 při předstírané intoxikaci u čteného textu.	72
5.33	Změny u parametru C_1 při předstírané intoxikaci u volné řeči.	73
5.34	Změny u parametru C_2 při předstírané intoxikaci u čteného textu.	73
5.35	Změny u parametru C_2 při předstírané intoxikaci u volné řeči.	73
5.36	Změny u parametru <i>Počet nul</i> při předstírané intoxikaci u čtené řeči.	74
5.37	Změny u parametru <i>Počet nul</i> při předstírané intoxikaci u volné řeči.	74
5.38	Test na intoxikaci v aplikaci ALCTest.	75

SEZNAM TABULEK

1.1	Účinky koncentrace alkoholu v krvi [1]	8
2.1	Hodnota pásem prvních tří formantů pro české samohlásky [4]	14
3.1	Amplituda řečového signálu u Sobella [10]	28
3.2	Rychlost čtení u Sobella [10]	29
4.1	Dostupné informace o nahrávaných subjektech	41
5.1	Změna průměrných hodnot <i>Energie</i> v procentech mezi 0‰ a 1‰ u zkoumaných nahrávek	51
5.2	Změny průměrných hodnot F_0 v procentech mezi 0‰ a 1‰ u zkoumaných nahrávek	52
5.3	Změny průměrných hodnot formantů F_1 až F_3 mezi 0 a 1‰ u mužských nahrávek samohlásek	57
5.4	Změny průměrné hodnoty antiformantů M_1 až M_3 mezi 0 a 1‰ u nahrávek mužských samohlásek	57
5.5	Změny průměrných hodnot formantů F_1 až F_3 mezi 0 a 1‰ u nahrávek samohlásek u žen	58
5.6	Změny průměrných hodnot antiformantů M_1 až M_3 mezi 0 a 1‰ nahrávek samohlásek u žen	58
5.7	Změny průměrných hodnot kepstrálních koeficientů v procentech mezi 0 a 1‰ u mužských nahrávek	59
5.8	Změny maximálních průměrných hodnot kepstrálních koeficientů u mužských nahrávek	60
5.9	Změny průměrných hodnot kepstrálních koeficientů u nahrávek žen.	61
5.10	Změny maximálních průměrných hodnot kepstrálních koeficientů u nahrávek žen	62
5.11	Změny průměrných hodnot jitterů v procentech mezi 0‰ a 1‰ u mužských nahrávek	64
5.12	Změny průměrných hodnot shimmerů v procentech mezi 0‰ a 1‰ u mužských nahrávek	65
5.13	Změny průměrných hodnot jitterů v procentech mezi 0‰ a 1‰ u ženských nahrávek	66
5.14	Změny průměrných hodnot shimmerů v procentech mezi 0‰ a 1‰ u ženských nahrávek	67
5.15	Změny průměrných hodnot počtu průchodů nulou v procentech mezi 0‰ a 1‰	68

ÚVOD

Diplomová práce Detekce alkoholu v řečovém signálu se zabývá široce přijímanou hypotézou, že intoxikace alkoholem ovlivňuje způsob, jakým člověk mluví. Intoxikace alkoholem je individuální u každého jedince. Vliv alkoholu je závislý nejen na pohlaví, váze a množství vypitého alkoholu, ale také na aktuální kondici. V nižších dávkách působí povzbudivě (pocit euforie, uvolnění a ztráta zábran), ve větších dávkách může dojít k projevům agresivity, typická je nesouměrnost pohybů (vrávorání, pády). Dochází k rozpoznatelným změnám v řeči vnějším pozorováním. Řeč je nesrozumitelná, setřelá, dysartrická.

Fatální mohou být následky při řízení vozidel. V české republice je užití alkoholu při řízení vozidla postihováno od pokuty 2500kč a zákazu řízení od 6 měsíců až po nepodmíněný trest 1 roku vězení, ztráty 7 bodů, pokuty až 50000kč a zákazu řízení na 1-2 roky. Pokud tedy nedojde k vážné dopravní nehodě. Pravděpodobnost nehody do 1 promile se zvyšuje až 7x. Úmrtnost na silnicích zaviněná alkoholem se v EU v roce 2018 snížila v porovnání s rokem 2010 z 63 usmrcených na 1 milion obyvatel na hodnotu 49 na 1 milion obyvatel.

Alkohol za volantem je v jedenácti zemích EU zakázán.¹ Zjištění intoxikace alkoholem se provádí až po nehodě nebo namátkou. Klademe si tedy otázku, zda lze provádět detekci alkoholu preventivně ještě před nastartováním motoru například pomocí analýzy řeči řidiče.

Práce se zaměřuje na významné publikace, které byly zpracovány za účelem popisu změn v řečovém signálu po požití alkoholu. Využití se projevilo například v případě kapitána ropného tankeru Exxon Valdez, který dne 24. března 1989 ztroskotal na pobřeží Aljašky s plným nákladem ropy. Kapitán byl v době krize pod vlivem alkoholu. Forenzní analýza zaznamenaného hovoru ukázala, že spektrum hovoru bylo nakloněno ve směru, který byl považován za indikátor opilosti.

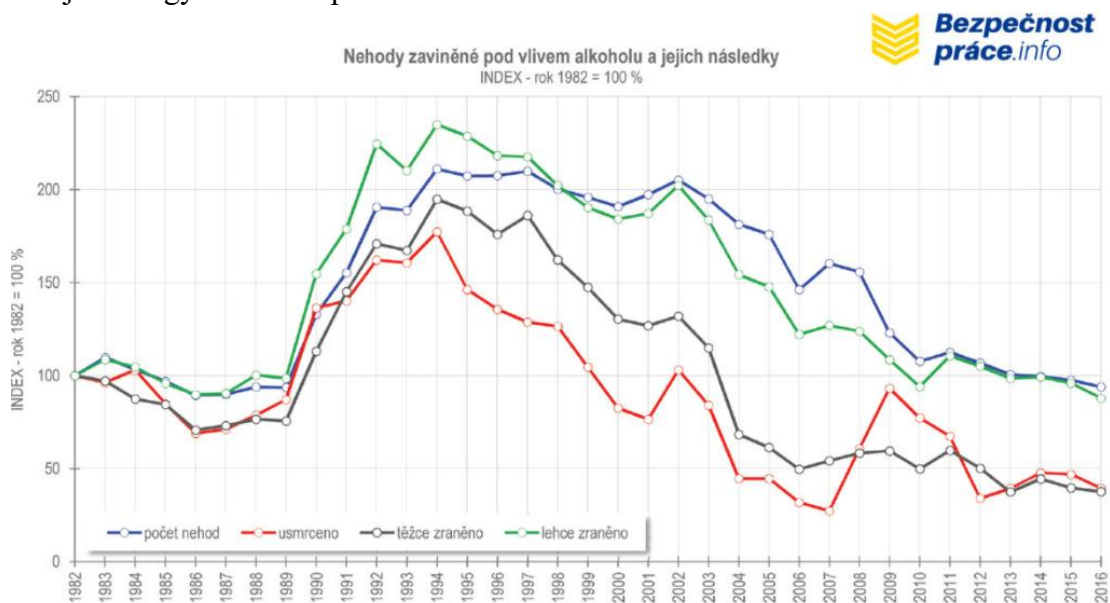
Práce se dále zabývá vytvořením vlastní databáze nahrávek střízlivé a opilé řeči. Databáze obsahuje nahrávky samohlásek, rozmlouvadel, čtené a volné řeči od 40 subjektů.

Pomocí vytvořené aplikace ALCTest v programu MATLAB získává 40 parametrů z každé nahrávky k dalšímu výzkumu. Dále zkoumá vliv alkoholu na jednotlivé parametry, u všech získaných nahrávek, během zvyšující se intoxikace od 0‰ do 1‰. Pomocí porovnání změn v parametrech se snaží práce najít možnost detekce alkoholu z řečového signálu. Práce dále zkoumá změny parametrů u předstírané opilé řeči.

¹ Povolená hranice alkoholu v krvi 0.5‰ platí v Belgii, Bosně, Finsku, Francii, Holandsku, Islandu, Itálii, Lotyšsku, Makedonii, Maltě, Monaku, Německu, Portugalsku, Rakousku, San Marinu,

1. ALKOHOL A JEHO VLIV [1]

Od roku 2007 do roku 2017 způsobili v České republice řidiči pod vlivem alkoholu nebo jiné drogy 58 398 dopravních nehod z toho zemřelo 756 osob.



Obrázek 1.1 Nehody zaviněné pod vlivem alkoholu a jejich následky [1]

Když hladina alkoholu v těle člověka stoupá, zvyšují se zároveň negativní účinky na centrální nervový systém. Alkohol je absorbován přímo skrze stěny žaludku a tenkého střeva. Poté přechází do krevního oběhu, kde se hromadí, dokud není metabolizován játry. Hladina alkoholu se měří třemi způsoby.

Miligram na litr mg/l – měří obsah alkoholu v dechu. 1 mg/l znamená, že 1 litr vydechnutého vzduchu obsahuje 1 mg alkoholu.

Promile ‰ – měří obsah alkoholu v krvi. 1 ‰ představuje tisícinu, takže v 1 litru krve se nachází 1 ml čistého alkoholu. V České republice je to oficiální zákonná jednotka určená pro dokazování alkoholu v krvi. Převod ‰ na mg/l se dělá v poměru 2:1, tzn. že 0,2 ‰ = 0,1 mg/l.

BAC ‰ – měří procentuální podíl alkoholu v krvi. Tato jednotka se používá v Americe, Kanadě nebo Austrálii. Často se s touto jednotkou můžete setkat také v alkohol testech. Převod ‰ na ‰ se dělá v poměru 1:10, tzn. že 0,1 ‰ BAC je 1 ‰. Převod ‰ na mg/l se dělá v poměru 1:5, tzn. že 0,1 ‰ BAC je 0,5 mg/l.

BAL ‰ [36] - je pouze orientační hodnota odečítaná z tabulky. Používá v Americe. Určuje počet miligramů alkoholu ve 100 ml krve.

1.1 Účinky koncentrace alkoholu v krvi

Tabulka 1.1 Účinky koncentrace alkoholu v krvi [1]

Koncentrace alkoholu v krvi	Efekt	Účinky na řízení
0,2 ‰	Ztráta úsudku, relaxace, změna nálady, zvýšená teplota těla.	Pokles vizuálních funkcí (rychlé sledování pohybujícího se cíle), snížená schopnost provádět dvě věci najednou.
0,5 ‰	Ztráta kontroly malých svalů (zaostřování očí), zhoršený úsudek, snížená ostrážitost, uvolnění inhibice – snížení aktivity.	Snížená koordinace, snížená schopnost sledovat pohyblivé předměty, potíže s řízením, snížená odezva na situace nouzového řízení.
0,8 ‰	Špatná koordinace svalů (rovnováha, řeč, vidění, doba reakce a sluch), je těžké odhalit nebezpečí; zhoršené rozhodování, sebeovládání, uvažování a paměť.	Zhoršená koncentrace, ztráta krátkodobé paměti, regulace rychlosti, snížená schopnost zpracování informací (detekce signálu, vizuální vyhledávání), narušené vnímání.
1 ‰	Jasně zhoršení reakční doby a kontroly, neuspokojivá řeč, špatná koordinace a zpomalené myšlení.	Snížená schopnost udržovat správnou pozici v jízdním pruhu, snížená schopnost brzdění.
1,5 ‰	Daleko méně svalové kontroly, může dojít ke zvracení významná ztráta rovnováhy.	Podstatně zhoršené schopnosti řízení vozidla, včetně pozornosti a zpracování vizuálních a sluchových informací.
3 ‰	Zde už mluvíme o těžké opilosti. Neschopnost samostatné chůze, velmi špatně srozumitelná řeč, poruchy chování, zvýšené riziko agrese, ztráta kontroly a paměti.	Nemožné řídit. Absolutní ztráta jakýchkoliv předpokladů řízení motorového vozidla.
4 ‰	Tato hodnota už představuje vážnou intoxikaci. Projevuje se mimo předchozí příznaky také pomalým dýcháním.	Nemožné řídit. Absolutní ztráta jakýchkoliv předpokladů řízení motorového vozidla.
5 ‰	V tomto případě už hrozí i smrt. Pro zhruba 50 % lidí je to smrtelná dávka. Velmi často dochází ke komaťóznímu stavu	Nemožné řídit. Absolutní ztráta jakýchkoliv předpokladů řízení motorového vozidla.

1.2 Doba účinku alkoholu

Alkohol se po požití dostává z trávicího systému přímo do krve. Tělo ho bere jako cizí látku, kterou se svým obranným způsobem snaží odbourat. Vylučování probíhá prostřednictvím dechu, potu a moči. Zbylý alkohol je zpracováván v trávicím systému.

Rychlost vylučování alkoholu z těla závisí na rychlosti přijímání alkoholu, výšce, váze a pohlaví. Vylučování alkoholu z krve neovlivňuje spánek, sport, káva ani tučná jídla.

Popis výpočtu zbytkového alkoholu v krvi [2]

Toto je stručný popis algoritmu, který se dá využít pro výpočet zbytkového alkoholu v krvi. Každý konzumovaný nápoj je charakterizován dvěma parametry — svým objemem a obsahem alkoholu. Na základě těchto parametrů můžeme vypočítat množství alkoholu obsažené v nápoji pomocí následujícího vzorce:

$$hmotnost = (objem * obsah alkoholu * 0,8) / 100 [g] , \quad (1.1)$$

Výsledná hmotnost v gramech je celkové množství čistého etanolu, který je v nápoji obsažen. Objem nápoje je v tomto vzorci zadán v mililitrech a hodnota 0,8 zastupuje hustotu etanolu. Aktuální hodnotu promile lze vypočítat pomocí vzorce:

$$promile = hmotnost / (váha * r) [g/kg] \quad (1.2)$$

Kde hmotnost je celkové množství etanolu v gramech a váha je váha konzumenta v kilogramech. Konstanta r určuje podíl vody v organismu. Pro muže nabývá hodnoty 0,7 a pro ženy pak 0,6. Lidský organismus má schopnost alkohol z těla odbourávat. Pokles hladiny alkoholu v krvi za jednu hodinu se označuje jako faktor β . Pomocí následujícího vzorce lze vypočítat hmotnost alkoholu, který tělo odbourá za jednu hodinu:

$$hmotnost = váha * \beta [g] \quad (1.3)$$

Hmotnost je zde celkové množství etanolu v gramech, které se z těla odbourá za jednu hodinu. Váha konzumenta je v kilogramech a konstanta β je podle odborných pramenů pro muže 0,1 a pro ženy 0,085 a udává, kolik gramů alkoholu se odbourá za jednu hodinu na 1 kilogram živé váhy.

1.3 Stanovení alkoholu v organismu [3]

Při stanovení hladiny alkoholu v krvi je třeba rozlišovat, pro jaké účely se toto stanovení provádí. Zcela jiné nároky na zjištění alkoholu v organismu jsou v různých profesích. V mnoha případech je rozhodující jednoduchost a rychlost zjištění alkoholu v organismu, byť je toto provedeno na úkor přesnosti - kontrola na silnici, pracovištích a pro diagnostické účely v medicíně.

Opačné požadavky, kde vystupuje především přesnost a jednoznačnost stanovení hladiny, jsou v důkazních řízeních, mající za následek postih vyšetřované osoby (pokuta, správní řízení, rozvázání pracovního poměru, náhrady škod, plnění pojistek a soudní řízení všeho druhu).

Metody zjišťování alkoholu v organismu

V první řadě zpravidla nastupuje metoda vyhledávací, tj. detekční trubička či dechové analyzátoary. K orientačnímu zjištění ovlivnění alkoholem se používá zkouška dechu detekční trubičkou DETELCOL nyní ALTEST. Na těle trubičky je vyznačena dělicí ryska, která určuje hranici 0,8 g.kg-1 alkoholu v krvi u kontrolované osoby.

V praxi, při posuzování vlivu alkoholu na schopnost bezpečně řídit motorové vozidlo, zaujímá klinické vyšetření druhé místo za laboratorně zjištěnou hladinou alkoholu v krvi.

Odebrání vzorku krve se provádí co nejdříve po nehodě. Jestliže však došlo k nehodě ihned po požití alkoholu, pak je třeba odebrat vzorek krve po 90 minutách po požití alkoholu, protože resorpční doba je od 30 do 90 minut.

Přítomnost alkoholu v organismu lze prokázat i vyšetřením jiných tělních tekutin, zejména moči. Přesnost výsledku vyšetření je stejná jako u krve, ovšem následná interpretace takového výsledku se zaměřením na stupeň ovlivnění vyšetřované osoby je značně komplikovaný.

1.4 Užitá metoda zjištění alkoholu v organismu

Pro stanovení hladiny bude použito dechové zkoušky, pro kontrolu může být použito alkoholové kalkulačky implementované ve vyvíjené aplikaci ALCTest a online kalkulačky ze stránek <https://www.idnes.cz/auto/alkulacka> pro výpočet alkoholu v krvi. Dechová zkouška je provedena za pomoci Alkohol testeru *Dräger Alcotest 7410*, který používá i POLICIE České republiky a Městské policie. Přesnost měření je $\pm 0,05\%$ (při pozitivním měření 0,00 až 1,00‰) a $\pm 5\%$ z měřené hodnoty nad 1,00‰. Rozsah měření je od 0.00‰ – 3,00‰.

Doba ustálení alkoholu v krvi se odhaduje okolo 20 minut po intoxikaci. Po této době je vhodné provádět test hladiny alkoholu a nahrávky pro studii.

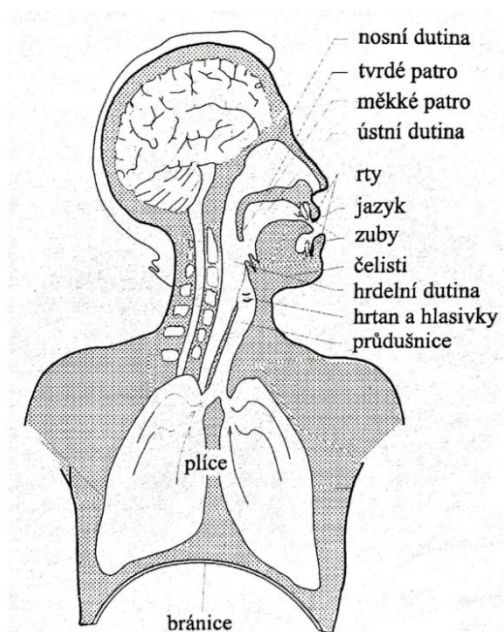
Odběry krve v tomto případě provedeny nebudou, neboť jsou značně nákladné na školený personál. Odběry by musely být prováděny pod odborným dohledem a ve speciálním zdravotnickém zařízení.

Profesionálně užívané metody na zjištění intoxikace jedince jsou značně nákladné a mnohdy se provádí až po nehodě. Detekce alkoholu z řečového signálu by mohla být cesta k zlevnění prvotní detekce bez užití přístroje na detekci alkoholu z dechu, který se musí pravidelně kalibrovat a udržovat.

2. ZPRACOVÁNÍ ŘEČOVÝCH SIGNÁLŮ

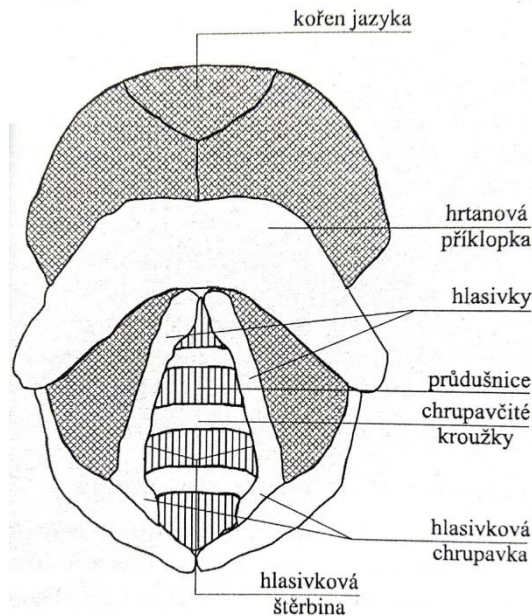
2.1 Vytváření řeči člověkem [4]

Pro vytváření řeči existuje v lidském těle několik skupin orgánů, které se nazývají řečové orgány. Hlasový trakt lze rozdělit na tři základní ústrojí: dechové, hlasové a artikulační. Trvání dechu má vliv na to, jak dlouhý úsek řeči lze vytvořit bez přerušení. Síla výdechového proudu má vliv na sílu hlasu a částečně i na jeho výšku. Nádech se projevuje jako pauza řeči.



Obrázek 2.1 Hlasový trakt člověka [4]

Hlasové ústrojí je uloženo v hrtanu. Z hlediska tvorby řeči nejdůležitější část hlasového ústrojí tvoří hlasivky. Frekvence kmitání hlasivek se označuje F_0 a nazývá se fundamentální frekvence nebo frekvence základního hlasivkového tónu. V této práci používáme označení základní hlasový tón.



Obrázek 2.2 Hrtan a hlasivky [4]

Převrácenou hodnotou $T_0 = 1/F_0$ nazýváme periodou základního hlasového tónu. Udává se, že základní hlasový tón má rozsah asi 60-400Hz. Při normální řeči se F_0 pohybuje zhruba v rozmezí jedné oktávy. U mužů se pohybuje asi mezi 80-160Hz s průměrnou hodnotou 132Hz, a u žen pak mezi 150-300Hz s průměrnou hodnotou 223Hz a u dětí dokonce v rozmezí 200-600Hz. Pro hluboký mužský hlas může F_0 klesnout i na 50Hz a naopak pro ženský hlas může vystoupat až nad 400Hz. Pro cvičený hlas není výjimkou základní frekvence převyšující i 1000Hz. Mírné kolísání délky periody se nazývá *jitter* (definován jako střední rozdíl délek sousedních period, dělený střední délkou periody a obvykle se vyjadřuje v procentech) a je závislý na duševním stavu mluvčího. Kolísání amplitudy hlasivkových pulsů se označuje jako *shimmer* (definován středním rozdílem mezi amplitudami sousedních hlasivkových pulsů a vyjadřuje se v dB).

Při normální řeči se *jitter* pohybuje přibližně mezi 0,5%-1,0% a *shimmer* mezi 0,04-0,21 dB. Změny nejsou postřehnutelné, neboť změny jsou malé.

Artikulační ústrojí je poslední ústrojí, které se podílí na tvorbě řeči. Skládá se z nadhrtanových dutin (hrdelní, ústní a nosní) a z artikulačních orgánů (jazyk, rty, měkké patro, zuby, tvrdé patro, čelisti a hrtan).

Akustická energie se soustředí kolem určitých frekvencí, kterým se říká formantové frekvence. Oblasti koncentrace akustické energie se nazývají formanty a označují se čísly počínaje formantem s nejnižší frekvencí, F_1, F_2, \dots, F_n . Pokud se do vytváření řeči zapojí i nosní dutina, dochází navíc vlivem jejich antirezonančních vlastností k potlačení některých frekvenčních oblastí, tzv. antiformanty (značí se M_1, M_2, \dots, M_m).

2.2 Fonetický popis českého jazyka [4]

Pro potřeby zkoumání řeči byly zavedeny fonologické jednotky v hierarchickém členění:

promluva → fráze → slova → slabiky → difony → fonémy → alofony

Promluva - celek řeči mezi dvěma absolutními pauzami.

Fráze - vyznačena větou intonací a relativními pauzami.

Slovo - významová jednotka. Evropské jazyky používají přibližně 45000 - 50000 slov.

Slabika - nejmenší jednotka, na kterou jde rozčlenit souvislou promluvu bez narušení její sdělnosti. Evropské jazyky používají přibližně 2500 - 3500 slabik.

Foném - nejmenší významotvorná jednotka. Evropské jazyky mají okolo 40 fonémů.

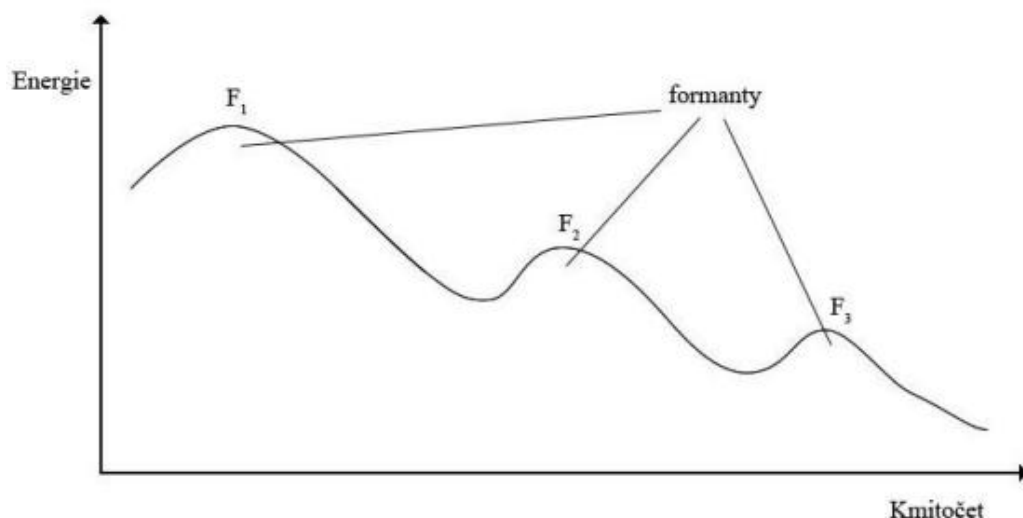
Difon - přechodné zvuky kterými lze popisovat vliv dané hlásky na hlásku předchozí.

Alofon - nejmenší zvukově rozlišitelné jednotky řeči. Slouží k rozpoznávání různých variant jednoho fonému v odlišných promluvách.

Základní jednotku řeči tvoří foném neboli hláska. Jednotlivé hlásky pak složí slovo. Fonémy dělíme na dvě základní skupiny: samohlásky (vokály) a souhlásky (konsonanty).

Samohlásky (vokály)

Ve spisovné češtině nalezneme 10 samohláskových fonémů, vznikají v ustálené poloze řečového traktu a jsou znělé. Jsou charakteristické téměř periodickým průběhem a relativně vysokou energií signálu. V průběhu kmitočtového spektra samohlásek jsou zřetelné jednotlivé rezonanční kmitočty, tzv. formanty. Ke každé samohlásce můžeme určit i více než pět formantů F_n . Pro určení samohlásky postačí pouze první dva. První tři formanty [7] mohou být považovány za rezonanční kmitočty největších dutin hlasového traktu s přiřazením: F_1 - dutina hrdelní, F_2 - dutina ústní a F_3 - dutina nosní.



Obrázek 2.3 Formanty ve spektru samohlásek [5]

V tabulce 2.1 jsou uvedeny hodnoty kmitočtů prvních tří formantů pro české samohlásky. Formantová struktura samohlásek je odrazem artikulace. První dva

formanty vykazují největší rozdíly v umístění pro různé samohlásky. Rozdíly v hodnotách dalších formantů jsou již menší.

Tabulka 2.1 Hodnota pásem prvních tří formantů pro české samohlásky [4]

Formanty Samohlásky	F_1 [Hz] pásmo	F_2 [Hz] pásmo	F_3 [Hz] pásmo
i,í	300-500	2000-2800	2600-3500
e,é	480-700	1560-2100	2500-3000
a,á	700-1100	1100-1500	2500-3000
o,ó	500-700	850-1200	2500-3000
u,ú	300-500	600-1000	2400-2900

Souhlásky (konsonanty)

Oproti samohláskám je jejich výrazným rysem šum ve spektru, čímž se výrazně ztíží jejich identifikace. Vznikají vlivem překážky v turbulentním proudění vzduchu. Souhlásky reprezentují přechodové stavy, mají kratší dobu trvání a podle způsobu artikulace je můžeme rozdělit do na explozivny (p,b,t,t',d,d',k,g,m,n,ň),frikativy (f,v,s,š,z,ž),afrikáty (c,č) a vibranty (r,ř).

2.3 Zpracování řečového signálu [6]

Základem většiny metod analýzy akustického signálu řeči je předpoklad, že se jeho vlastnosti v průběhu času mění pomalu. Tento předpoklad vede na aplikaci krátkodobé analýzy, při nichž se úseky řečových signálů vydělují a zpracovávají tak, jako by to byly krátké oddělené zvuky. Tyto mikrosegmenty mají typickou délku 10ms.

Princip kódování tvaru vlny

Pulzní kódová modulace (PCM) – řeč se zpravidla zpracovává získáváním zvukových kmitů pomocí mikrofону. Dále se signál digitalizuje pomocí dvou kroků, a to vzorkováním a kvantizací s kódováním. Na frekvenci vzorkování $F_V=1/T$ jsou kladena jistá omezení vyplývající z Nyquistova vzorkovacího teorému. Vzorkovací kmitočet by měl být minimálně dvojnásobný. Každý vzorek je kvantován převodníkem A/D a jako výstup generuje odpovídající kódovou reprezentaci.

Zpracování ve frekvenční oblasti

V časové oblasti a i ve frekvenční oblasti se pracuje s představou, že signál je přibližně stacionární, a proto je účelné mluvit o krátkodobé spektrální analýze. Nejčastěji používané principy jsou založeny na aplikaci krátkodobé Fourierovy transformace.

V tomto případě je čas i frekvence diskrétní a získané koeficienty se využívají dále zejména ve spektrálních analyzátozech řeči, v řečových syntetizérech, a systémech rozpoznávání řeči, nebo jsou prvním stupněm zpracování různých parametrizačních technik. Nehodí se jako vstupní příznaky pro řečové klasifikátory vzhledem k nepodstatnému snížení informační redundance ve srovnání s PCM. Fourierova

transformace se dá použít jako proces lineární filtrace, je možné volbou vhodného okénka navrhnout pásmový filtr či soustavu pásmových filtrů a zkoumat spektrální vlastnosti řečového signálu na výstupu úzkých frekvenčních pásem.

Formáty pro záznam řečových signálů [8]

Existuje celá řada formátů pro záznam hlasu či hudby. Nejpoužívanějším je formát MP3 využívající ztrátovou kompresi. V profesionální sféře dominují formáty, které nejsou postiženy nevýhodami ztrátové komprese. Rozšířené formáty jsou např. WAV nebo FLAC. Pro jednokanálový záznam hlasu dostačuje vzorkovací frekvence 22kHz.

Formát WAV [9]

Formát WAV je standardní formát pro záznam zvuku vyvinutý společností IBM a Microsoft pro ukládání audia do PC. Je zde aplikován formát RIFF (Resource Interchange File Format) pro ukládání dat v „blocích“. Je to hlavní formát používaný v systémech Microsoft Windows pro nezpracovaný a obvykle nekomprimovaný záznam zvuku. Ke kódování formátu WAV se využívá lineárně pulzní kódová modulace (LPCM). Zvuk se ukládá do dvou kanálů se vzorkovací frekvencí 44 100Hz a 16bity na vzorek. Ke komprimaci zvuku se používá jakýkoliv kodek ACM (Audio Compression Manager). Pro cíle práce bude využit právě tento způsob záznamu právě pro jeho nekomprimovanou formu.

2.4 Parametry řečového signálu

Digitalizovanou řeč můžeme brát za „lokálně stacionární“ stochastický proces. V praxi se ukazuje, že tento proces probíhá dostatečně pomalu na odhadnutí jeho lokální vlastnosti z hodnot sledovaného signálu.

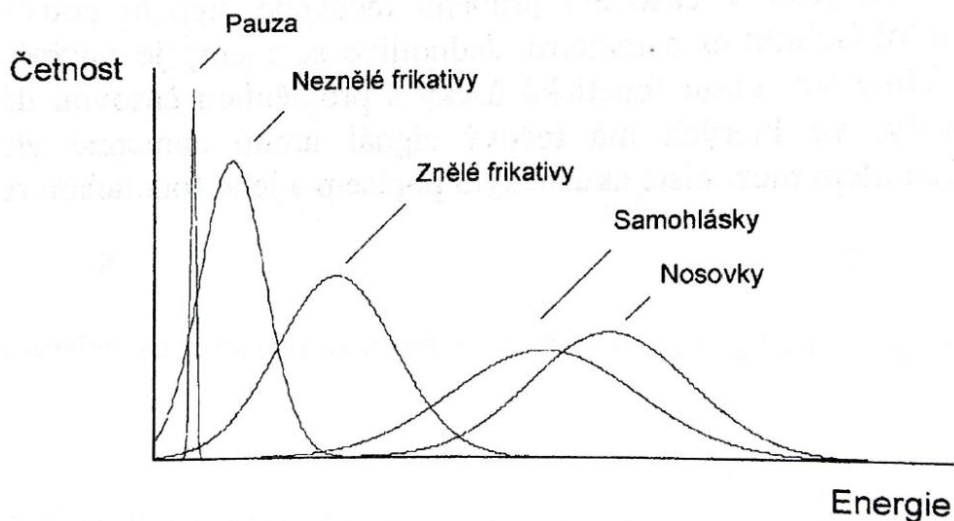
Navzorkovaný řečový signál obsahuje mnoho irelevantních informací. Je nutné potlačit celkovou redundanci řeči a zdůraznit relevantní vlastnosti řečového signálu pro cílové rozpoznávání. Příznaky získáme pomocí vhodných metod analýzy.

Energie signálu [7]

Energie diskrétního řečového signálu $s(n)$ na jednom segmentu o délce N vzorků je dána vztahem

$$E = \sum_{n=1}^N s^2(n) \quad (2.1)$$

Energie signálu E je u znělých hlásek zřetelně vyšší než u hlásek neznělých. Může být použita u rozlišení mezi znělými a neznělými hláskami a k rozpoznání přechodu mezi znělým a neznělým úsekem. Energie signálu se nejčastěji používá jako kritérium pro přítomnost řeči a klidu.



Obrázek 2.4 Rozdělení fonémů do skupin podle energie řečového signálu [7]

Energie signálu může být použita k hrubému dělení fonémů do pěti skupin: samohlásky, nosovky, znělé frikativy, neznělé frikativy a pauz v řeči. Z obrázku 2.4 je patrné, že skupiny samohlásek a nosovek lze dobře oddělit od skupiny neznělých frikativů a pauz.

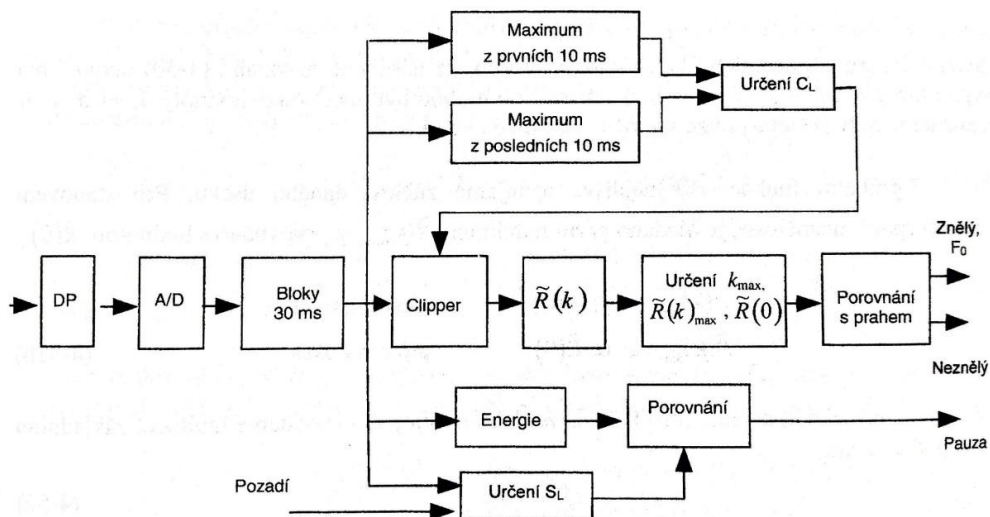
Kmitočet základního tónu řeči

Základní tón řeči můžeme považovat za hlavní parametr, který určuje charakteristiku mluvího. Jeho střední hodnota je vnímána jako výška hlasu.

Hodnoty F_0 leží obvykle v intervalu 50-400 Hz. U mužů je udáván základní kmitočet 120 Hz, u žen téměř dvojnásobný 210 Hz.

K určení základního tónu se využívá metoda AMDF (Average Magnitude Difference Function). Princip spočívá v nahrazení operace násobení při výpočtu autokorelační funkce operací odečítání. První výrazné minimum pak odpovídá svou polohou periodě základního tónu.

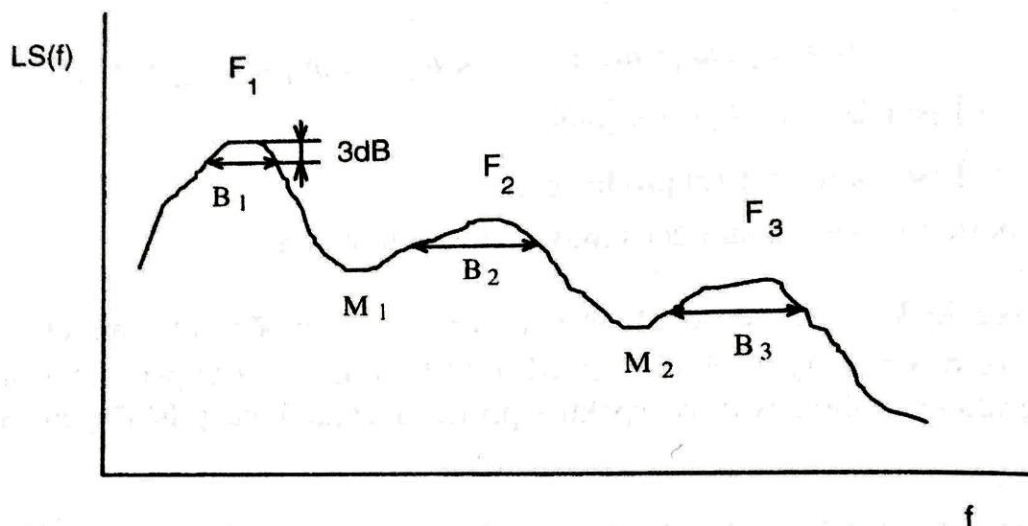
Metoda AMDF je sice jednoduchá, avšak v malé míře potlačuje vliv formantů. Odstranění této nevýhody lze dosáhnout vyříznutím střední části signálu, kterou využívá metoda Center-Clipping obrázek 2.5.



Obrázek 2.5 Blokové schéma metody Center-Clipping [7]

Formantové příznaky

Při zpracování řeči se často používají příznaky z kmitočtové oblasti odvozené od formantů. Je známo, že první tři formantové kmitočty obsahují důležité informace o charakteru samohlásek a znělých souhlásek. Kmitočty formantů F_i , šířky formantů B_i a kmitočty antiformantů M_i jsou ve vzájemné souvislosti znázorněny na obrázku 2.6.



Obrázek 2.6 Jednoduché parametry řeči založené na vlastnostech spektra [7]

Jako samostatné příznaky můžeme použít následující parametry krátkodobého spektra:

Kmitočet prvního formantu:

$$x_i = F_1 \quad (2.2)$$

Tímto parametrem lze od sebe odlišit např. foném „i“ a foném „a“.

Střední kmitočet tří formantů:

$$x_i = (F_1 + F_2 + F_3)/3 \quad (2.3)$$

Hodnoty druhého a třetího formantu nelze v některých případech snadno určit, proto se

někdy používá střední hodnota. Pomocí tohoto parametru lze rozlišit spektrálně blízké fonémy „i“ a „e“.

Střední šířka pásma tří formantů:

$$x_i = \frac{B_1 + B_2 + B_3}{3} \quad (2.4)$$

Znělé fonémy mají obvykle užší pásma formantů než neznělé nebo pauzy. Tento příznak může sloužit k rozlišení znělosti.

Kmitočet prvního antiformantu (vztaženo k F_1):

$$x_i = M_1 - F_1 \quad (2.5)$$

Poloha kmitočtů formantů lze určit podle maxima v logaritmickém spektru. Podobně také minima ve spektru je tedy možné považovat za parametry. Jako příznak se nejčastěji používá první minimum po formantu F_1 .

Velikost prvního formantu (vztaženo k 0 dB):

$$x_i = LS(F_1) - LS(0) \quad (2.6)$$

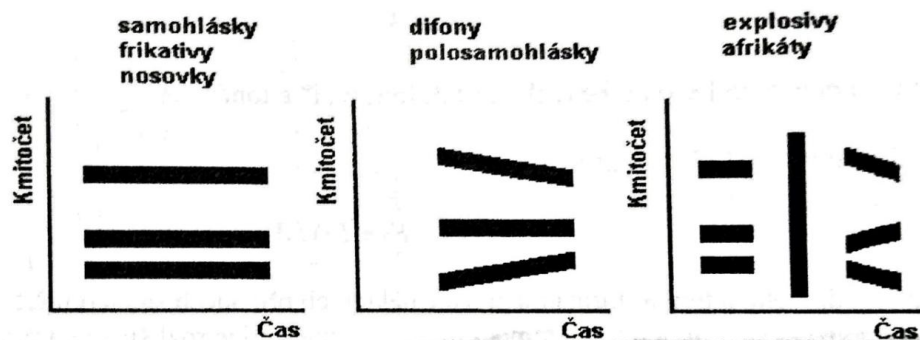
U fonému „a“ nabývá vysokých hodnot.

Velikost prvního formantu (vztaženo k minimu):

$$x_i = LS(F_1) - LS(M_1) \quad (2.7)$$

U fonému „a“ je tato hodnota velmi malá, u nosovek naopak velmi velká.

Podle průběhů formantových kmitočtů v čase lze určit artikulaci sousedních hlásek. Můžeme pozorovat tři typy průběhů: stacionární, spojitý a přerušovaný.



Obrázek 2.7 Stacionární, spojitý a přerušované spektrum různých fonémových skupin [7]

Kromě metod založených na prohledávání obálky spektra a vyhledávání lokálních maxim, můžeme určit formanty také přímo výpočtem při LPC analýze. Póly z_i v rovnici 2.8 jsou vázány s formantovými kmitočty F_i a odpovídajícími šířkami pásem B_i vztahem

$$z_i = \exp\left(-\frac{\pi B_i}{f_{vz}} + \frac{j2\pi F_i}{f_{vz}}\right) \quad (2.8)$$

z něhož plyne

$$F_i = \frac{f_{vz}}{2\pi} \arg(z_i) \quad (2.9)$$

$$B_i = -\frac{f_{vz}}{\pi} \ln |z_i| \quad (2.10)$$

Jitter a Shimmer [37]

Mírné kolísání délky periody se nazývá *jitter* (definován jako střední rozdíl délek sousedních period, dělený střední délkou periody a obvykle se vyjadřuje v procentech) a je závislý na duševním stavu mluvčího. Kolísání amplitudy hlasivkových pulsů se označuje jako *shimmer* (definován středním rozdílem mezi amplitudami sousedních hlasivkových pulsů a vyjadřuje se v dB).

Jitter (local)

je lokální absolutní rozdíl mezi po sobě jdoucími intervaly

$$jitter(s) = \sum_{i=2}^N |T_i - T_{i-1}| / (N - 1) \quad (2.11)$$

,kde T_i je doba trvání i -tého intervalu a N je počet intervalů.

Jitter (relativní)

Definujeme průměrnou dobu T

$$průměrná\ doba(s) = \sum_{i=1}^N T_i / N \quad (2.12)$$

Relativní jitter se pak spočítá podílem jitrů lokálního a průměrné doby.

$$jittter = \frac{jitter(s)}{průměrná\ doba(s)} \quad (2.13)$$

Výsledkem je hodnota mezi 0 a 2 nebo mezi 0 a 200 procenty.

Jitter RAP (Relative Average Perturbation)

Nejříve se definuje absolutní průměrná chyba AP.

$$absAP(s) = \sum_{i=2}^{N-1} |T_i - (T_{i-1} + T_i + T_{i+1}) / 3| / (N - 2) \quad (2.14)$$

Hodnota RAP se pak rovná podílu $absAP(s)$ a $průměrné\ doby(s)$.

$$RAP = \frac{absAP(s)}{průměrná\ doba(s)} \quad (2.15)$$

Výsledkem je hodnota mezi 0 a 2 nebo mezi 0 a 200 procenty.

JitterPPQ5 (five-point Period Perturbation Quotient)

JitterPPQ5 získáme, definujeme-li absolutní PPQ

$$absPPQ5(s) = \sum_{i=3}^{N-2} |T_i - (T_{i-2} + T_{i-1} + T_i + T_{i+1} + T_{i+2}) / 5| / (N - 4) \quad (2.16)$$

a podělíme průměrnou dobou.

$$PPQ = \frac{absPPQ5(s)}{průměrná\ doba(s)} \quad (2.17)$$

Shimmer (local)

je průměrný absolutní rozdíl amplitud jdoucích po sobě dělených průměrnou amplitudou.

Shimmer (local, dB)

je průměrný absolutní rozdíl logaritmu amplitud o základu 10 jdoucích po sobě násobených 20. Hodnota 0.350 dB je hranice pro patologie.

Shimmer (apq3)

Jedná se o tříbodový koeficient rozptylu amplitudy. Spočítá se jako průměrný absolutní rozdíl mezi amplitudou a průměrem sousedních amplitud, dělený průměrnou amplitudou.

Shimmer (apq5)

Jedná se o pětibodový kvocient amplitudy. Spočítá se jako průměrný absolutní rozdíl mezi amplitudou a průměrem čtyř nejbližších sousedů, dělený průměrnou amplitudou.

Lineární predikce (LPC)

Nejčastější a nejvýkonnější metoda analýzy řečového signálu je lineární predikce, známá pod zkratkou LPC (linear predictive coding). Metodu nelze jednoznačně zařadit do časové nebo kmitočtové oblasti.

Princip metody LPC spočívá v předpovědi n -tého vzorku řečového signálu pomocí lineární kombinace určitého počtu předcházejících váhových vzorků téhož signálu podle vztahu

$$\hat{s}(n) = \sum_{m=1}^M a_m \cdot s(n - m), \quad (2.18)$$

kde a_m jsou predikční (LPC) koeficienty a M značí řád prediktoru (počet koeficientů). Pokud jsou hodnoty a_m dobře nastaveny, je hodnota předpovězeného signálu lineární kombinací předchozích hodnot. Hlavním cílem predikční analýzy je výpočet koeficientů a_m . Pokud máme určeny koeficienty LPC, jsou jimi jednoznačně stanoveny charakteristiky periodického signálu. Krátké úseky řečového signálu lze tedy výstižně popsat omezeným počtem 6-18 koeficientů.

Charakter řečového signálu se neustále mění, proto může jedna sada koeficientů a_m platit ve vyhovující míře jen pro jeden segment. Výpočet koeficientů a_m cestou minimalizace střední chyby můžeme použít metodu autokorelační, která dává stabilní řešení.

Obvyklý postup znamená určit celkovou chybu přes celý segment ε jako součet kvadrátů chyb jednotlivých vzorků a poté optimalizovat koeficienty tak, aby celková chyba byla minimální.

$$\varepsilon = \sum_{n=1}^N e^2(n), \quad (2.19)$$

kde chyba odhadu n -tého vzorku $e(n)$ je dána rozdílem následujícího vzorku od skutečné hodnoty.

$$\varepsilon = \sum_{n=1}^N e^2(n) = \sum_{n=1}^N \left[s(n) - \sum_{m=1}^M a_m \cdot s(n-m) \right]^2 \quad (2.20)$$

Parciální derivace $\partial\varepsilon/\partial a_m = 0$ pro každý koeficient a_m vedou na lineární systém M rovnic pro určení M neznámých koeficientů a_m :

$$\begin{aligned} \frac{\partial}{\partial a_\mu} \left\{ \sum_{n=1}^N \left[s(n) - \sum_{m=1}^M a_m \cdot s(n-m) \right]^2 \right\} &= 0 & \mu = 1, \dots, M \\ \sum_{n=1}^N \left[s(n) - \sum_{m=1}^M a_m \cdot s(n-m) \right] \cdot [-s(n-\mu)] &= 0 \\ \left[\sum_n s(n) - \sum_n \sum_m a_m \cdot s(n-m) \right] \cdot [-s(n-\mu)] &= 0 \\ \sum_m a_m \sum_n s(n-m)s(n-\mu) &= \sum_n s(n)s(n-\mu) \\ \sum_m a_m \sum_n s(n-m)s(n-m+m-\mu) &= R(\mu) \\ \sum_m a_m R(|m-\mu|) &= R(\mu) \end{aligned} \quad (2.21)$$

Výsledný vztah můžeme použít již k výpočtu nebo jej přepsat do maticového tvaru a ze soustavy rovnic vypočítat predikční koeficienty a_m .

$$\begin{pmatrix} R(0) & R(1) & R(2) & \dots & R(M-1) \\ R(1) & R(0) & R(1) & \dots & R(M-2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ R(M-1) & R(M-2) & R(M-3) & \dots & R(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix} = \begin{pmatrix} R(1) \\ R(2) \\ \vdots \\ R(M) \end{pmatrix} \quad (2.22)$$

Matice autokorelačních koeficientů je tzv. Toeplitzova typu (symetrická a pozitivně semidefinitivní). Vyznačuje se tím, že všechny řádky matice lze získat cyklickým posunutím řádku prvního.

Na základě uvedených vlastností existuje pro soustavu lineárních rovnic velmi efektivní způsob řešení ve tvaru tří rekurzivních rovnic:

$$\text{Počáteční hodnoty} \quad e(0) = R(0)$$

$$a_{0,m} = 1$$

$$a_{m,n} = \left[R(m) - \sum_{i=1}^{m-1} a_{i,m-1} R(m-i) \right] / e(m-1) \quad (2.23a)$$

$$a_{i,m} = a_{i,m-1} - a_{m,n} a_{m-i,m-1} \quad m = 1, \dots, M \quad (2.23b)$$

$$e(m) = (1 - a_{m,n}^2) e(m-1) \quad i = 1, \dots, m-1 \quad (2.23c)$$

První část rovnice pozbývá smysl pro $m=1$ a je třeba ji vypustit. Dále pak určíme prediktor s jedním koeficientem $a_{1,1}=a_1$, potom následuje prediktor 2.řádu s dvěma koeficienty $a_{1,2}, a_{2,2}$ a dále řád prediktoru postupně zvyšujeme až do konečného řádu M s koeficienty $a_{1,M}, a_{2,M}, \dots, a_{M,M}$. Rozšířením sady koeficientů v každém kroku o další koeficient dochází vždy k přepočtu a změně hodnot původních koeficientů. Na závěr výpočtu přejdeme na obvyklé jednosložkové indexování $a_{1,M}, a_{2,M}, \dots, a_{M,M} \rightarrow a_1, a_2, \dots, a_M$.

LPC spektrum [7]

Predikční koeficienty mohou sloužit nejen pro předpověď nového vzorku signálu, ale obsahují také spektrální vlastnosti řečového signálu.

Pomocí z -transformace můžeme získat normované LPC spektrum $S(f)$ s jednotkovým zesílením podle vztahu

$$S(f) = \left| \frac{1}{1 - \sum_m a_m \cdot z^{-m}} \right|_{z=\exp\left(\frac{j2\pi f}{f_{vz}}\right)}^2 \quad (2.23)$$

$$= \frac{1}{\left| 1 - \left(a_1 e^{-\frac{j2\pi f}{f_{vz}}} + a_1 e^{-\frac{j4\pi f}{f_{vz}}} + \dots + a_1 e^{-\frac{jM\pi f}{f_{vz}}} \right) \right|^2}$$

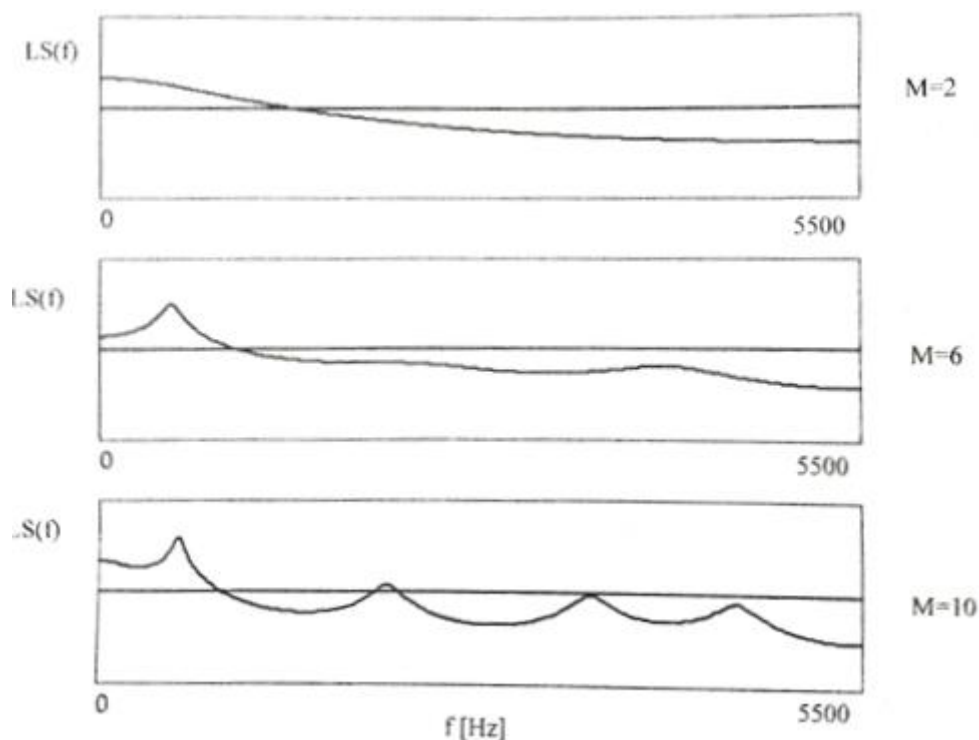
kde f_{vz} je kmitočet vzorkování řečového signálu $s(t)$ a f je proměnný kmitočet s maximální hodnotou $0,5f_{vz}$.

Standardně se při zpracování řečových signálů používá převodu do logaritmické soustavy. Spektrální hodnoty v decibelech označujeme jako logaritmické spektrum $LS(f)$.

Se zvyšujícím se řádem prediktoru M se LPC spektrum přibližuje FFT spektru. LPC spektrum má tvar vyhlazeného průběhu FFT spektra. Vhodný počet potřebných predikčních koeficientů může být určen předpisem:

$$M = f_{vz} + 4, \quad (2.25)$$

kde vzorkovací kmitočet je udán v kHz.



Obrázek 2.8 Vliv řádu prediktoru M na LPC spektrum [7]

Dlouhodobé LPC spektrum [7]

Dlouhodobá spektra odrážejí přenosové vlastnosti hlasového traktu daného mluvčího. Dlouhodobé spektrum můžeme použít na rozpoznávání mluvčích, a jak se ukázalo, dokonce také na rozpoznávání některých psychických stavů u mluvčích.

Při výpočtu nejdříve určíme koeficienty autokorelační funkce (AKF) $R_j(k)$ pro všechny segmenty podle vztahu

$$R(k) = \sum_{n=1}^{N-k} s(n) \cdot s(n+k) \quad \text{pro } k = 0, 1, \dots, N-1 \quad (2.26)$$

$$R(k) = 0 \quad \text{pro } |k| \geq N \quad (2.27)$$

pro j -tý segment a z nich průměrné AKF koeficienty pro celou promluvu o délce J segmentů.

$$\bar{R}(k) = \frac{1}{J} \sum_{j=1}^J R_j(k) \quad (2.28)$$

Z průměrných AKF koeficientů získáme průměrné LPC koeficienty \bar{a}_m a průměrné dlouhodobé LPC spektrum použitím stejných vztahů jako při krátkodobém spektru. Kvalita LPC spektra jako obrazu hlasového traktu je závislá na délce promluvy a na její fonetické struktuře.

Vlastnosti dlouhodobého LPC spektra nám udávají možnost k využití pro normalizaci promluvy při rozpoznávání řeči nezávisle na mluvčích. Pro tento účel vystačíme někdy již s predikcí druhého řádu a příslušné LPC koeficienty určíme z jednoduchých vztahů

$$\bar{a}_1 = \frac{\bar{R}(1)\bar{R}(0) - \bar{R}(1)\bar{R}(2)}{\bar{R}(0)^2 - \bar{R}(1)^2} \quad (2.29)$$

$$\bar{a}_2 = \frac{\bar{R}(2)\bar{R}(0) - \bar{R}(1)\bar{R}(1)}{\bar{R}(0)^2 - \bar{R}(1)^2} \quad (2.30)$$

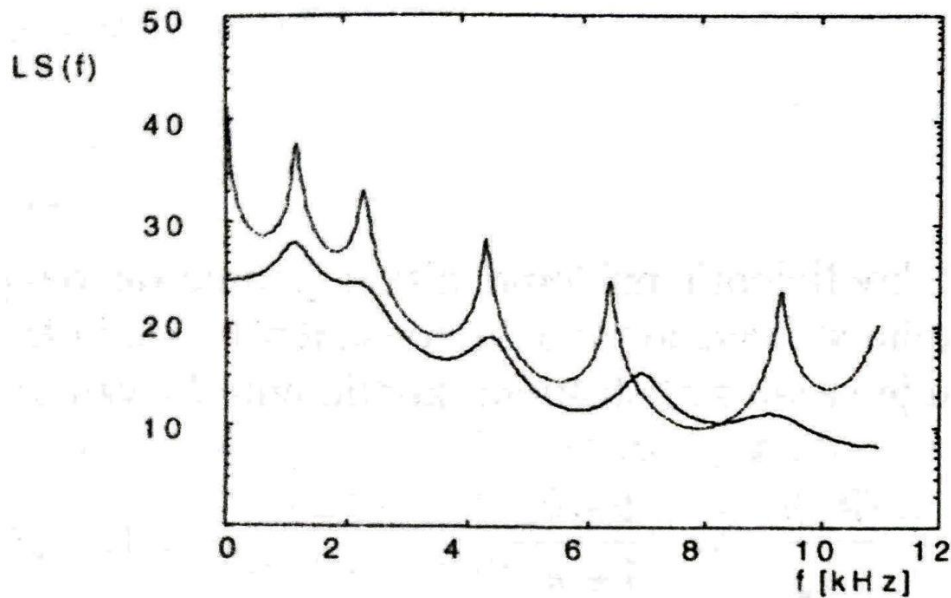
Normalizaci provedeme formou transformace AKF koeficientů $R_j(k)$ v každém j -tém segmentu řeči

$$R_j^n(k) = R_a(0)R_j(0) + \sum_{m=1}^M R_a(m)[R_j(|k-m|) + R_j(|k+m|)], \quad (2.31)$$

kde

$$R_a(m) = \sum_{i=0}^{M-m} \bar{a}_i \bar{a}_{i+m} \quad (2.32)$$

Normalizované AKF koeficienty $R_j^n(k)$ pak použijeme na výpočet dalších příznaků, popřípadě k rozpoznávání.

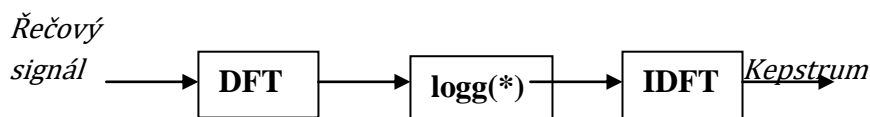


Obrázek 2.9 Vliv normalizace dlouhodobým LPC spektrem na spektrum hlásky „a“ [7]

Kepstrum

Kepstrální analýza patří do homomorfní skupiny metod nelineárního zpracování řečových signálů. Tyto metody se hodí pro oddělování signálů, které vznikly konvolucí nebo násobením více složek. Používá se k oddělení budících a přenosových parametrů řečového signálu.

Komplexní kepstrum představuje inverzní Fourierovu transformaci logaritmu Fourierova obrazu vstupního signálu.

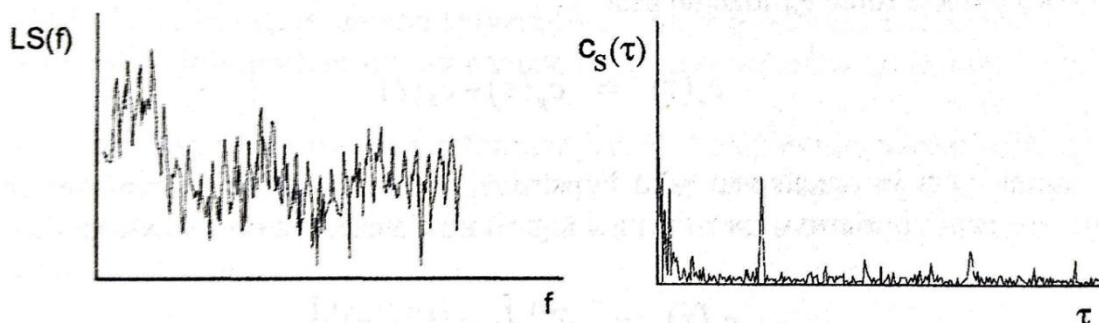


$$\begin{array}{llll}
 s(n)= & S(f)= & \log |S(f)| = & c_s(\tau)= \\
 g(n).h(n)= & G(f).H(f) & \log |G(f)| + \log |H(f)| & c_g(\tau) + c_h(\tau) \\
 \sum g(k).h(n-k) & & &
 \end{array}$$

konvoluce (časová oblast)	součin (kmitočtová oblast)	součet (kvefrenční oblast)
------------------------------	-------------------------------	-------------------------------

Obrázek 2.10 Systém k určení kepstra řečového signálu [7]

Na obrázku 2.11 je znázorněno kepstrum znělého segmentu řeči. Výrazná špička představuje základní periodu a nepravidelná funkce v okolí počátku odráží vliv formantů.



Obrázek 2.11 Logaritmické spektrum a odpovídající kepstrum znělého segmentu [7]

Kepstrální koeficienty c_m mohou být vypočítány také z LPC koeficientů pomocí rekurzivních vztahů

$$c_m = a_m + \sum_{k=1}^{m-1} \frac{k}{m} c_k a_{m-k} \quad \text{pro } m = 1, \dots, M \quad (2.33)$$

$$c_m = \sum_{k=m-M}^{m-1} \frac{k}{m} c_k a_{m-k} \quad \text{pro } m = M + 1, M + 2, \dots, \quad (2.34)$$

kde M je řád prediktoru.

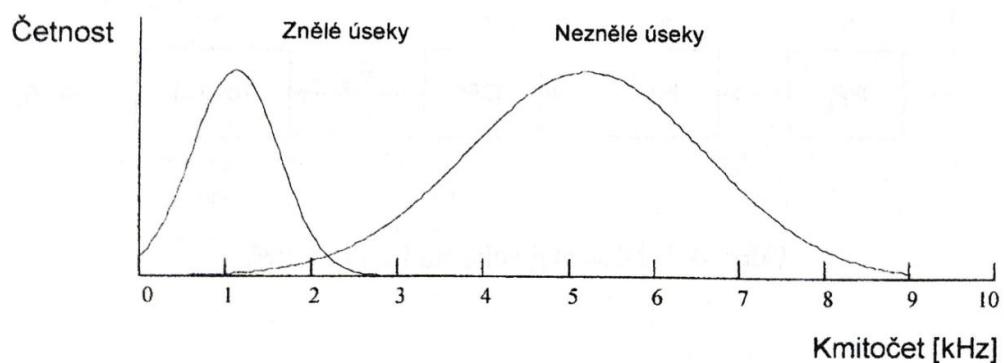
Metody kepstrální analýzy se prokázaly jako poměrně spolehlivé při rozpoznávání fonémů, ověřování mluvčích, určování stresu a jiných psychických stavů. Nevýhodou zůstává vysoká početní náročnost.

Počet průchodů nulou

U vzorkovaného signálu nastane průchod nulou, pokud mají dva sousední vzorky různé znaménko. Z pozorování řečového signálu zjistíme, že energetické části znělých hlásek leží převážně pod hodnotou kmitočtu 3 kHz, u neznělých hlásek jsou naopak výraznější vyšší kmitočty. Vysoké kmitočty vedou k vysokému počtu průchodu nulou. Můžeme odhadnout rozdělení energie v kmitočtovém spektru.

Zlepšením výsledků se dosáhne při měření počtu průchodů nulou v jednotlivých kmitočtových pásmech, namísto měření celého signálu. Pro rozpoznávání řeči se používá obvykle 8-20 pásem, pro přenos řeči pomocí vokodéru 10 -32 pásem. Často se používá banka filtrů s logaritmickými pásmy. Tím je možno hrubě odhadnout spektrální vlastnosti signálu.

Na základě metod krátkodobé funkce středního počtu průchodu signálu nulou byly vypracovány postupy na určení začátku a konce jednotlivých slov.



Obrázek 2.12 Četnost průchodu nulou u znělých a neznělých úseků řeči. [7]

3. VÝZNAMNÉ PUBLIKACE

Studie se zaměřují na různé parametry řeči od změn v intenzitě hlasu, rychlosti čtení, přeruky [10], přes změny základní hlasové frekvence F_0 [19], parametry tempa hlasu (rychlost slabik, rychlost artikulace, pauz), výřečnost [11], [12], nebo se zabývají záznamem samohlásek, kde se provádí podrobná analýza formantových frekvencí [13], [16]. Další studie zkoumá odstup signálu od šumu (SNR), poměr kmitočtů prvního a druhého formantu (F_1/F_2) [16], rychlost změn frekvencí F_0, F_1, F_2 [14]. *SNR a F_0* pomohly ke správnému rozdělení jednotlivých mluvčích. Další ze studií se zabývá zkoumáním prozodických řečových charakteristik [17] pro úspěšnou klasifikaci. Využívá pro tyto účely základní frekvenci, energii, *jitter* a *shimmer*. Pro úspěšnou detekci intoxikace se analyzovali rázové impulsy samohlásek [18].

V případové studii, která byla provedena po nehodě ropného tankeru Exxon Valdez [15] u Prince William Sound na Aljašce, byl prozkoumán záznam z rozhovorů a záznamů z komunikace mezi přístavem a lodí. Zkoumají se spektrální změny ve výslovnosti u anglického „d“ a „t“, frikativu „s“ v opakujícím se jméně „Exxon Valdez“ a číslovkách „třináct a šestnáct“. Zaměřila se také na změny základního hlasového tónu, rychlosti řeči a *jitteru*.

3.1 Řečové korpusy

Pro širší využití byla vytvořena sada Alcohol Language Corpus (ALC) [20], která obsahuje záznamy intoxikované a střízlivé německé řeči rozdělené podle osobních údajů (věk, pohlaví, hmotnost, výška, region, kuřák/nekuřák, návyk na alkohol). Korpus poskytuje nahrávky různých typů řeči (přečtená, příkazová, spontánní monology a dialogy). U nahrávek v intoxikovaném stavu se přečtená část skládá z pěti čísel, pěti adres, dvou jazykolamů, tři vybraných speciálních jazykolamů a pěti ovládacích příkazů užívaných v automobilu. Spontánní řeč obsahuje pět ovládacích příkazů, tři monology a dva dialogy. Délka monologů a dialogů je omezena na 60 sekund. Celkem je zaznamenáno 30 položek pro intoxikovaný stav a celková doba záznamu trvá nejvíce 15 minut, aby se zabránilo významným změnám způsobeným únavou nebo změnou naměřené hladiny alkoholu v krvi. Pro střízlivý stav bylo zaznamenáno 60 položek a doba záznamu trvá 30 minut. Aby se vyhovělo statistickým požadavkům analýzy, je celkový počet mluvčích stanoven na hodnotu 200-250. Nahrávky jsou rozděleny dále na čtyři věkové skupiny 22-27, 28-35, 36-50, 50+. Před zkouškou si každý mluvčí vybral hodnotu BAC, kterou chtěl subjekt intoxikací dosáhnout. Bylo odhadnuto množství požadovaného alkoholu pro každou osobu. Po konzumaci subjekt čekal 20 minut, než podstoupil test BRAC a test krve (BAC). Rozsah hodnoty BAC je v rozsahu 0.28 - 1.75 na ml. ALC využívají studie [16],[19],[21],[22].

Další studie využívají svých nahrávek.

Studie Sobella [10] rozděluje 16 subjektů přečtením standardního textu po konzumaci tří různých dávek alkoholu: nulové, střední a vysoké dávky.

Studie Künzela [11] zahrnuje 33 mužských jedinců, kteří popsali obrázek nejprve ve střízlivém stavu a následně na různých úrovních intoxikace.

U studie Cooneyho [13] se využilo 12 dobrovolníků, kteří vytvořili nejdříve 5s nahrávku samohlásky „a“, následně přečetly dvakrát jednu z vybraných vět a třikrát zopakovaly slova „me“, „you“ a „to“. Jednou za půl hodiny po intoxikaci byla změřena hladina alkoholu.

Studie Klingholze [14] vyhodnotila řeč 11 německy mluvících mužů (ve věku 25-35 let) po konzumaci alkoholu ve srovnání s 5 střízlivými kontrolami. Jedenáct mužských subjektů muselo číst připravený text.

Ve Sigmundově studii [18] je testováno 12 mužských jedinců. Vzorkem jsou samohlásky klasifikovány do dvou tříd „střízlivý“ a „intoxikovaný“ ($BAC > 1,0$ na ml).

Ve Levitově studii [17] se použila sbírka alkoholizovaných vzorků řeči sestavené na Policejní akademii v Hesensku v Německu. Obsahuje 120 záznamů (cca 87 minut) německé verze bajky "Slunce a severní vítr", kterou přečetlo 33 mužských mluvčích v různých stupních alkoholizace s hladinou alkoholu v krvi mezi 0 a 2,4 promile.

Pro specifické účely detekce alkoholu v řečovém signálu pro český jazyk je zapotřebí provést několik nahrávek mluvčích statisticky vyhovující požadavkům. Dosavadní korpusy neobsahují dostatek nahrávek v intoxikovaném stavu. Pro statistiku neintoxikovaného stavu řečníků by mohl být vhodný korpus EuroParl: European Parliament Proceedings Parallel Corpus 1996-2011 [23] obsahující 668,595 vět a 13,761,381 slov nebo Český národní korpus [24] se spontánním mluveným slovem (přes 7 milionů slov).

3.2 Parametry v časové oblasti

Parametry časového signálu vychází z hodnot vzorku nikoliv z hodnot spektra.

Energie signálu

V práci Sobella [10] byla amplituda (dB) řeči změřena na třech úrovních: střízlivý, mírná opilost $BAL \geq 0.79\%$ a vysoká opilost $Peak\ BAL < 0.79\%$. Jak je patrné z tabulky 3.1, hodnota amplitudy se v intoxikovaném stavu snížila oproti hodnotě amplitudy ve střízlivém stavu.

Tabulka 3.1 Amplituda řečového signálu u Sobella [10]

Míra opilosti	žádná	mírná $BAL \geq 0.79\%$	vysoká $BAL < 0.79\%$.
Amplituda[dB]	73.59	72.83	73.19

V dalších studiích se energie hlasu používá jako dílčí parametr v kombinaci s dalšími parametry.

Rychlost řeči

Změnami v tempu řeči ve střízlivém a intoxikovaném stavu se zabývají studie [10], [11], [12], [15], [19], [25]. V práci Sobella [10] bylo zjištěno, že při intoxikaci dochází k delšímu trvání čtení předloženého textu viz tabulka 3.2. Při čtení se při vysoké intoxikaci zhoršuje i artikulace. Pozorovatelé se shodli, že hlas při intoxikaci zrazuje.

Tabulka 3.2 Rychlost čtení u Sobella [10]

Míra opilosti	žádná	mírná BAL $\geq 0.79\%$	vysoká BAL $< 0.79\%$.
Rychlost čtení [s]	31.19	31.88	34.19

V práci Künzela [11] jsou prezentovány výsledky polospontánní řeči. Proměnné zahrnují výřecnost a parametry tempa hlasu (rychlost slabik, rychlost artikulace a pauz). Výsledky nasvědčují tomu, že všichni mluvčí vytvářejí při intoxikaci více vět, než když jsou střízliví. Parametry tempa hlasu jsou také vážně ovlivněny úrovněmi BRAC nad 0,08%. Tempo řeči, pokud jde o rychlost slabik, je obecně zpomalené. U nižších úrovní intoxikace však u některých subjektů dochází k opačnému účinku. Při měření rychlosti artikulace byly pauzy definovány jako přerušování signálu minimální délkou 100ms, pokud jsou vnímány také vyškoleným posluchačem. Na základě všech výše uvedených zjištění se zdá, že neexistuje žádný jednoduchý lineární vztah úrovní intoxikace a verbálními nedostatky, ale místo toho se objevuje třicestný rozdíl.

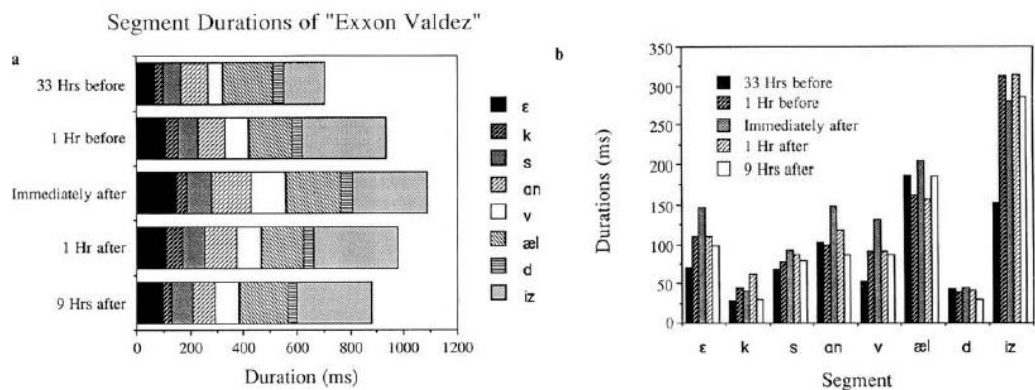
Pod 0,08% je účinek do značné míry euforický: Množství textu vytvořeného v popisu obrázku se zvyšuje, zatímco čas na jeho vyslovení není podstatný.

Při „střední“ úrovni intoxikace (0,12 až 0,2%) se množství textu zvyšuje, míra slabik a artikulace však výrazně klesá.

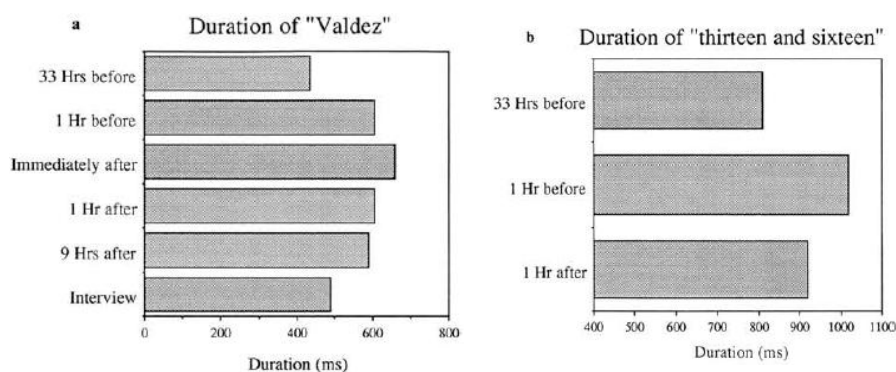
Při extrémních úrovních intoxikace ($> 0,2\%$) se množství textu podstatně nezvyšuje, zatímco doba trvání jejího vyslovení se dramaticky zvyšuje.

Studie Holliena [12] není veřejně dostupná. Primárními použitými znaky byly intenzita řeči, rychlost mluvy a schopnost plynulé řeči. Byly zjištěny menší změny mezi pohlavími, ale chyběla jim statistická významnost. Nakonec, i když lze vyvodit závěr, že určité změny v řečových suprasegmentálech nastanou jako funkce rostoucí intoxikace, nelze tyto vzorce považovat za univerzální, protože několik subjektů (přibližně 20%) nevykazovalo žádné (nebo negativní) změny.

V Johnsonově [15] případové studii nehody ropného tankeru Exxon Valdez kapitánova řeč v záznamech pořizovaných v době nehody je pomalejší než v záznamech pořizovaných 33 hodin před nehodou. Trvalé změny jsou možná nejspolehlivějšími parametry, které se našli v nahrávkách NTSB (National Transportation Safety Board), a naznačují, že kapitán Hazelwood mluvil v době nehody pomaleji než obvykle. Tyto změny v trvání jsou v souladu s laboratorními nálezy hlášenými pro řeč produkovanou pod vlivem alkoholu, v hlučném prostředí, ve stresu (u některých subjektů) a v simulacích emocí strach, hněv a smutek.

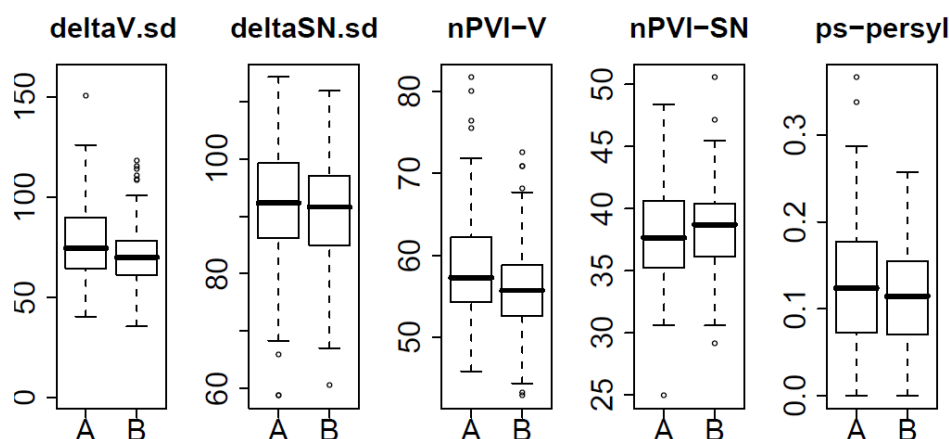


Obrázek 3.1 Trvání segmentů řeči ve frázi Exxon Valdez v různých časech nahrávání. a) Kumulativní doby trvání označující celkové prodloužení doby trvání. b) Doba trvání seskupené podle segmentů, které ukazují, u kterých segmentů se doba trvání zvýšila v době nehody [15]



Obrázek 3.2 a) Délka slova Valdez z pásek NTSB ve srovnání se stejným slovem vytvořeným v podobné pozici v televizním rozhovoru. b) Trvání fráze třináct a šestnáct ze záznamů pořízených třikrát v době nehody [15]

V Schielově studii [19] rozumíme funkci řeči jako časové vzorce znělých a neznělých segmentů v rámci nepřetržité řeči. Rytmické funkce jsou rozděleny na segmenty souhláskové, vokální a segmenty ticha (pauz). Předpokládá se, že poloha jádra slabiky je uprostřed shluku samohlásek. Vzdálenosti mezi jednotlivými jádry jsou větší než 500ms. Zkoumá se 17 různých rytmických funkcí, ale práce se zaměřila na 5 funkcí: Směrodatná odchylka (SD) trvání samohlásek ($\Delta V.sd$), SD vzdálenost mezi slabikami ($\Delta SN.sd$), průměrný rozdíl trvání po sobě jdoucích shluků samohlásek ($nPVI-V$), průměrný rozdíl vzdáleností po sobě jdoucích slabik ($nPVI-SN$) a pauzy menší než 1s ($ps-persyl$). V této studii je upozorněno na významný vliv mezi styly čtené a spontánní řeči.



Obrázek 3.3 5 rytmických funkcí pro intoxikovaný (A) a neintoxikovaný (b) projev [19]

V Schielově [25] práci bylo zjištěno, že intoxikované subjekty mají tendenci k hyperartikulaci, jak je patrné i z Obrázku. Studie nám dokazují, že rychlost řeči se zvyšuje s mírnou intoxikací a při vyšší intoxikaci se rychlost řeči zpomaluje.

3.3 Parametry ve frekvenční oblasti

Formantové příznaky

Studie Cooneyho [13] se zabývala analýzou hlasových frekvencí a získané výsledky z analýzy ukazují, že neexistuje žádná souvislost se změnami formantových frekvencí a úrovní intoxikace. Dospěli k závěru, že není pozorována žádná korelace mezi úrovní konzumovaného alkoholu a změnami charakteristik lidského hlasu.

Studie Klingholze [14] měřila poměr frekvencí 1. a 2. formantu, rychlost změn F_1 a F_2 . Parametr F_1/F_2 popisující artikulaci se u jednotlivců silně lišil. Měnil se pouze s vysokou hladinou alkoholu v krvi. Také ale považuje aplikaci akustické analýzy ve forenzní medicíně za neúčinnou.

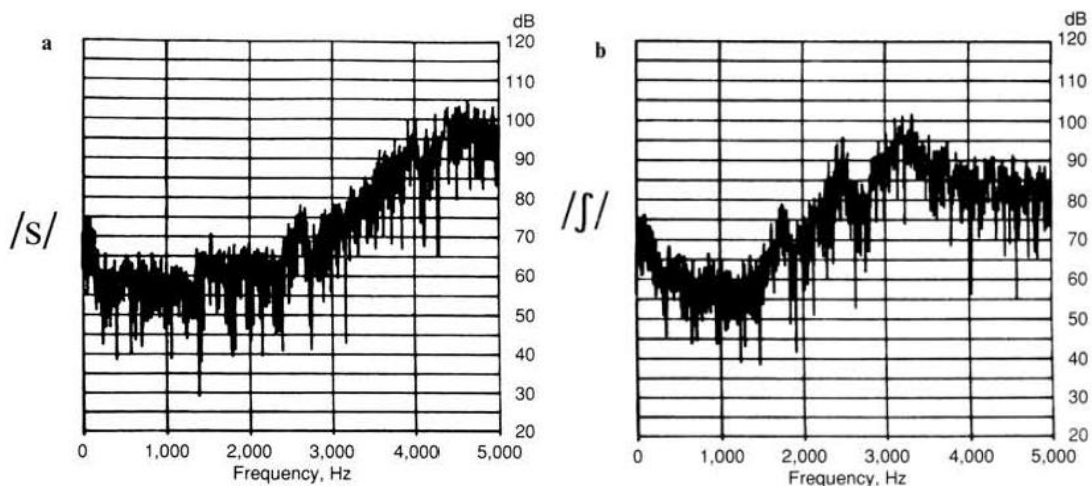
Ve studii Schiella [16] jsou analyzovány frekvence formantů F_1 až F_4 . Testuje se hypotéza, že samohlásky jsou více soustředěny v prostoru F_1/F_2 při intoxikaci, bylo ale zjištěno, že subjekty mají tendenci k hyperartikulaci. Využívá nástroj pro automatickou segmentaci MAUS (Munich Automatic Segmentation)² a poté formant tracker balíčku SNACK3. Metoda je využita vzhledem k množství řečníků, až 131, z databáze ALC. Je odvozen medián F_m a čtvrt kvantilová vzdálenost F_{qq} formantových frekvencí. Průměr těchto 8 hodnot pro všechny samohlásky v rámci jedné kombinace (pohlaví, věk, míra intoxikace...) se otestuje pomocí opakovaného měření. F_{1m} se významně zvyšuje u alkoholizované řeči. Ukazuje se, že je to významné pro hlásku „a“. Další F se výrazně neliší. F_{1qq} vykazuje slabý nárůst pro samohlásku „u“. F_{4qq} se významně zvyšuje pro všechny kombinace faktorů.

² https://www.phonetik.uni-muenchen.de/forschung/publikationen/ICPhS99_Schiel.pdf

³ <http://www.speech.kth.se/snack/>

Další příznaky

V Johnsonově případové studii [15] se uvádí změna ve frekvenčním výkonovém spektru. Spektrum „s“ ze záznamu pořízeného bezprostředně po nehodě a 1 hodinu po nehodě mají energetické vrcholy v oblasti od 2000 do 3000Hz. Spektrum „s“ zaznamenané 9 hodin po nehodě vrchol v oblasti 2000 až 3000Hz nemá. Dá se tedy usuzovat spektrální změna artikulace z „s“ na „ʃ“ ve slově „sea“ v záznamech v době hodinu před nehodou a v době nehody.



Obrázek 3.4 Výkonové spektrum a) /s/ a b) /ʃ/ [15]

Ve stejné studii se setkáváme s výpočtem *jitteru*. Vzhledem k množství šumu v pozadí se musel algoritmus⁴ upravit pomocí opravné techniky a low-pass filtrem (k odstranění vysokofrekvenčního šumu). Hodnota *jitteru* byla vyšší v čase blízko nehody. Dalo by se předpokládat, že nehoda byla faktorem vyvolávající stres, a tak je možné, že pozorované změny jsou výsledkem psychického stresu.

Je dobře známou skutečností, že různé kvality mluvené řeči mohou být ovlivněny faktory, jako je stres [27] a emoce [28].

3.4 Kmitočet základního tónu řeči

Ve Künzelově studii [11] je zjištěno, že vliv na průměr F_0 je zcela jasný: zatímco subjekty s méně než 0,08% alkoholu v krvi nevykazují žádnou významnou změnu, významné zvýšení F_0 je znatelné při vyšších úrovních opilosti. Totéž platí pro variabilitu F_0 .

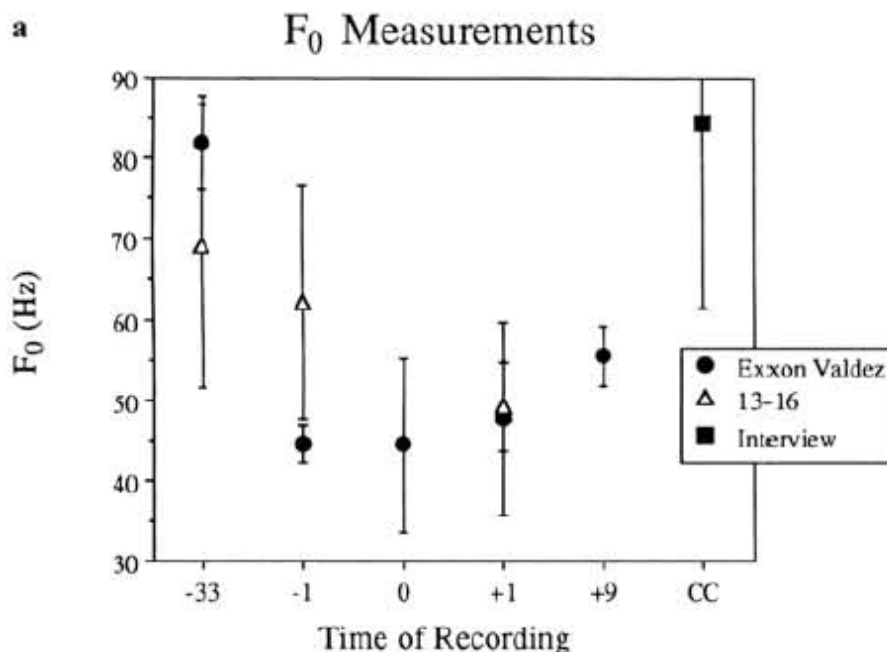
Studie Holliena [12] shrnuje zvýšení F_0 po intoxikaci jako statisticky významnou změnu. Zatímco studie Cooneyho [13] dospěla k názoru, že neexistuje žádná souvislost mezi základní frekvencí a úrovní konzumace alkoholu a že není pozorována žádná korelace mezi konzumací alkoholu a změnami charakteristik lidského hlasu.

Práce Klingholze [14] se zmiňuje o využití F_0 jako parametru pro využití ke správné diskriminaci jednotlivých subjektů. Podle studie má dlouhodobé hlasové úsilí podobné

⁴ Pinto, N. B.; Time, I. R.: Unification of perturbation measures in speech signals. J. acoust. Soc. Am. 87: 1278-1289 (1990).

účinky jako intoxikace alkoholem. Akustickou analýzu ve forenzní medicíně považuje za neúčinnou.

V Johnsonově případové studii [15] je zmíněno, že suprasegmentální účinek alkoholu je pozorován jako zvýšení základní frekvence F_0 , pokles F_0 (jen u některých subjektů). Frekvence hlasového tónu F_0 je proměnlivější při intoxikované řeči ve srovnání s kontrolním stavem. Zvýšenou variabilitu může způsobovat extrémnější intonace a chvění hlasu způsobenou stresem.



Obrázek 3.5 Průměrná F_0 napříč frází *Exxon Valdez*, *třináct* a *šestnáct* a z jedné věty v televizním rozhovoru [15]

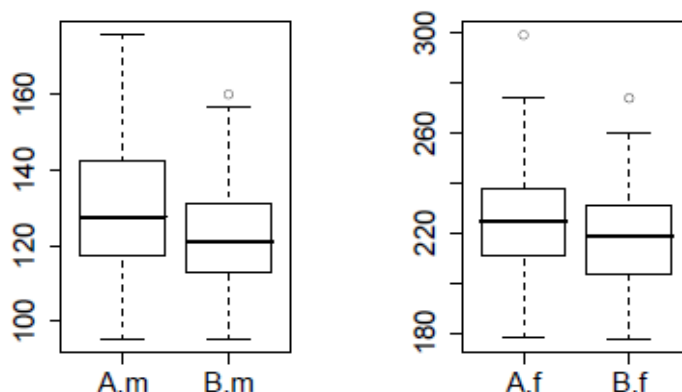
Obrázek 3.5 ukazuje F_0 zprůměrované napříč frází *Exxon Valdez* ze všech vzorků, frází *třináct*, *šestnáct* ze tří záznamů NTSB a jednu větu z televizního rozhovoru. Osa x zobrazuje čas nahrávky, kde 0 značí dobu nehody a CC značí dobu televizního rozhovoru. Měření se provedlo z každé ze čtyř samohlásek v *Exxon Valdez*. Každý bod pro *třináct*, *čtrnáct* je založen na měření aspoň dvou samohlásek. V interview se zprůměrovala věta „I would say the same for the state of Alaska, They come after me, hamer and tong.“

Jak je patrné z obrázku 3.5., F_0 je dramaticky nižší v době nehody. Toto zjištění je totožné pro intoxikaci, zármutek a stres.

V Schielově studii [19] zjišťujeme že, ve většině případů je u F_0 hlášeno zvýšení po požití alkoholu. Základní frekvence se vypočítala pomocí detekčního algoritmu⁵ Vincenta Schaefera a odstranily se části, které se vyhodnotili jako neznělé. Ze zbývajících částí byl vypočten medián f_m a čtvrt-quantilový rozsah Δf_{qq} pro každého mluvčího. Obrázek 3.6 ukazuje hodnoty f_m intoxikované a střízlivé řeči samostatně pro mužské a ženské subjekty. Pohlaví řečníků nemá žádný významný vliv alkoholu na

⁵ Schafer-Vincent K (1983) Pitch period detection and chaining: method and evaluation. *Phonetica* 1983, Vol 40, pp. 177-202.

změny v řečovém signálu. U ženských subjektů se spíše jednotně zvyšuje průměrná základní frekvence, zatímco u mužských subjektů se pohybuje v obou směrech.



Obrázek 3.6 F_0 intoxikované (A) a střízlivé (B) řeči dělená pro muže(m) a ženy(f) [19]

Studie Baumeisterové [26] zkoumá kontury projevu základní frekvence F_0 odvozené z korpusu ALC. Medián F_0 , rozsah F_0 a kontury F_0 jsou analyzovány při intoxikaci v závislosti na pohlaví a věku. Kontury se porovnávají jak přímo (střední kvadratická chyba, statistická korelace nebo euklidovská vzdálenost ve spektrálním prostoru kontury), tak pomocí parametrizace kontury pomocí diskretní kosinové transformace a prvního a druhého momentu spodního spektra kontury. Výsledky částečně potvrzují dřívější zjištění, tj. průměr a rozsah F_0 jsou většinou zvýšeny intoxikací, a také naznačují, že většina řečníků nenásleduje obecný trend, ale vykazuje idiosynkratické změny F_0 . Kontury F_0 se významně liší od intoxikace, ale podrobnější analýza nemohla přiřadit tyto změny konkrétním obecným změnám formy, jako je pokles nebo zakřivení. Výsledky naznačují, že není možné předpovědět intoxikaci z F_0 u jediného modelu napříč různými řečníky. Místo toho je navržen model závislý na mluvčím, který zohledňuje chování jednotlivých mluvčích.

Jak je patrné z předešlých studií dochází v určité míře ke zvýšení základního hlasového tónu F_0 a ke zvětšení horního rozsahu pro jednotlivé mluvčí.

3.5 Patenty

Jedním z využití detekce alkoholu v řečovém signálu je použití keprálních koeficientů střízlivé a intoxikované řeči pro zjišťování opilosti řidiče. Podle výsledku se vozidlo nastartuje, nebo zablokuje [29].

Další patent detekuje rozdíl signálů pomocí rychlé Fourierovy transformace původního signálu a rozdílového signálu [30]. Hledá se rozdíl strmosti ve frekvenční doméně. Podle toho rozhodne řidič nebo operátor na dálku, zda je řidič intoxikován a zda jde zabránit možné nehodě.

Hideakiho patent porovnává slova s databankou slov a při změně se vypočítá pravděpodobnost od prahové hodnoty [31]. Pokud zkoumané slovo neprojde testem, upozorní systém vozu řidiče o nepovoleném řízení vozidla. Pokud jsou odchylky

vysoké, jednotka zakáže nastartování vozu. Test lze provést pouze u registrované osoby, nezaregistrovaná osoba vozidlo nenastartuje.

U patentu Leea [32] dochází ke kontrole hlasu prostřednictvím telekomunikační sítě propojené se službou hlasové analýzy, která uživateli prostřednictvím displeje oznámí stav intoxikace. Využívá se spektrální analýzy a 13 funkcí (informace o základním hlasovém tónu, energii, pět formantů...), které se uloží do uživatelské databáze k rozpoznání jedince. Detekuje se koncentrace alkoholu od 0.05% BAC.

4. IMPLEMENTACE

4.1 Test alkoholu

Za účelem vypočtení požadované hodnoty alkoholu testovaného subjektu bylo využito alkoholové kalkulačky implementované v aplikaci ALCTest. Pro přesné stanovení hladiny alkoholu v krvi pozorovaného subjektu byl zapůjčen profesionální alkohol tester Dräger Alcotest 7410.



Obrázek 4.1 Alkohol tester Dräger Alcotest 7410

Nejdříve se zapsala informace o hmotnosti, věku a pohlaví subjektu. Pomocí výpočtu (1.1) a (1.2) se zjistilo potřebné množství alkoholu k dosažení požadované intoxikace. Po konzumaci nápoje, dle vlastního výběru testované osoby (víno, pivo, ...), se čekalo 20 minut na ustálení hladiny alkoholu v krvi testované osoby. Po uplynutí této doby se za pomoci alkohol testeru změřila přesná hladina alkoholu. Pokud došlo k nedosažení požadovaného limitu, bylo podáno další menší množství nápoje a opět se čekalo 20 minut. Pokud se stalo, že limit byl překročen, doporučilo se testovanému subjektu napít se pitné vody a měření se zopakovalo. Ve většině případů se pomocí vypitého nealkoholického nápoje dosáhlo snížení měřené hodnoty.

Testování obsahovalo 5 kroků, první měření se provedlo při hodnotě 0‰, druhé měření se provedlo po dosažení hodnoty z rozsahu 0‰-0.3‰, další z rozsahu 0.3‰-0.6‰, 0.6‰-0.9‰ a pokud byla testovací osoba ochotna, poslední měření bylo v rozsahu 0.9‰ a výše. Pokud byla testovaná osoba profesí herec nebo se v uměleckém prostředí pohybovala, byla požádána i o nahrávku hrané opilosti ještě před konzumací alkoholického nápoje.

Z tabulky 4.1 je patrné, že se Alkohol tester začal používat až od vzorku 15. Bylo to způsobeno nefunkčností předešlého alkohol testeru. Proto se přistoupilo k výpočtům z několika alkoholových kalkulaček, převážně z alkoholové kalkulačky implementované v aplikaci ALCTest, která byla vložena do aplikace pro účely určení potřebného množství alkoholu k dosažení potřebné hranice intoxikace.

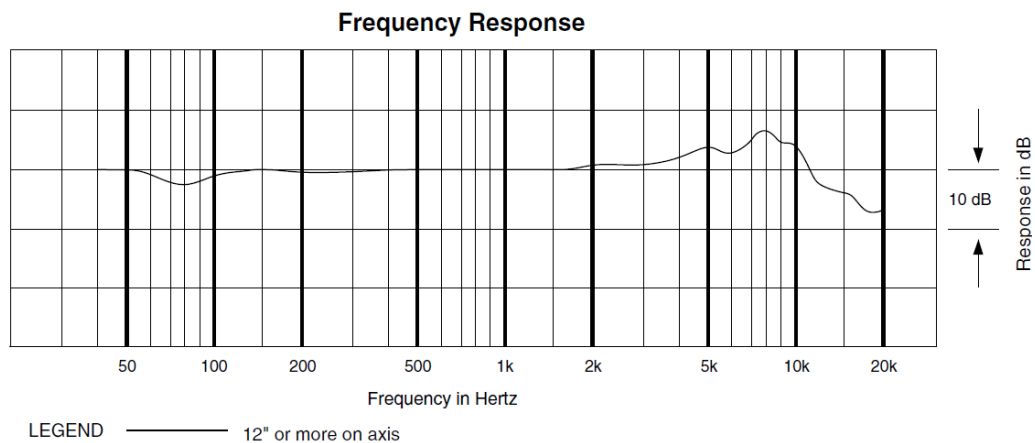
4.2 Použitá technika pro získání databáze

Pro nahrávání řečového signálu byl využit zapůjčený mikrofon Audio-Technica PRO 37.



Obrázek 4.2 Použitý mikrofon Audio-Technica PRO 37 a externí zvuková karta M-Audio M-track Solo.

Kondenzátorový mikrofon je potřeba napájet fantomovým napájením. K tomuto účelu byla pořízena externí zvuková karta M-Audio M-Track Solo se vzorkovací frekvencí 48kHz/16bit. Zvuková karta se propojila přes USB 2.0 k počítači s nahrávacím programem a na vstupu signálu 1 se nastavila hodnota 8 pro každou nahrávku a testovanou osobu.



Obrázek 4.3 Kmitočtová charakteristika mikrofону Audio-Technica PRO 37 [37]

4.3 Databáze nahrávek

Pro účely zkoumání vlivu alkoholu na analyzované parametry bylo nutné vytvořit databázi nahrávek, které obsahují střízlivou, intoxikovanou a předstíranou intoxikovanou řeč. Pro tyto účely bylo osloveno více jak 50 osob.

Nahrávka byla prováděna v domácím klidném prostředí za pomoci vyvíjené aplikace ALCTest a nahrávací techniky. Nahrávky jsou dále uloženy do jednoho mono kanálu ve formátu WAV se vzorkovací frekvencí 44100 Hz/16bit.

Nahrávky jsou pro rozdělení označeny podle dalších parametrů.

Označení jednotlivých nahrávek k uložení

1. Parametr - Nahrávaná část:

Samohlásky: **A, Á, E, É, I, Í, O, Ó, U, Ú,**

Telefonní čísla: **T**

Rozmlouvadla: **B, S, D, N**

Čtený text: **R**

Volný text: **F**

2. Parametr - Pořadové číslo 01-40

3. Parametr - Pohlaví M, Z

4. Parametr - Věk A(18-24), B(25-34), C(35-44), D(45-54), E(55-64), F(65+)

5. Parametr - Hmotnost A(45-54), B(55-64), C(65-74), D(75-84), E(85-94), F(95+)

6. Parametr - Intoxikace A(0‰), B(0‰-0.3‰), C(0.3‰-0.6‰), D(0.6‰-0.9‰), E(0.9‰+), P(Předstíraná)

Nahrávka ve výsledku může být **Á12MBCD.wav.**, kde se uložila nahrávka samohlásky **Á** do souboru **12** od mužského subjektu **M** ve věku mezi 25-35 lety **B** o hmotnosti mezi 65-74 kg **C** a dosažená intoxikace byla v rozmezí **D** 0.6 - 0.9‰.

Při všech nahrávkách lze takto získat od jedné osoby 102 vzorků. Pokud se nenahrává předstíraná intoxikovaná řeč, obsahuje soubor 85 nahrávek. Celkem tímto způsobem bylo získáno 3237 nahrávek od 23 mužských dobrovolníků a 17 ženských dobrovolnic. Bylo utvořeno 40 souborů, každý pro jednoho nahrávajícího.

V tabulce 4.1 je vidět přehled testovaných osob, jejich váha, věk, dosažená intoxikace a hodnoty promile v jednotlivých bodech měření a nahrávání. Nejmladší testovaná osoba měla 18 let a nejstarší 62 let. Průměrný věk je 38let. U mužů je průměrný věk 39,5 roků a u žen 35,9 roků. Předstírané nahrávky se povedlo získat od 6 mužských subjektů.

Text nahrávek

Samohlásky: (každá samostatně), dlouhá asi 3s:

A Á E É I Í O Ó U Ú

Telefonní čísla:

123 456 789

602 686 800

721 454 260

778 797 346

Rozmlouvadla: 3x každé samostatně rychle:

Blb v úprku brk u krbu

Šťávička ze švestiček

Dalajláma v lomu láme slámu

Nesnese se se sestrou

Čtený text:

DOBŘÝ VOJÁK ŠVEJK NA POLICEJNÍM ŘEDITELSTVÍ

Sarajevský atentát naplnil policejní ředitelství četnými oběťmi. Vodili to jednoho po druhém a starý inspektor v přijímací kanceláři říkal svým dobráckým hlasem: „Von se vám ten Ferdinand nevyplatí!”

Když Švejka zavřeli v jedné z četných komor prvního patra, Švejk našel tam společnost šesti lidí. Pět jich sedělo kolem stolu a v rohu na kavalci seděl, jako by se jich stranil, muž v prostředních letech.

Švejk se počal vyptávat jednoho po druhém, proč jsou zavřeni. Od těch pěti sedících u stolu dostal takřka úplně stejnou odpověď: „Kvůli Sarajevu!” — „Kvůli Ferdinandovi!” — „Kvůli té vraždě na pana arcivévodovi!” — „Pro Ferdinanda!” — „Za to, že pana arcivévodu odpravili v Sarajevu!”

Šestý, který se těch pěti stranil, řekl, že s nimi nechce nic mít, aby na něho nepadlo nijaké podezření, on že tu sedí jen pro pokus loupežné vraždy na pantátovi z Holic.

Volný text: nahrávka trvající 1 minutu

Vyprávění na téma „Moje nejlepší dovolená“.



Obrázek 4.4 Ilustrační fotografie z průběhu nahrávání.

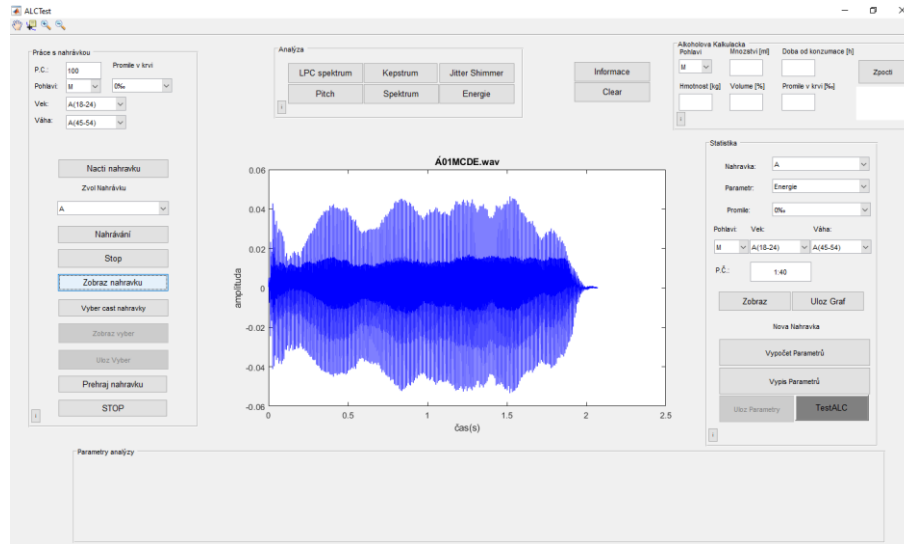
Z důvodu rušení pohybem papíru byl do programu ALCTest implementován celý čtený text a při nahrávání se daly instrukce číst přímo z obrazovky počítače.

Tabulka 4.1 Dostupné informace o nahrávaných subjektech

P.Č.	pohlaví	věk	kg	0‰-0.3‰	0.3‰-0.6‰	0.6‰-0.9‰	0.9‰+	Předstíraná	Test
1	M	37	75	0.12	0.5	0.88	1.300	ano	ALCTest
2	Z	32	73	0.15	-	-	-	ne	ALCTest
3	M	36	97	0.2	0.4	0.75	-	ne	ALCTest
4	Z	35	60	0.12	0.25	-	-	ne	ALCTest
5	M	40	98	0.12	0.46	0.79	1.00	ne	ALCTest
6	M	48	70	0.29	0.58	0.87	-	ano	ALCTest
7	Z	49	72	0.3	-	-	-	ne	ALCTest
8	M	40	103	0.2	0.5	0.75	-	ne	ALCTest
9	M	34	80	0.18	0.45	0.7	0.95	ano	ALCTest
10	M	43	98	0.2	0.42	0.76	0.97	ano	ALCTest
11	Z	39	56	0.25	0.5	0.75	1.000	ne	ALCTest
12	M	33	79	0.29	0.58	0.76	1.045	ne	ALCTest
13	Z	28	63	0.27	0.47	0.67	1.012	ne	ALCTest
14	M	43	90	0.18	0.44	0.7	0.95	ne	ALCTest
15	M	53	110	0.23	0.48	0.68	1.250	ne	Dräger
16	M	47	84	0.16	0.39	0.79	1.090	ne	Dräger
17	M	24	90	0.17	0.41	0.87	1.080	ne	Dräger
18	Z	37	73	0.3	0.57	0.8	1.020	ne	Dräger
19	M	43	99	0.11	0.36	0.86	1.070	ne	Dräger
20	M	34	105	0.21	0.52	0.62	1.150	ne	Dräger
21	Z	31	94	0.28	0.54	0.82	1.104	ne	ALCTest
22	M	32	102	0.07	0.32	0.89	1.030	ne	Dräger
23	Z	33	60	0.12	0.37	0.61	0.9	ne	Dräger
24	M	62	98	0.24	0.37	0.61	-	ne	Dräger
25	M	30	86	0.19	0.47	0.85	1.180	ne	Dräger
26	Z	27	65	0.11	0.3	0.75	-	ne	Dräger
27	M	38	83	0.14	0.48	0.72	0.9	ano	Dräger
28	Z	31	73	0.06	0.3	0.55	0.89	ne	Dräger
29	M	33	90	0.15	0.58	0.68	0.9	ne	Dräger
30	Z	32	72	0.08	0.43	0.8	1.00	ne	Dräger
31	Z	37	60	0.23	0.53	0.89	1.360	ne	Dräger
32	Z	38	58	0.11	0.41	0.83	1.400	ne	Dräger
33	M	50	80	0.1	0.36	0.79	0.9	ne	Dräger
34	Z	18	55	0.28	0.31	0.76	0.95	ne	Dräger
35	M	34	70	0.1	0.4	0.69	0.94	ne	Dräger
36	M	34	86	0.21	0.38	0.89	1.190	ano	Dräger
37	Z	47	78	0.24	0.43	0.73	0.91	ne	Dräger
38	Z	58	60	0.3	0.44	0.65	1.020	ne	Dräger
39	Z	39	62	0.2	0.46	0.84	0.9	ne	Dräger
40	M	41	105	0.23	0.45	0.87	1.180	ne	Dräger

4.4 Program ALCTest

Pro zpracování diplomové práce byla v prostředí GUI Matlab vyvíjena aplikace ALCTest. Obsahuje nahrávací prostředí, prostředí analýzy, statistické prostředí a alkoholovou kalkulačku. Další informace a grafy se vyobrazují ve středu aplikace nebo ve spodní části. Každá sekce má vlastní informační tlačítko, kde se dají přečíst potřebné informace a instrukce.



Obrázek 4.5 Aplikace ALCTest

Sekce Práce s nahrávkou

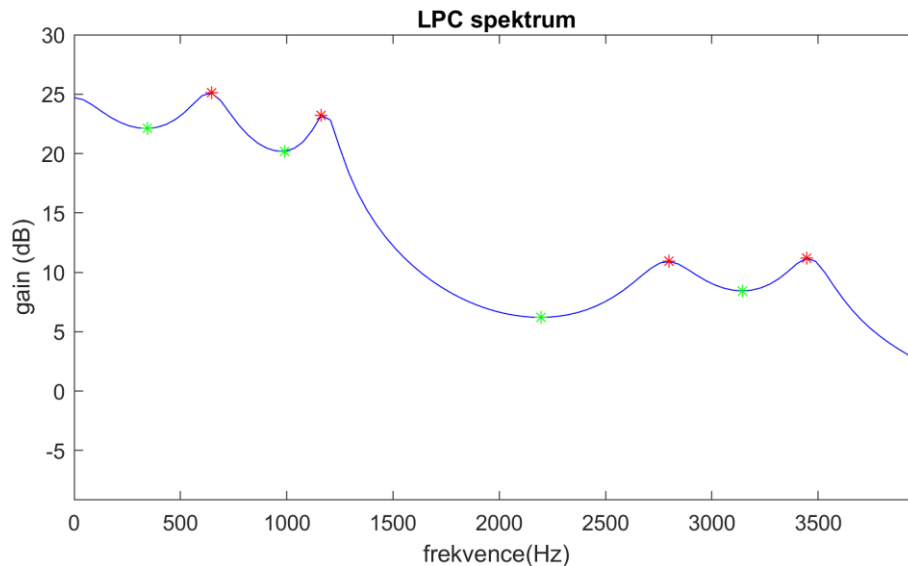
V sekci Práce s nahrávkou je možné zvolit požadovanou nahrávku a po vložení všech potřebných údajů o testované osobě ji vytvořit pomocí tlačítek „Nahrávání“ a „Stop“. Před každým počátkem nahrávání se na prostřední obrazovce vypíše samostatný text a 3 vteřinový odpočet do začátku nahrávání, který je doprovázen krátkým zvukovým signálem. Stejná nahrávka se může provést 3x, poté se objeví informace o existenci již třetí totožné nahrávky.

Pomocí tlačítka „Načti nahrávku“ lze načíst libovolnou nahrávku ve formátu wav. Načtenou nahrávku lze zobrazit pomocí tlačítka „Zobraz nahrávku“. Pomocí tlačítka „Vyber část nahrávky“ lze ořezat nahrávku určením začátku a konce nahrávky. „Ulož výběr“ uloží ořez předešlé nahrávky. Tento výběr lze opět zobrazit pomocí „Zobraz nahrávku“ nebo „Zobraz Výběr“. Tato část je vhodná pro zkracování delších nahrávek nebo k ořezání začátků a konců neúčinných částí nahrávky.

K přehrání načtené nahrávky slouží tlačítko „Přehraj nahrávku“ a „STOP“.

Sekce Analýza

V sekci Analýza lze provést LPC analýzu pomocí tlačítka „LPC spektrum“. V prostřední části se vyobrazuje LPC spektrum s označením formantů a antiformantů. V sekci Parametry analýzy se vypisují hodnoty F_1 až F_3 , průměr prvních tří formantů, poměr frekvencí F_1 a F_2 , M_1 až M_3 a počet průchodů nulou za dobu nahrávky.



Obrázek 4.6 LPC spektrum aplikace ALCTest

Před výpočtem provedeme převzorkování a normalizaci signálu:

```
%% převzorkování
y=resample(y,10000,Fs);
Fs=10000;
%% normalizace
if abs(max(y))>abs(min(y))
    ymax=(max(y));
else
    ymax=abs(min(y));
end
y=y/ymax;
```

Pro výpočet LPC spektra se využije příkazu:

```
a=lpc(x1,M);
```

Pro výpočet formantů a šířek pásem se využije postupu:

```
r=roots(a);
r = r(imag(r)>=0);
angz = atan2(imag(r),real(r));

[frqs,indices] = sort(angz.*(Fs/(2*pi)));
bw = -1/2*(Fs/(2*pi))*log(abs(r(indices)));
```

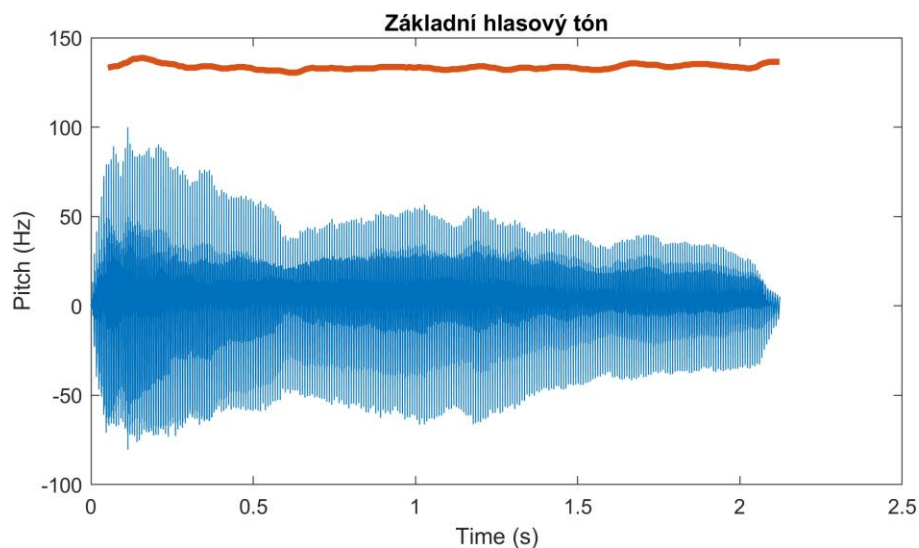
Pro výpočet průměru prvních tří formantů a podílu prvního a druhého formantu se využije postupu:

```
f13=(ffreq(1)+ffreq(2)+ffreq(3))/3;
f12=(ffreq(1)/ffreq(2));
```

Pro určení počtu průchodu nulou v nahrávce se používá postupu:

```
p1=sign(x); % urceni znamenska v jednotlivych prvcich vektoru y
for ind=1:length(x)-1
if p1(ind)+p1(ind+1)==0
z(ind)=1;
else
z(ind)=0;
end
end
end
pocet=sum(z);
```

Po stlačení tlačítka „Pitch“ se zobrazí načtená nahrávka a základní hlasový tón F_0 . V sekci Parametry analýzy se vypíše průměrná hodnota F_0 jeho maximum a minimum.

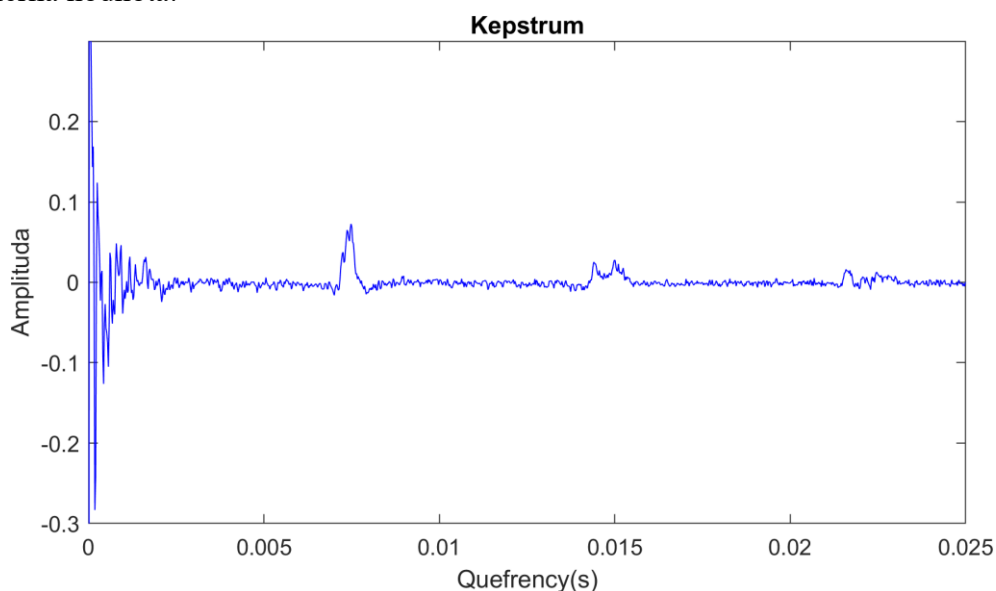


Obrázek 4.7 Základní hlasový tón aplikace ALCTest

Pro výpočet základního hlasového tónu se využije příkazu:

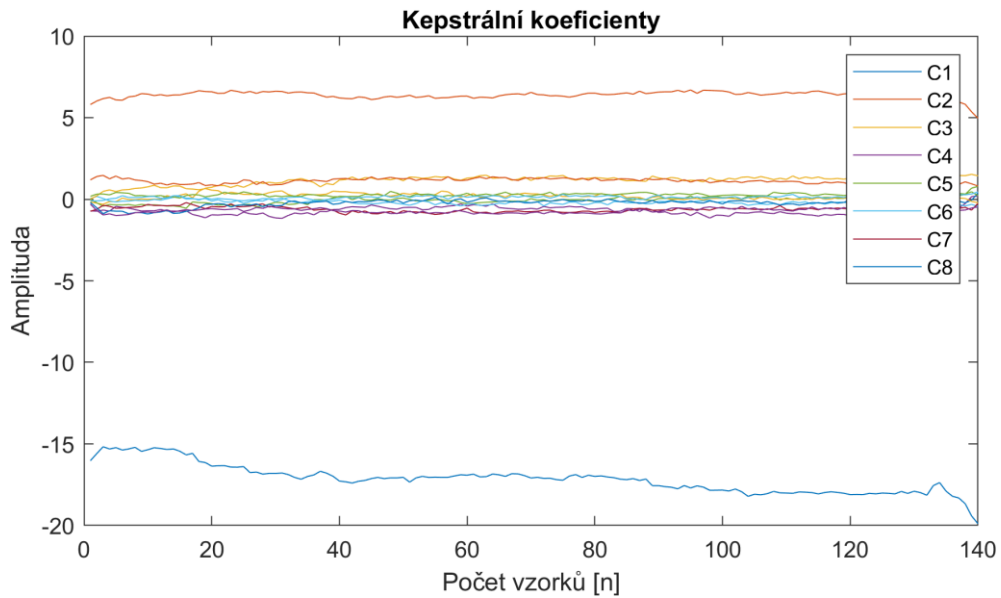
```
[f0,idx] = pitch(x,Fs);% Základní hlasový tón
```

Po prvním stlačení tlačítka „Kepstrum“ se vykreslí Kepstrum, ze kterého se dá odečíst např. základní hlasový tón. V obrázku 4.8 je patrný peak mezi časem 0.005 a 0.01. V sekci Parametry analýzy se vypíše maxima kepstrálních koeficientů C_1 až C_8 a jejich průměrná hodnota.



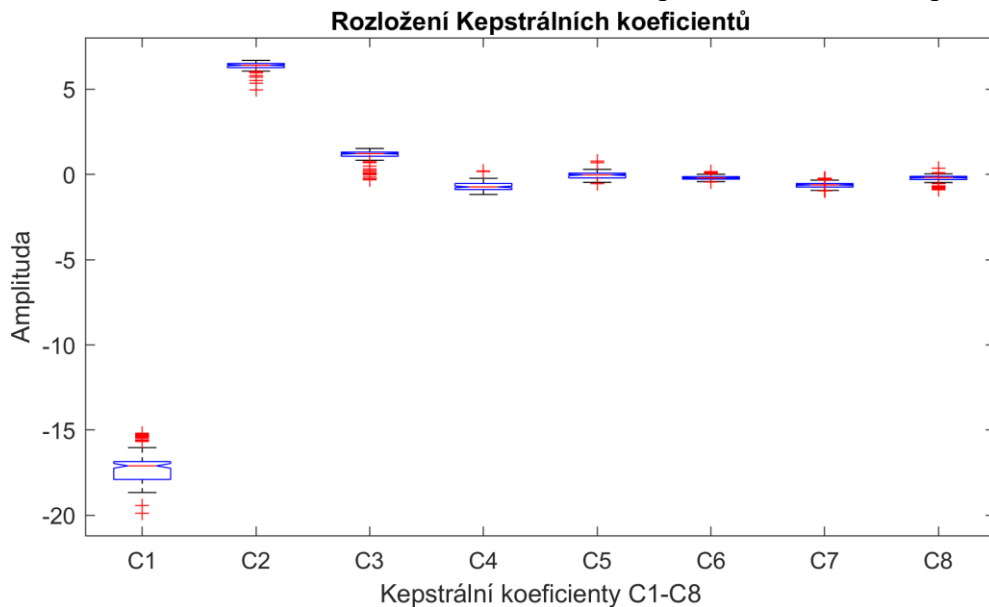
Obrázek 4.8 Kepstrum aplikace ALCTest

Po druhém stisknutí tlačítka „Kepstrum“ se vykreslí kepstrální koeficienty C_1 až C_8 .



Obrázek 4.9 Kepstrální koeficienty aplikace ALCTest

Po třetímu stlačení tlačítka “Kepstrum” se vyobrazí rozložení hodnot kepstrálních koeficientů C_1 až C_8 . Po dalším stisknutí tlačítka se opět zobrazí reálné Kepstrum.



Obrázek 4.10 Rozložení kepstrálních koeficientů v aplikaci ALCTest

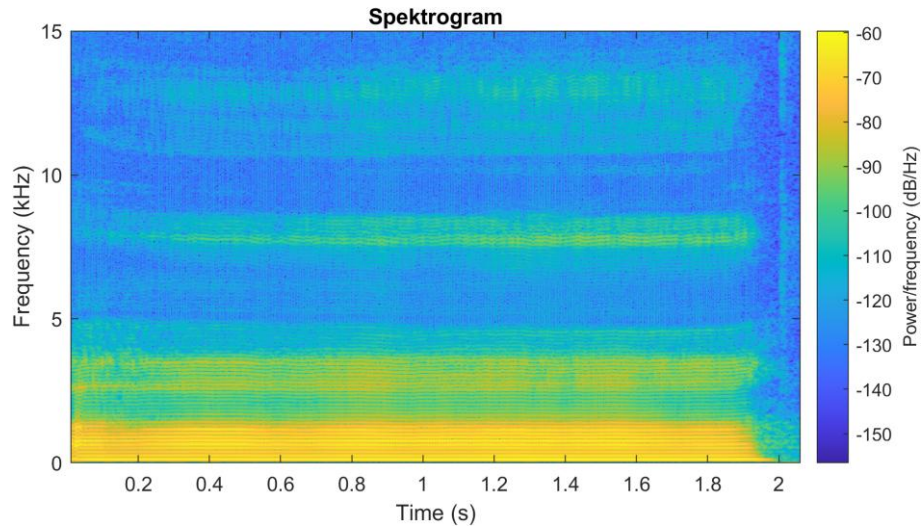
Pro výpočet kepstra se využije příkazu:

```
g = rceps(audioIn); %vypocet realneho kepstra
```

Pro výpočet kepstrálních koeficientů se využije příkazů:

```
windowLength = round(0.03*fs);%velikost okna 30ms
overlapLength = round(0.015*fs);%presah 15ms
S = stft(audioIn,"Window",hann(windowLength,"periodic"),
"OverlapLength", overlapLength,"FrequencyRange","onesided");
S = abs(S);
filterBank = designAuditoryFilterBank(fs,'FFTLenght',windowLength);
melcc = filterBank*S;
melSpec = cepstralCoefficients(melcc);
```

Pomocí tlačítka “Spektrum” je vyobrazen spektrogram nahrávky v čase.

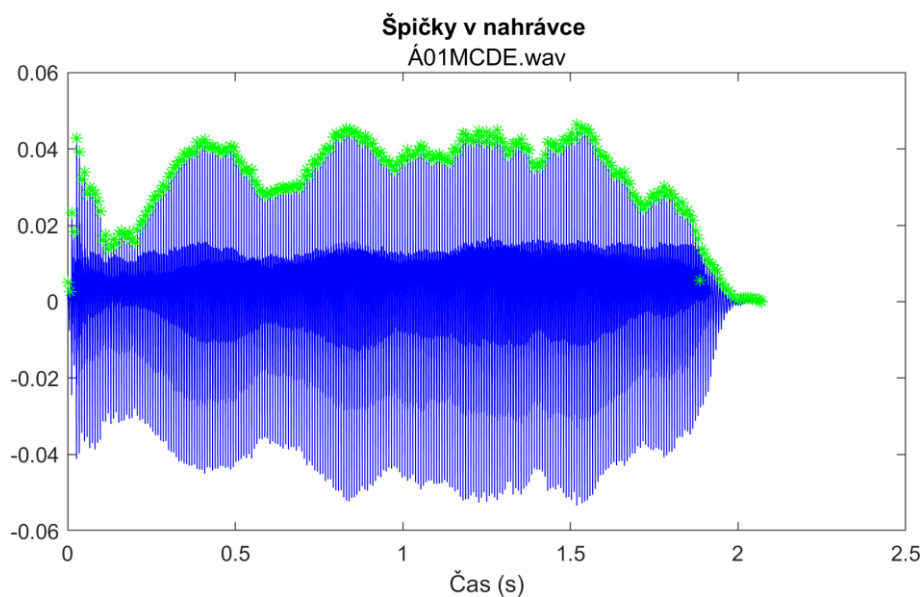


Obrázek 4.11 Spektrogram nahrávky aplikace ALCTest

Pro zobrazení spektrogramu se využívá postupu:

```
segmentlen = 1000;  
noverlap = 900;  
NFFT = 1280;  
Fs=handles.audioSampleRate;  
y=handles.audioBpS;  
spectrogram(y, segmentlen, noverlap, NFFT, Fs, 'yaxis')
```

Pomocí tlačítka „Jitter Shimmer“ se vykreslí signál s označením špiček v nahrávce. Podle těchto špiček se vypočítává *JitterABS*, *JitterREL*, *JitterRAP*, *JitterAPQ5* a *ShimmerLOK*, *ShimmerDB*, *ShimmerAPQ3*, *ShimmerAPQ5*. Jejich hodnoty se vypíší v sekci Parametry analýzy.



Obrázek 4.12 Zobrazení špiček v nahrávce pro výpočet *Jitteru* a *Shimmeru*.

Pro výpočet *Jitteru* a *Shimmeru* se nejdříve využije nalezení znělých nahrávek v hlase za pomoci detekce hlasu:

```
detectSpeech(data, Fs);
```

Využijeme nalezení základního hlasového tónu k vybrání relevantních špiček v nahrávce:

```
[peaks, t0]=findpeaks(x, Fs, 'MinPeakDistance', t0mean*0.93);
```

Špičky se ještě vyberou z rozsahu okolí základního hlasového tónu. Odečítáním sousedních špiček získáme sadu period $tt(k)$ a amplitud $peaks$, které se následně použijí pro výpočet *Jitteru* a *Shimmeru*.

```
%% jitterABS
for k=2:length(t0)
    pcd(k)=(abs(tt(k)-tt(k-1)));
end
pcdch=(pcd(2:end));
lengthpcdch=length(pcdch);
pch=sum(pcdch);
jittabs=pch/(lengthpcdch);

%% jitterREL
pdc=sum(tt);
pdch=pdc/(lengthpcdch);
jittrel = (jittabs/pdch)*100;

%% jitter(RAP)
for m=2:l-1
    pdirap1(m-1)=(abs(tt(m)-((tt(m-1)+tt(m)+tt(m+1))/3)));
end
pdirap2=sum(pdirap1);
pdirap=pdirap2/(l-2);
jittrap = (pdirap/pdch)*100;

%% jitterppq5
for i= 3:(l-2)
    pom_apq(i-2) = (abs((tt(i))- ((tt(i-2)+tt(i-1)+tt(i)+tt(i + 1) +
    tt(i+2) ) / 5)));
end
jittppq5 = (((sum((pom_apq)))/(l-4))/(pdch))*100;

%% shimm_local
peaks=abs(peaks);
lpeaks=length(peaks);

for pp = 2:lpeaks
    shimmloc1(pp)=(abs(peaks(pp)-peaks(pp-1)));
end
shimmloc2=shimmloc1(2:end);
shimmloc3=sum(shimmloc1);
lpeak3=length(shimmloc2);
shimmloc4=shimmloc3/lpeak3;
ppeaks=sum(peaks)/lpeaks;
shimm_local = (shimmloc4/pppeaks)*100;

%% shimmdB
peaksdb=pow2db(peaks);
for i = 2:(length(peaks))
    pom(i-1) = abs(peaksdb(i)-(peaksdb(i-1)));
```

```

end
pomdb=sum(pom);
shimmdb = ((sum(pomdb))/(length(pom)));

%% shimm_apq3
for i = 2:(length(peaks) - 1)
pom_apq3(i-1) = (abs(peaks(i) - ((peaks(i-1)+peaks(i)+
peaks(i+1))/3)));
end
shimm_apq3 = (((sum(pom_apq3))/(length(peaks) - 2))/(mean(peaks))) *
100;

%% shimm_apq5
for i = 3:(length(peaks) -2)
pom_apq5 = (abs((peaks(i)) - ((peaks(i-2)+peaks(i-
1)+peaks(i)+peaks(i+1)+peaks(i+2))/5)));
end
shimm_apq5 = (((sum(pom_apq5))/((length(peaks)-4))/(mean(peaks)))
*100;

```

Pomocí tlačítka “Energie” vypočítá aplikace energii v nahrávce a vypíše ji v sekci Parametry analýzy.

Pro výpočet Energie v nahrávce se použije postupu:

```

DT = class(x);
E = abs(sum(x.^2));
E(E==0) = realmin(DT);
logE = log(E);

```

Všechny takto vypočtené parametry se uloží. V sekci “Statistika” je možné je nechat všechny vypsát a uložit do souboru data.mat a data.xls.

Sekce Statistika

V sekci Statistika lze vykreslit parametr zprůměrovaný ze všech nahrávek nebo jen z jednoho souboru po zadání údajů o nahrávce, výběru zkoumaného parametru, určité míry intoxikace, pohlaví, věku a váhy. U věku, váhy a promile lze zadat hodnotu vše. Tím se vykreslí trend parametru na změně intoxikace od 0‰ do 1‰. A pokud jsou vybrané všechny osoby, které mají všechny stupně intoxikace, je možné po druhém stisknutí tlačítka vyobrazit kvartilový graf.

V podsekci Nová nahrávka lze vypočítat všechny parametry zkoumané nahrávky jedním tlačítkem „Výpočet Parametrů“ a nechat je všechny vypsát. Tlačítkem „Ulož Parametry“ dojde k připsání dalšího řádku do souboru data.mat a data.xls, ve kterém jsou zapsány všechny parametry získané ze všech nahrávek. Tlačítko je znepřístupněno, aby nedošlo k nechtěnému přihrání nahrávky. Dá se otevřít pokud 5x stiskneme tlačítko „i“ v příslušné sekci. Dalšími stisky tohoto tlačítka dojde i k zobrazení skrytých tlačítek pro hromadné zpracování dat.

Tlačítko „ALCTest“ provádí porovnání všech získaných parametrů z nové nahrávky s daty z databáze data.mat. Z databáze se vyberou pouze hodnoty střízlivých nahrávek. Maxima a minima se porovnají s hodnotami parametrů získaných z nové nahrávky. V procentech se vypíše, kolik parametrů nové nahrávky je mimo limit.

Soubor data.mat má doposud 3238 řádků. U každé nahrávky se uložilo 40 změřených parametrů: *Energie*, $F_{0\text{mean}}$, $F_{0\text{max}}$, $F_{0\text{min}}$, F_1 , F_2 , F_3 , F_{13} , F_{12} , M_1 , M_2 , M_3 , bw_1 , bw_2 , bw_3 , *jittABS*, *jittREL*, *jittRAP*, *jittPPQ5*, *shimmLOC*, *shimmDB*, *shimAPQ3*, *shimAPQ5*, $C_{1\text{max}}$, $C_{2\text{max}}$, $C_{3\text{max}}$, $C_{4\text{max}}$, $C_{5\text{max}}$, $C_{6\text{max}}$, $C_{7\text{max}}$, $C_{8\text{max}}$, $C_{1\text{mean}}$, $C_{2\text{mean}}$, $C_{3\text{mean}}$, $C_{4\text{mean}}$, $C_{5\text{mean}}$, $C_{6\text{mean}}$, $C_{7\text{mean}}$, $C_{8\text{mean}}$ a počet nul.

Soubor data.mat a celá sekce statistika se dále používají pro zobrazení výsledků práce. Dílčí výsledky budou dále rozebrány v další kapitole.

U každého měření se provádí statistický výpočet. Pro tyto účely byly vybrány jen hodnoty, které obsahují všechny úrovně intoxikace. Podle tabulky 4.1 se výpočet provádí z 19 souborů nahrávek u mužských subjektů a ze 13 souborů nahrávek u ženských subjektů. Pro každou úroveň intoxikace, každý parametr a nahrávku se vypočítá průměr, modus, medián, rozpětí, směrodatná odchylka, rozptyl, šikmost a špičatost. Hodnoty jsou uloženy v souboru datastatic.mat.

The image shows two screenshots of a spreadsheet. The first screenshot, titled 'metadata', displays a table with 19 columns and 30 rows. The columns are labeled with parameters like 'Name', 'Nahravka', 'Pohlavi', 'Vek', 'Hmotnost', 'Stupenintox', 'Energie', 'F0mean', 'F0max', 'F0min', 'F1', 'F2', 'F3', 'F13', 'F12', 'M1', 'M2', 'M3', 'bw1', 'bw2', 'bw3', 'jittABS', 'jittREL', 'jittRAP', 'jittPPQ5', 'shimmLOC', 'shimmDB', 'shimAPQ3', 'shimAPQ5', 'C1max', 'C2max', 'C3max', 'C4max', 'C5max', 'C6max', 'C7max', 'C8max', 'C1mean', 'C2mean', 'C3mean', 'C4mean', 'C5mean', 'C6mean', 'C7mean', 'C8mean', and 'nul'. The second screenshot, titled 'metadastatic', displays a table with 15 columns and 28 rows. The columns are labeled with statistical parameters: 'Name', 'nahravka', 'pohlavi', 'parametr', 'PrumerA', 'PrumerB', 'PrumerC', 'PrumerD', 'PrumerE', 'ModusA', 'ModusB', 'ModusC', 'ModusD', 'ModusE', and 'MedianA'.

Obrázek 4.13 Soubor data.mat a datastatic.mat s uloženými parametry

Sekce Alkoholová kalkulačka

V sekci Alkoholová kalkulačka je možné určit pomocí výpočtu (1.1.) a (1.2.) přibližnou hodnotu alkoholu v krvi po požití alkoholického nápoje. Po vybrání pohlaví testované osoby a zapsáním hmotnosti, množství zkonsumovaného nápoje, doby od konzumace, procenta alkoholu v nápoji a stávající promile v krvi, se může přistoupit k předběžnému výpočtu. Přesnost této kalkulačky nelze příliš určit. U testované osoby hladina intoxikace závisí na více parametrech.

Alkoholova Kalkulacka			
Pohlavi	Mnozství [ml]	Doba od konzumace [h]	
M	500	0.5	Zpocit
Hmotnost [kg]	Volume [%]	Promile v krvi [‰]	
85	7	0	0.39916 ‰

Obrázek 4.14 Alkoholová kalkulačka v aplikaci ALCTest

5. VÝSLEDKY NAMĚŘENÝCH HODNOT

Pro porovnání naměřených hodnot byly vybrány nahrávky, které obsahují všechny míry intoxikace. Z tabulky 4.1 je z mužských nahrávek vybráno 18 subjektů a z ženských nahrávek 13. U subjektu 22 došlo k nevytvoření nahrávky N, proto byl z dalšího měření také vyjmut. Všechny grafy jsou uloženy v příloze.

5.1 Energie nahrávky

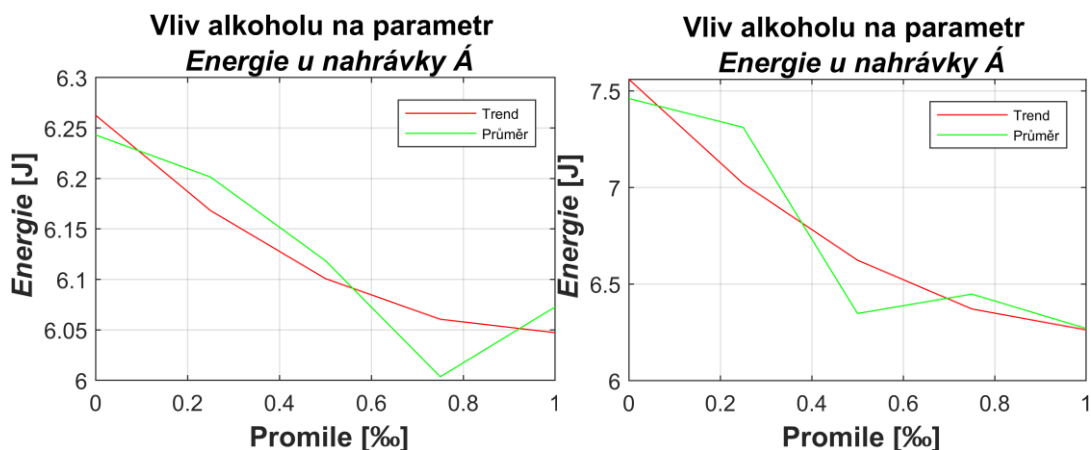
Pro porovnání energie nahrávky jsou vybrány všechny mužské subjekty, všechny hodnoty o hmotnosti, věku a hodnoty při všech pěti úrovních intoxikace. V grafu se zobrazuje průměrná Energie [J].

U samohlásky Á je z grafu 5.1 patrné, že se energie mírně snižuje při nárůstu intoxikace. U většiny pořízených nahrávek se Energie nahrávky s narůstajícím opojením zvyšuje, jak je patrné z tabulky 5.1. U čteného textu se Energie sníží až o 10,4%. K nárůstu energie dochází u samohlásky U.

Při výběru ženských subjektů dochází u samohlásek a u textu k obdobnému efektu. U dlouhých samohlásek Á, Í a Ú se celková energie snižuje. U rozmlouvadel dochází ke snižování energie jen u rozmlouvadla „Nesnese se se sestrou“. Nárůst pozorujeme u samohlásky U o 38.8% a pokles u samohlásky E o 19.3%.

Tabulka 5.1 Změna průměrných hodnot *Energie* v procentech mezi 0‰ a 1‰ u zkoumaných nahrávek

n\p	$\Delta E(M)[\%]$	$\Delta E(Z)[\%]$	n\p	$\Delta E(M)[\%]$	$\Delta E(Z)[\%]$
A	4,27	4,26	Ú	-0,50	-2,57
Á	-2,73	-14,92	T	0,26	-10,58
E	-1,35	-19,33	B	-0,04	5,64
É	4,63	1,40	S	7,60	18,12
I	2,29	4,04	D	2,65	1,89
Í	-3,82	-11,34	N	-4,91	-14,27
O	13,16	38,80	R	-10,43	-16,95
Ó	0,80	8,14	F	6,74	22,24
U	14,62	21,00			



Obrázek 5.1 Vliv alkoholu u samohlásky *Á* na parametr *Energie* [J] u mužů a, a žen b,

5.2 Základní hlasový tón F_0

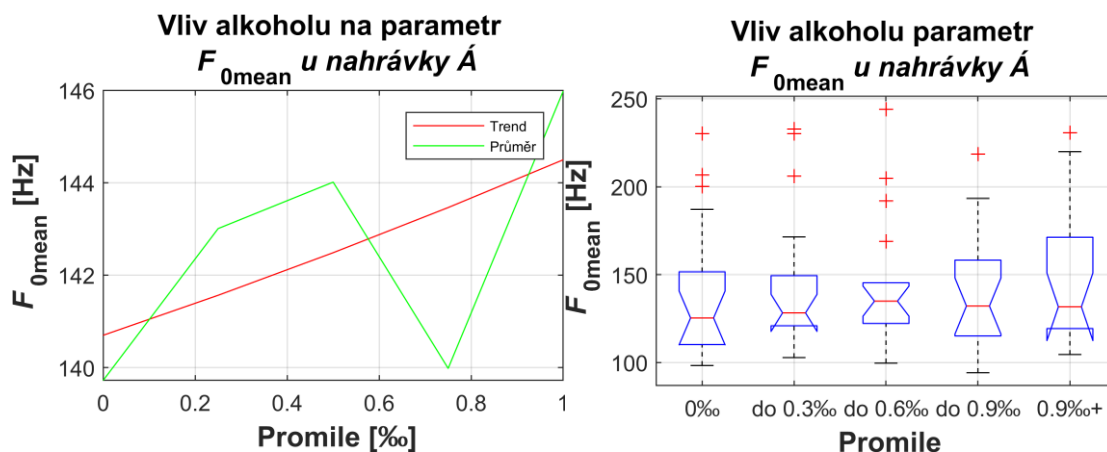
Průměr základního hlasového tónu F_0 u mužů u samohlásky *Á* při intoxikaci roste ze 140 Hz na 146 Hz, u samohlásky *É* roste a pak klesá, stejně u samohlásky *Í* se nejdříve zvýší a pak opět klesá. U samohlásky *Ó* F_0 klesá a u samohlásky *Ú* se zvyšuje ze 140 Hz na 146 Hz stejně jako u samohlásky *Á*. U čteného textu se základní hlasový tón zvyšuje až o 9%. U volné promluvy se základní hlasový tón snižuje. U kratších nahrávek však dochází k různým změnám. V nahrávkách kde je více samohlásek dochází ke zvýšení F_0 a v nahrávkách, kde je více neznělých částí dochází ke snížení F_0 .

Pro zkoumání je výhodnější pozorovat nejnižší měřený základní hlasový tón, kde dochází u samohlásek k výraznějšímu zvýšení F_0 . U čteného textu dochází k průměrnému zvýšení minima až o 70%. U volné promluvy dochází ke snížení až o 26%.

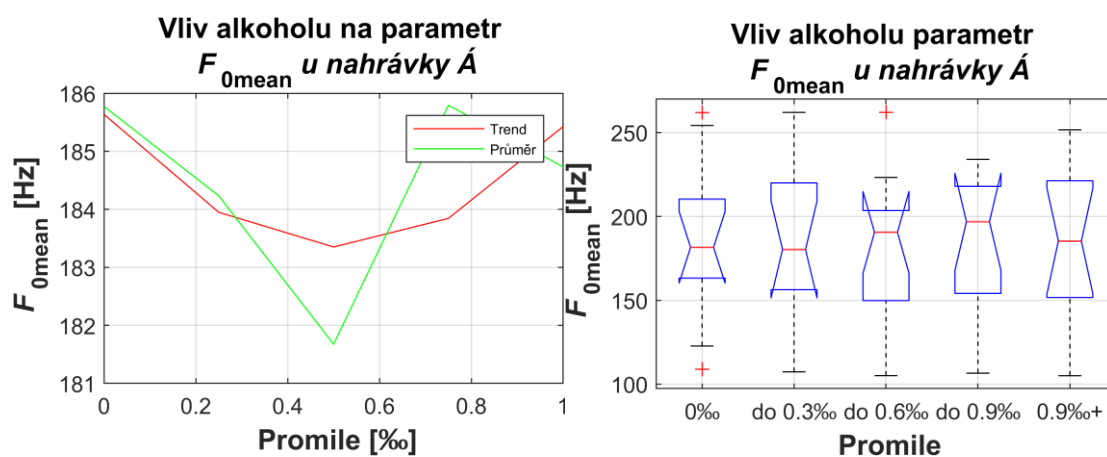
U žen docházíme k obdobným výsledkům. Lehké snížení lze pozorovat u samohlásky *Á*. Větší snížení průměrného základního hlasivkového tónu lze sledovat u samohlásky *Ó*. Stejně jako u mužských subjektů, dochází k rozdílu mezi čteným a volným textem.

Tabulka 5.2 Změny průměrných hodnot F_0 v procentech mezi 0‰ a 1‰ u zkoumaných nahrávek

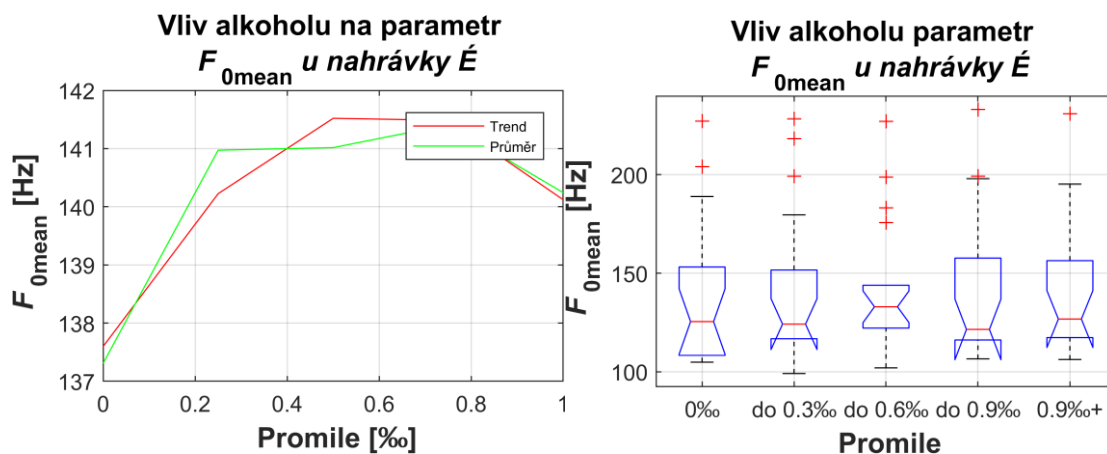
n\p	$\Delta F_{0pr}(M)[\%]$	$\Delta F_{0pr}(Z)[\%]$	n\p	$\Delta F_{0pr}(M)[\%]$	$\Delta F_{0pr}(Z)[\%]$
A	6,25	5,32	Ú	4,65	1,29
Á	4,48	-0,17	T	5,14	5,37
E	15,15	4,30	B	-3,46	7,03
É	2,14	0,01	S	-4,75	-6,64
I	5,84	4,34	D	3,91	15,45
Í	4,89	10,07	N	6,81	13,59
O	6,69	1,31	R	9,08	7,85
Ó	-2,39	-6,37	F	-3,93	-4,34
U	1,43	-4,18			



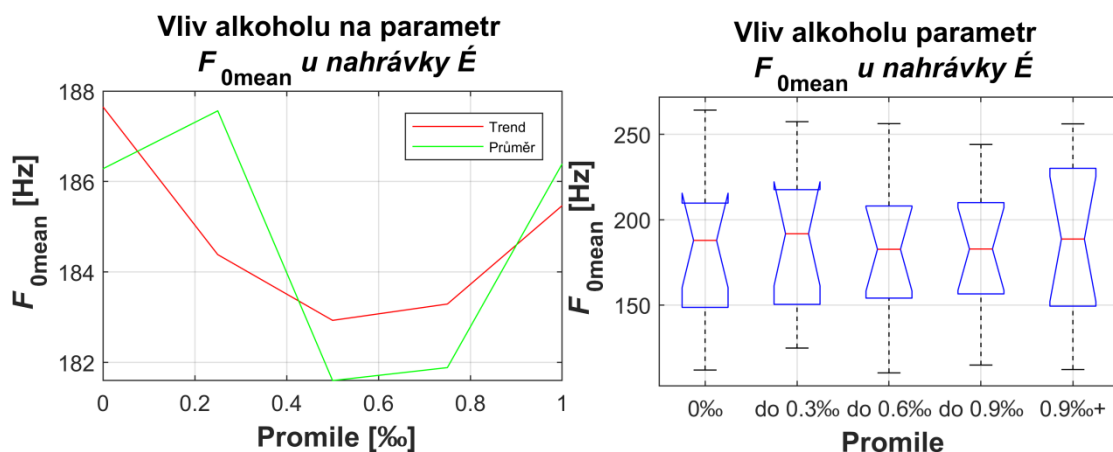
Obrázek 5.2 Vliv alkoholu u samohlásky Á na parametr F_0 u mužů.



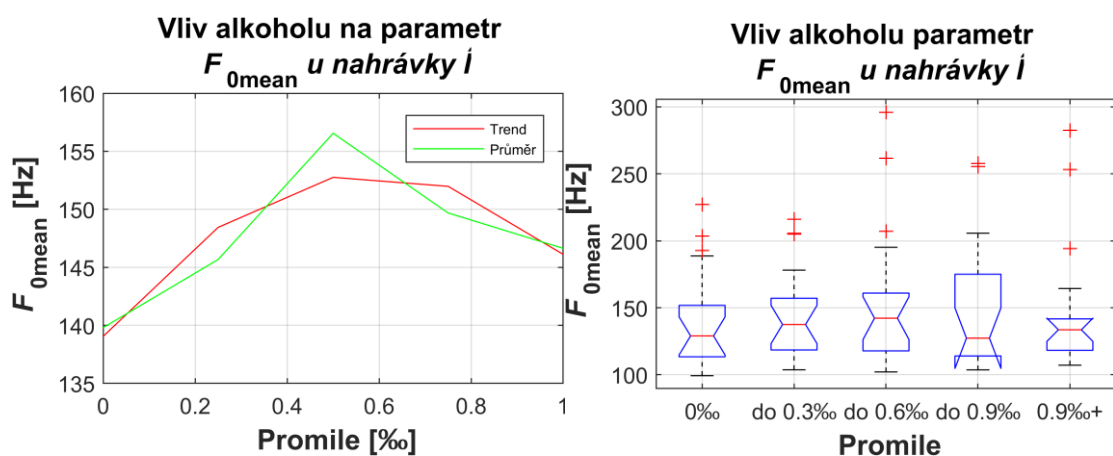
Obrázek 5.3 Vliv alkoholu u samohlásky Á na parametr F_0 u žen.



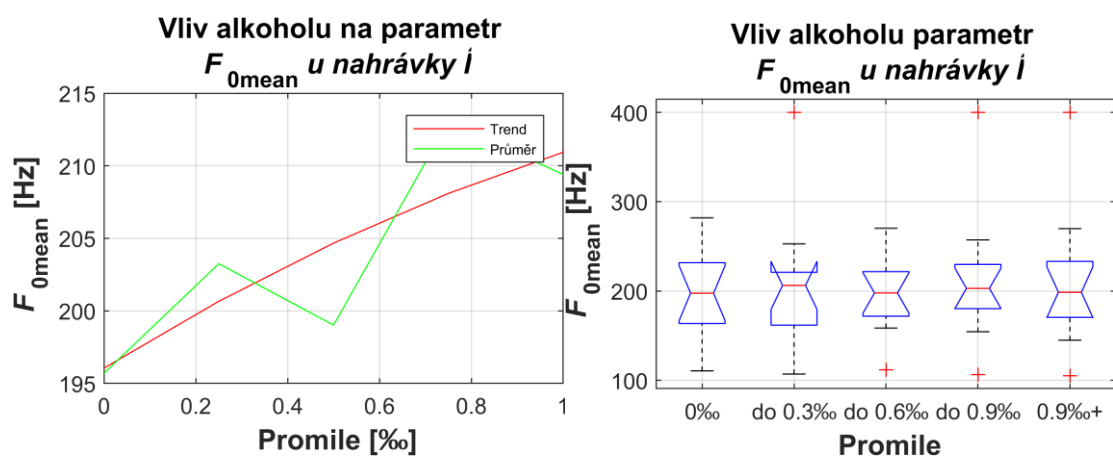
Obrázek 5.4 Vliv alkoholu u samohlásky Ě na parametr F_0 u mužů.



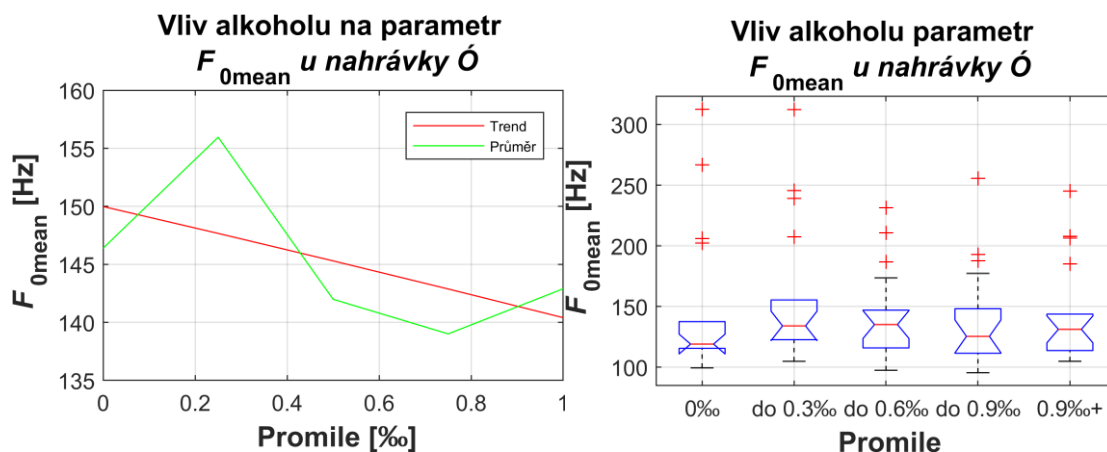
Obrázek 5.5 Vliv alkoholu u samohlásky \acute{E} na parametr F_0 u žen.



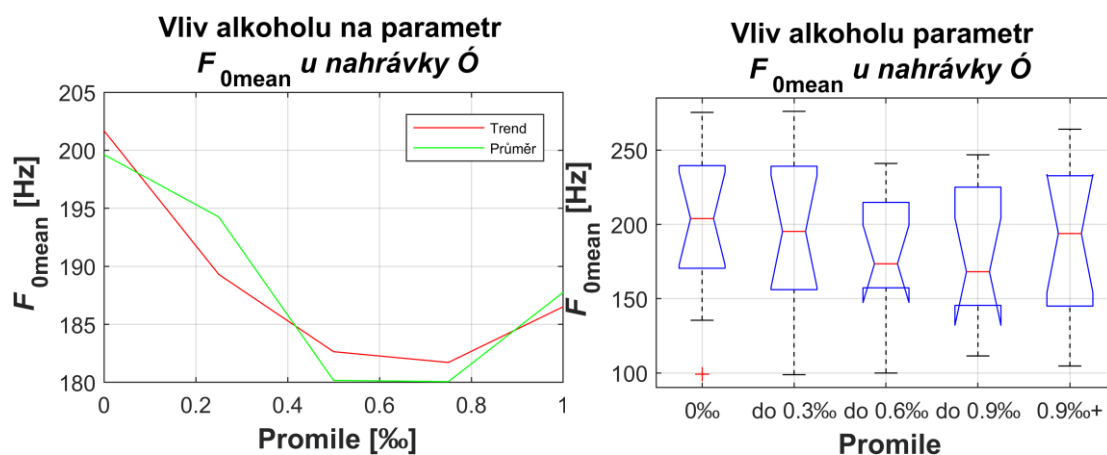
Obrázek 5.6 Vliv alkoholu u samohlásky \acute{I} na parametr F_0 u mužů.



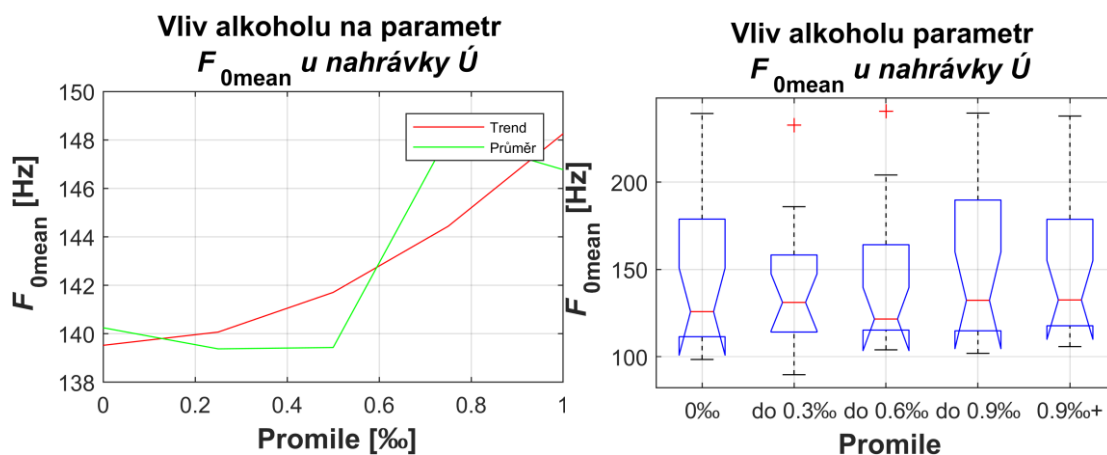
Obrázek 5.7 Vliv alkoholu u samohlásky \acute{I} na parametr F_0 u žen.



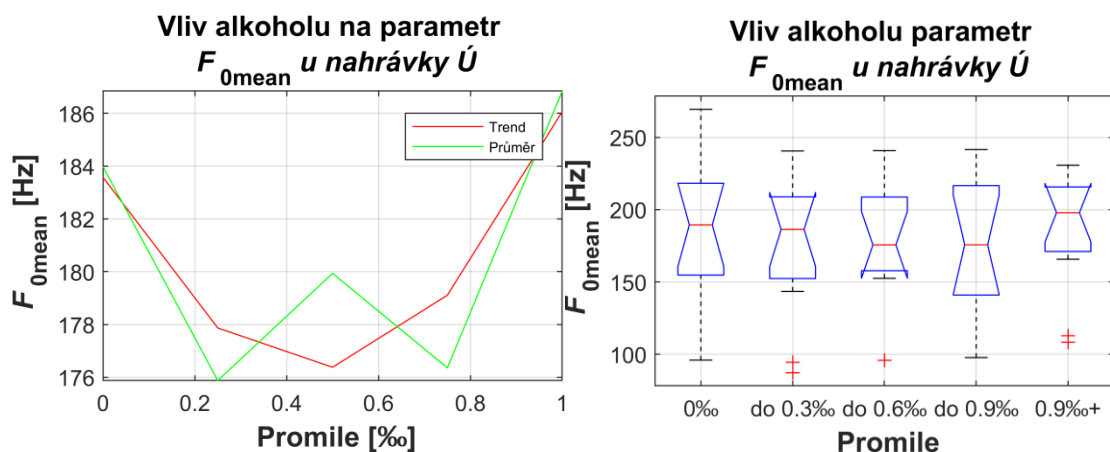
Obrázek 5.8 Vliv alkoholu u samohlásky Ó na parametr F_0 u mužů.



Obrázek 5.9 Vliv alkoholu u samohlásky Ó na parametr F_0 u žen.



Obrázek 5.10 Vliv alkoholu u samohlásky Ú na parametr F_0 u mužů.



Obrázek 5.11 Vliv alkoholu u samohlásky Ú na parametr F_0 u žen.

5.3 Formantové příznaky

U formantů porovnáváme změny frekvenčního posuvu, průměrnou hodnotu prvních tří formantů a jejich změnu. Zkoumáme změny v šířce pásem jednotlivých formantů a hodnoty prvních tří antiformantů.

Z tabulky 5.3 je možné vyčíst pokles průměrných hodnot o 19% u samohlásky Ú a samohlásky Í u formantu F_1 . O 17% klesá průměrná frekvence formantu F_2 u samohlásky Ú. Formant F_3 má největší pokles o 9% u samohlásky Ó a nárůst o 6% u samohlásky A.

U průměrné hodnoty prvních tří formantů dochází u samohlásky Ó k poklesu o 8% a u samohlásky A nárůst o 3%.

Vzdálenost mezi formanty F_1 a F_2 se mění nejvíce u samohlásky I, kde roste v průměru o 11% a u samohlásky E, kde se zmenší o 11%.

U změn průměrných šířek pásem nacházíme velmi vysokou hodnotu u šířky pásma formantu F_2 u samohlásky E až 1169%. Lze ji ale považovat za chybu způsobenou chybným výpočtem šířky pásem pro druhý formant. Větších hodnot nabývá šířka pásma pro 2. formant u samohlásky A a U. K průměrnému zúžení o 58% dochází u samohlásky U u šířky pásma formantu $F1$.

Z tabulky 5.4 zjistíme, že u antiformantů se snížila průměrná frekvence při intoxikaci u samohlásky O o 58% a zvýšila frekvence na antiformantu M_3 u samohlásky Á o 44%.

Z tabulky 5.5 a 5.5, kde jsou vypsány procentuální změny parametrů u nahrávek žen, můžeme vyčíst snížení průměrné hodnoty frekvence o 26% formantu F_2 u nahrávky É. Zvýšení průměrné frekvence o 45% sledujeme u formantu F_2 u samohlásky U.

U šířky pásma formantu F_3 , u samohlásky O, dosáhlo průměrné zvětšení 203%. U šířky pásma formantu F_2 dochází v průměru k největšímu snížení o 69% u samohlásky Í.

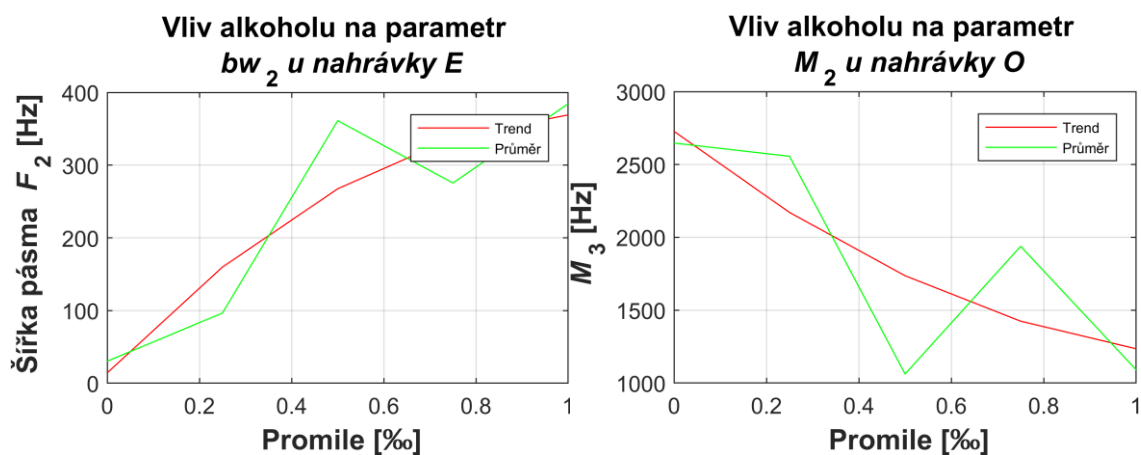
U antiformantů dochází ke zvýšení o 66% u samohlásky E u antiformantu M_3 a snížení u samohlásky I o 63%.

Tabulka 5.3 Změny průměrných hodnot formantů F_1 až F_3 mezi 0 a 1‰ u mužských nahrávek samohlásek

n\p	ΔF_1 [%]	ΔF_2 [%]	ΔF_3 [%]	ΔF_{13} [%]	ΔF_{12} [%]	Δbw_1 [%]	Δbw_2 [%]	Δbw_3 [%]
A	-13,01	4,93	6,66	3,43	-10,96	-43,97	262,67	86,75
Á	0,53	2,34	2,89	2,42	1,63	-6,92	-43,74	-13,42
E	-1,00	4,92	-1,92	0,45	-11,38	38,76	1169,39	-3,59
É	-6,40	0,34	-2,50	-1,96	-10,13	92,49	-44,40	-6,10
I	11,22	1,70	-1,42	0,54	11,70	-44,48	-48,65	5,48
Í	-19,23	-2,84	4,35	0,21	-6,81	-24,91	31,62	-27,56
O	-1,99	-0,84	1,87	0,78	-0,50	36,62	-48,94	103,01
Ó	4,18	-10,73	-9,45	-8,40	0,34	-5,04	-23,46	-21,98
U	-8,30	-1,16	-0,24	-1,16	-7,93	-58,28	363,61	-31,88
Ú	-19,70	-17,05	-0,54	-6,58	-5,60	21,61	-15,00	-47,55

Tabulka 5.4 Změny průměrné hodnoty antiformantů M_1 až M_3 mezi 0 a 1‰ u nahrávek mužských samohlásek

n\p	ΔM_1 [%]	ΔM_2 [%]	ΔM_3 [%]
A	2,95	14,25	22,66
Á	-11,63	16,53	44,36
E	29,51	5,58	17,98
É	30,69	-9,74	-5,87
I	5,20	8,78	7,43
Í	0,90	15,16	9,20
O	14,84	7,72	-58,78
Ó	-26,31	-21,42	6,82
U	19,80	-22,57	-38,25
Ú	-11,37	16,51	43,26



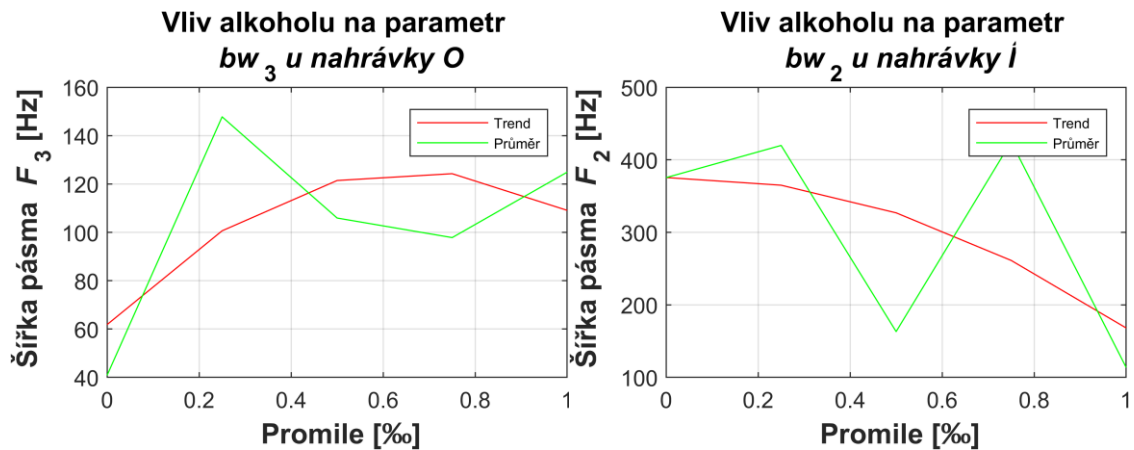
Obrázek 5.12 Nejvýraznější změny formantových příznaků u mužů při intoxikaci.

Tabulka 5.5 Změny průměrných hodnot formantů F_1 až F_3 mezi 0 a 1‰ u nahrávek samohlásek u žen

n\p	ΔF_1 [%]	ΔF_2 [%]	ΔF_3 [%]	ΔF_{13} [%]	ΔF_{12} [%]	Δbw_1 [%]	Δbw_2 [%]	Δbw_3 [%]
A	0,01	-9,83	7,31	0,88	30,36	-27,42	82,30	-60,22
Á	-6,28	5,56	3,04	2,71	-7,75	-37,50	-67,18	-13,38
E	39,40	13,25	-2,74	6,44	29,00	28,71	7,40	-48,35
É	-23,82	-26,71	-4,60	-13,79	-3,15	48,43	-39,76	-24,87
I	18,44	-12,06	-0,53	-3,06	38,28	134,84	-18,86	-42,04
Í	13,12	-17,41	0,60	-4,26	37,37	114,03	-69,48	16,83
O	4,79	-5,99	-2,46	-2,73	1,08	-19,75	-22,18	203,60
Ó	28,79	28,23	-4,33	6,72	4,99	0,01	-49,24	-57,43
U	28,68	45,84	1,25	13,54	-14,96	39,53	-63,50	-48,20
Ú	-25,24	24,04	3,71	5,45	-29,86	119,14	37,34	-29,18

Tabulka 5.6 Změny průměrných hodnot antiformantů M_1 až M_3 mezi 0 a 1‰ nahrávek samohlásek u žen

n\p	ΔM_1 [%]	ΔM_2 [%]	ΔM_3 [%]
A	12,51	19,07	-23,43
Á	-2,19	12,87	0,00
E	-11,21	26,59	66,17
É	14,87	38,82	-54,53
I	18,29	36,52	-25,09
Í	-7,70	-9,32	-31,32
O	34,54	-12,63	-57,43
Ó	0,74	26,89	12,41
U	22,70	12,99	-63,71
Ú	11,81	-0,51	-18,99



Obrázek 5.13 Nejvýraznější změny formantových příznaků u žen při intoxikaci.

5.1 Kepstrální koeficienty

U průměrných hodnot kepstrálních koeficientů ΔC_{pr} u mužů dochází k maximálnímu zvětšení u C_3 u nahrávky A o 6248%. Jak je patrné z grafu 5.14 jde o velmi malé hodnoty. K maximálnímu snížení průměrné hodnoty dochází u kepstrálního koeficientu C_7 o 1162% u nahrávky A.

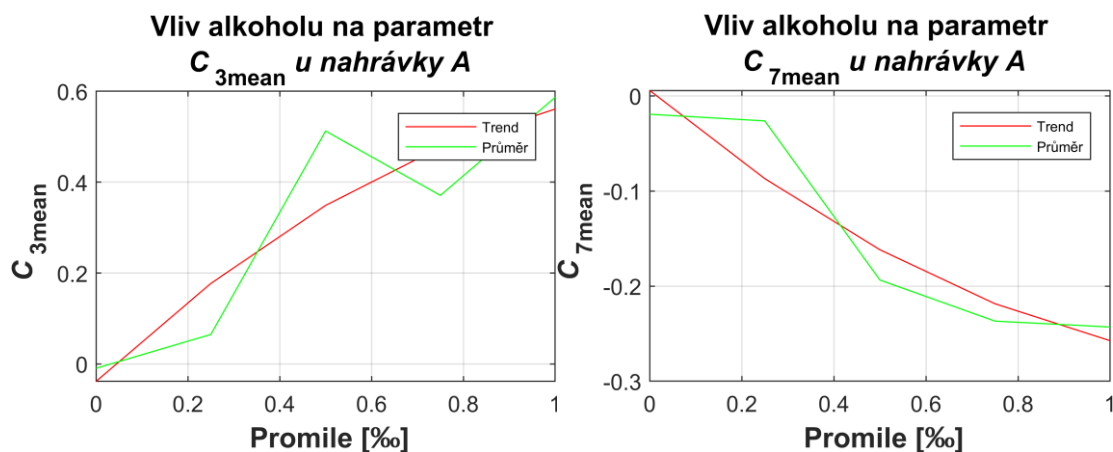
U průměrných maximálních hodnot dochází ke zvýšení hodnot u kepstrálního koeficientu C_3 u nahrávky A o 49% a ke snížení u koeficientu C_5 u nahrávky R o 25%. Maximální hodnoty kepstrálních koeficientů se snižují u čteného textu a některé se zvyšují u volné promluvy. Koeficienty C_4 , C_5 a C_6 se snižují u promluvy jak u maxim, tak v průměrné hodnotě.

U průměrných hodnot kepstrálních koeficientů ΔC_{pr} u žen dochází k maximálnímu zvětšení C_6 u nahrávky B o 2409% a maximálnímu snížení průměrné hodnoty u kepstrálního koeficientu C_8 o 1188% u nahrávky A.

U průměrných maximálních hodnot dochází ke zvýšení hodnot u kepstrálního koeficientu C_3 u nahrávky A o 59% a ke snížení u koeficientu C_3 u nahrávky E o 34%.

Tabulka 5.7 Změny průměrných hodnot kepstrálních koeficientů v procentech mezi 0 a 1‰ u mužských nahrávek

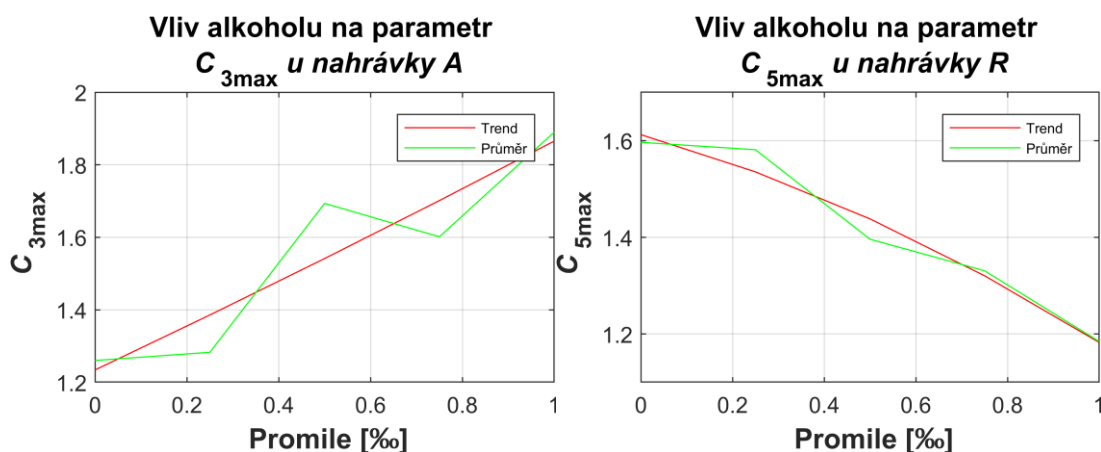
n\p	ΔC_{1pr} [%]	ΔC_{2pr} [%]	ΔC_{3pr} [%]	ΔC_{4pr} [%]	ΔC_{5pr} [%]	ΔC_{6pr} [%]	ΔC_{7pr} [%]	ΔC_{8pr} [%]
A	-4,46	-11,22	6248,00	121,06	109,83	-14,29	-1162,48	66,29
Á	1,71	-12,40	663,54	-49,95	194,99	-8,90	7,95	20,67
E	-1,03	-13,83	-32,53	7,68	69,54	248,76	-45,27	-15,60
É	-0,49	1,21	15,88	9,52	-8,47	-152,09	8,49	-4,67
I	-2,69	-13,44	-17,00	-18,21	26,05	0,42	14,90	34,18
Í	-3,89	-3,33	18,34	-5,70	-23,63	55,65	-4,72	-7,97
O	1,23	-4,64	3,01	-14,13	-11,87	-27,49	-129,36	-53,98
Ó	5,00	6,61	-44,49	-45,76	8,46	12,27	-24,37	-131,94
U	3,55	-3,44	10,82	63,25	-15,68	-50,66	-814,30	-135,01
Ú	0,57	-4,37	38,37	117,71	-92,72	-51,23	77,14	-182,89
T	-3,90	-3,53	11,17	-27,46	767,66	187,16	43,98	91,71
B	0,19	7,72	-34,67	-82,10	33,20	135,89	44,82	104,59
S	3,92	1,18	-23,82	-3,12	-45,24	-363,55	6,90	-20,91
D	-2,63	2,95	112,27	50,48	-5,82	-56,56	-4,37	5,14
N	-0,69	-10,65	4,92	43,93	-49,78	-23,80	92,37	26,36
R	-6,26	-7,57	57,48	-4,51	-23,53	102,07	0,92	438,20
F	1,09	5,40	34,33	-13,98	-180,40	-89,10	15,66	-28,20



Obrázek 5.14 Nejvýraznější průměrné změny keprálních koeficientů u mužů při intoxikaci.

Tabulka 5.8 Změny maximálních průměrných hodnot keprálních koeficientů u mužských nahrávek

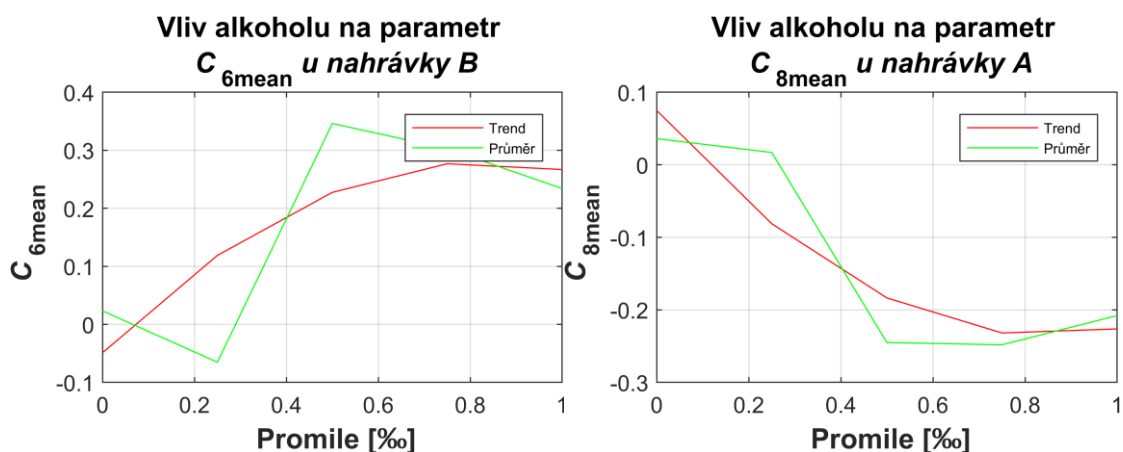
n\p	$\Delta C_{1\text{max}}$ [%]	$\Delta C_{2\text{max}}$ [%]	$\Delta C_{3\text{max}}$ [%]	$\Delta C_{4\text{max}}$ [%]	$\Delta C_{5\text{max}}$ [%]	$\Delta C_{6\text{max}}$ [%]	$\Delta C_{7\text{max}}$ [%]	$\Delta C_{8\text{max}}$ [%]
A	-3,13	3,02	49,94	16,05	3,77	1,20	-5,84	9,92
Á	1,95	-9,50	-2,09	-11,34	-9,03	-11,15	-10,84	-9,52
E	6,00	-7,21	10,89	-1,70	1,93	2,02	0,20	-1,95
É	-1,57	4,46	10,16	3,02	-5,64	6,17	-4,77	-0,33
I	-0,18	-8,72	-2,41	-5,79	-0,57	-9,13	16,81	-2,33
Í	-6,36	-7,31	-4,02	-6,82	-10,44	-3,00	-1,76	15,58
O	3,74	7,84	29,40	18,73	21,27	11,78	9,12	10,17
Ó	7,13	2,49	3,88	-1,83	9,89	-0,51	12,98	5,16
U	4,31	4,13	22,88	34,86	16,44	27,03	25,79	37,63
Ú	2,39	-4,34	5,60	-7,52	-20,75	-5,80	8,32	-11,29
T	-0,45	-8,40	-10,64	-6,40	-7,52	-8,92	-12,64	-2,03
B	3,80	10,61	-9,14	4,37	-3,74	16,52	11,56	15,79
S	3,39	12,26	1,16	13,18	3,75	8,76	-1,79	17,30
D	-2,63	1,96	13,86	4,14	-8,65	-11,36	-6,66	-4,87
N	0,70	-7,13	-7,66	-1,92	-1,54	-2,22	3,71	-4,45
R	-5,84	-14,43	-3,91	-14,80	-25,80	-9,17	-17,49	-12,00
F	0,25	7,22	7,42	-5,25	-6,22	-2,41	4,85	4,70



Obrázek 5.15 Nejvýraznější průměrné změny maximálních koeficientů u mužů při intoxikaci.

Tabulka 5.9 Změny průměrných hodnot koeficientů u náhrávek žen.

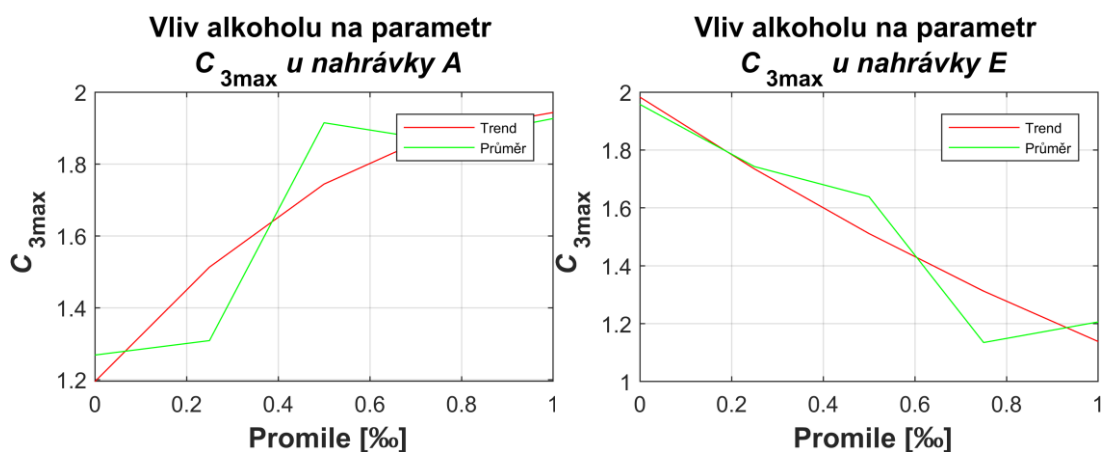
n\p	ΔC_{1pr} [%]	ΔC_{2pr} [%]	ΔC_{3pr} [%]	ΔC_{4pr} [%]	ΔC_{5pr} [%]	ΔC_{6pr} [%]	ΔC_{7pr} [%]	ΔC_{8pr} [%]
A	-0,43	-10,58	1110,98	281,85	-311,15	9,97	-179,79	-1188,51
Á	1,18	-20,08	283,49	34,73	-156,01	21,56	11,61	64,23
E	-4,63	-39,45	-97,83	247,29	65,02	904,96	252,75	179,07
É	-1,03	31,74	63,07	-28,40	-32,36	24,79	-377,96	-1036,60
I	4,06	-9,39	-33,05	-131,35	-23,05	1881,03	-21,95	43,96
Í	-4,05	8,18	-20,94	-162,95	22,97	263,19	31,58	86,20
O	8,82	10,01	-16,87	-199,21	46,22	-61,57	-345,90	-120,87
Ó	10,76	-2,73	-165,80	-43,57	-221,15	38,17	-173,85	72,27
U	7,37	-4,02	4,76	102,43	-115,63	-148,22	-68,37	-198,03
Ú	-1,63	-1,16	905,79	86,99	-156,87	-124,11	63,16	-139,46
T	-5,95	0,64	13,22	-35,89	70,61	199,41	110,70	913,71
B	7,10	31,46	-41,03	-290,89	-14,51	2409,41	47,01	73,22
S	10,34	12,26	-30,14	5,78	28,17	-197,71	-174,86	-33,17
D	2,74	2,81	30,29	15,47	-8,68	-64,45	-7,78	-7,32
N	-6,74	-18,92	18,72	1386,61	18,24	399,04	141,64	259,77
R	-8,42	2,26	74,46	6,04	-14,09	170,14	31,85	282,42
F	-1,61	20,67	527,96	135,57	-123,68	-93,81	10,98	-77,21



Obrázek 5.16 Nejvýraznější průměrné změny keprálních koeficientů u žen při intoxikaci.

Tabulka 5.10 Změny maximálních průměrných hodnot keprálních koeficientů u nahrávek žen

n/p	$\Delta C_{1\text{max}}$ [%]	$\Delta C_{2\text{max}}$ [%]	$\Delta C_{3\text{max}}$ [%]	$\Delta C_{4\text{max}}$ [%]	$\Delta C_{5\text{max}}$ [%]	$\Delta C_{6\text{max}}$ [%]	$\Delta C_{7\text{max}}$ [%]	$\Delta C_{8\text{max}}$ [%]
A	-3,10	-6,25	59,19	7,36	-6,76	1,72	-12,54	-10,14
Á	-0,56	-14,61	-0,24	-11,70	-21,89	-17,12	-6,56	-9,04
E	12,30	-35,42	-34,28	-19,71	-15,03	23,64	3,59	4,29
É	-4,72	13,42	3,37	-9,95	-10,89	-7,02	-7,99	-6,61
I	0,87	-5,42	-9,45	-23,07	-6,96	7,49	8,27	10,96
Í	-11,11	-4,56	-29,91	-27,23	-28,22	-1,56	-0,67	-20,42
O	6,63	54,12	50,47	40,47	44,73	33,00	15,82	21,35
Ó	12,35	-10,09	-34,06	-17,09	2,46	8,17	1,55	24,70
U	8,99	8,25	48,11	46,80	32,34	4,67	39,31	45,33
Ú	0,80	-2,54	37,20	-5,65	-14,36	-14,64	21,85	-14,73
T	-5,50	-13,58	1,87	-5,10	-8,66	9,03	2,09	-15,66
B	8,37	37,86	-7,35	4,81	26,62	34,11	35,04	16,01
S	14,49	28,87	2,38	32,45	25,65	-4,44	15,57	13,42
D	-0,70	3,57	2,77	-7,80	-6,01	-5,88	-1,08	3,10
N	3,68	-21,45	-16,67	-2,42	-14,40	-12,76	21,63	6,07
R	-9,64	-17,02	-5,13	-28,33	-27,71	-11,04	-16,26	-22,84
F	-6,43	28,68	26,15	36,98	12,44	1,74	19,05	19,11



Obrázek 5.17 Nejvýraznější průměrné změny maxim keprálních koeficientů u žen při intoxikaci.

5.2 Jitter a Shimmer

Hodnoty *Jitteru* a *Shimmeru* nejsou dostatečně přesně spočítané. Hodnoty oproti programu Praat jsou značně odlišné. Přesto byly změřené hodnoty započítány do statistiky a porovnány jejich změny při střízlivém a intoxikovaném stavu. Budeme uvažovat pouze samohlásky, neboť tam k chybě mohlo dojít nejméně, jelikož jde převážně po celou dobu nahrávky o znělou část, což může být jeden z důvodů odlišného výpočtu.

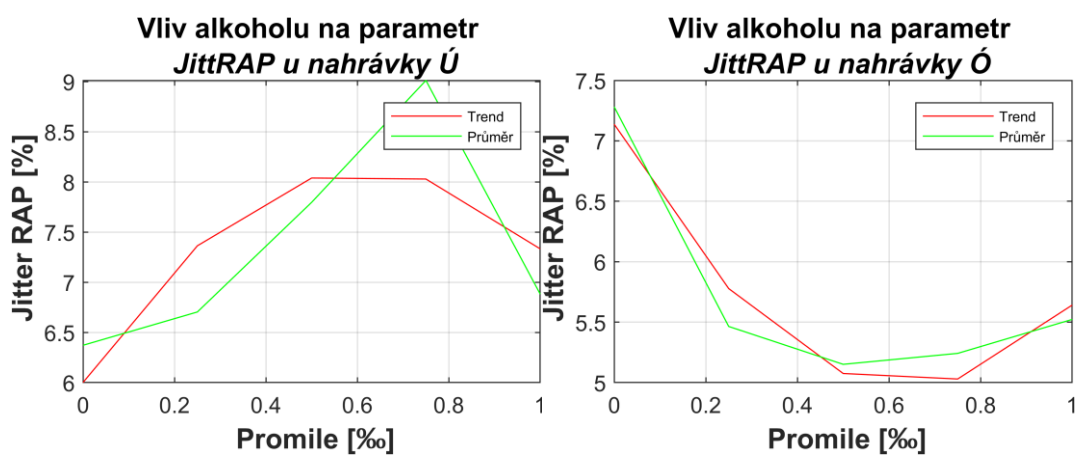
Pro ilustraci jsou v tabulkách 5.11 až 5.14 uvedeny všechny pořízené nahrávky a jejich změny v procentech mezi 0‰ a okolo 1‰.

U mužů se u samohlásek nejvíce zvyšuje hodnota *JittPPQ5* o hodnotu 8,5% u nahrávky A a nejvíce se snižuje *JitterPPQ5* o hodnotu 26% u nahrávky Ó. Hodnota *ShimmerAPQ5* se zvyšuje u samohlásky Í o 213% a snižuje u samohlásky U o 89%.

U žen se u samohlásek nejvíce zvyšuje *JitterREL* o hodnotu 31% u nahrávky E a nejvíce se snižuje *JitterABS* o hodnotu 31% u nahrávky Ó. Hodnota *ShimmerAPQ5* se zvyšuje u samohlásky É o 1404% a snižuje u samohlásky U o 91%.

Tabulka 5.11 Změny průměrných hodnot jitterů v procentech mezi 0‰ a 1‰ u mužských nahrávek

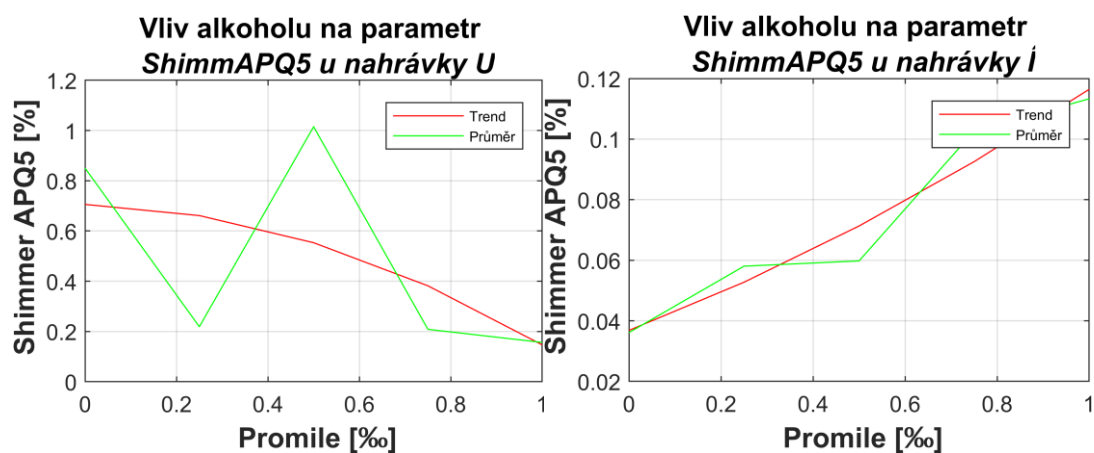
n/p	$\Delta jittABS$ [%]	$\Delta jittREL$ [%]	$\Delta jittRAP$ [%]	$\Delta jittPPQ5$ [%]
A	-4,30	-0,69	3,86	8,51
Á	-4,08	-1,75	-2,84	-4,42
E	-4,72	0,17	0,28	4,34
É	-11,11	-10,37	-9,15	-13,89
I	-7,78	-6,45	-6,51	-2,76
Í	7,41	7,69	7,44	2,10
O	-9,56	-11,78	-8,10	-5,54
Ó	-21,96	-23,15	-24,16	-26,02
U	-21,80	-21,80	-22,87	-17,15
Ú	4,33	7,09	8,01	6,20
T	-8,29	-8,12	-10,60	-10,88
B	-24,64	-21,23	-21,74	-21,66
S	-9,79	-6,92	-5,90	-3,03
D	-9,97	-10,90	-9,72	-12,02
N	3,76	1,01	-0,96	0,92
R	-7,36	-8,07	-10,71	-12,78
F	-12,83	-10,49	-8,73	-8,78



Obrázek 5.18 Nejvýraznější průměrné změny Jitteru u mužů při intoxikaci.

Tabulka 5.12 Změny průměrných hodnot shimmerů v procentech mezi 0‰ a 1‰ u mužských nahrávek

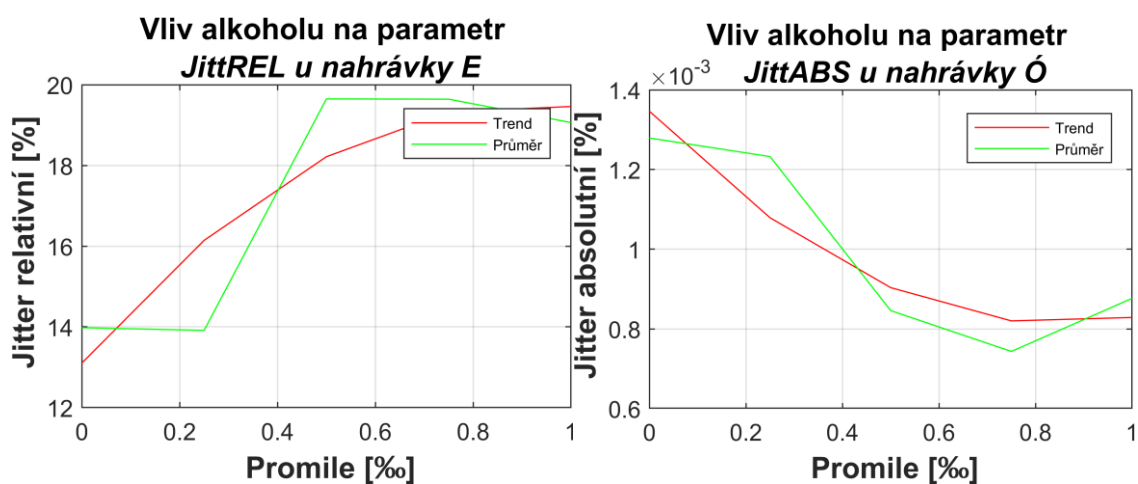
n/p	$\Delta shimmerLOC$ [%]	$\Delta shimmerdB$ [%]	$\Delta shimmerAPQ3$ [%]	$\Delta shimmerAPQ5$ [%]
A	-3,20	-3,46	-1,06	15,44
Á	4,44	0,61	-0,37	-19,05
E	11,30	3,58	4,35	-33,43
É	-17,25	-14,73	-17,40	34,01
I	-24,98	-25,90	-30,18	-52,86
Í	-15,21	-11,97	-19,98	213,76
O	-22,91	-20,09	-32,88	-31,09
Ó	-8,53	-19,92	-5,84	21,21
U	-13,95	-21,12	-15,93	-81,59
Ú	26,01	28,59	24,58	-14,16
T	-7,88	-2,21	-6,83	284,13
B	-7,46	-5,39	-10,28	-59,25
S	-0,72	-0,53	-1,34	-77,61
D	-16,80	-19,14	-17,77	-37,34
N	12,06	3,35	19,39	161,01
R	-11,13	-2,01	-9,71	4684,95
F	-6,27	-4,25	-8,70	-89,69



Obrázek 5.19 Nejvýraznější průměrné změny Shimmeru u mužů při intoxikaci.

Tabulka 5.13 Změny průměrných hodnot jitterů v procentech mezi 0‰ a 1‰ u ženských nahrávek

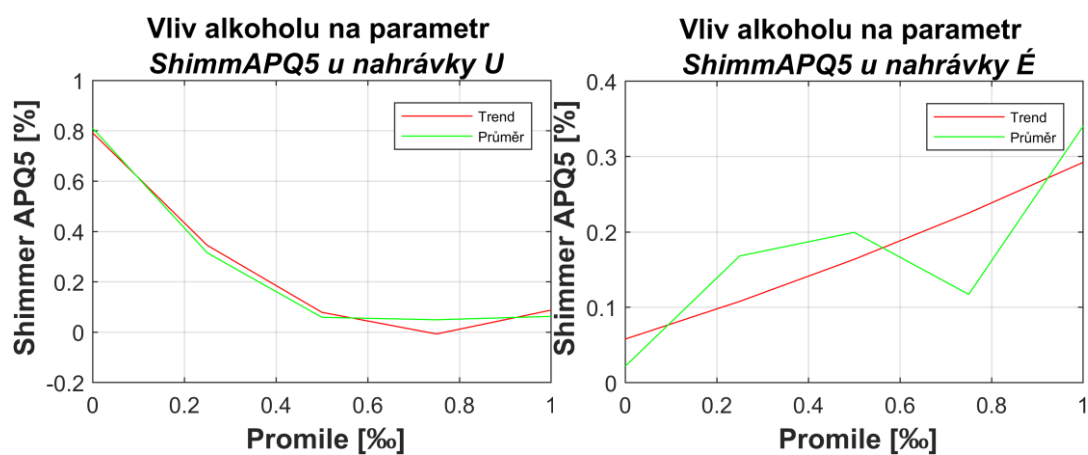
n/p	$\Delta jittABS$ [%]	$\Delta jittREL$ [%]	$\Delta jittRAP$ [%]	$\Delta jittPPQ5$ [%]
A	-6,81	-7,61	-5,09	-9,93
Á	15,10	12,82	10,97	11,34
E	23,46	31,54	31,34	19,54
É	0,49	-0,01	-0,97	0,53
I	-27,03	-23,86	-22,99	-21,27
Í	26,30	11,04	6,56	4,52
O	-3,42	-0,73	6,19	-0,25
Ó	-31,17	-29,93	-25,66	-28,27
U	-15,90	-5,46	1,38	-1,18
Ú	-10,52	-14,12	-17,49	-14,73
T	16,21	8,77	4,98	2,98
B	12,92	12,34	19,05	23,23
S	-10,74	-9,84	-8,85	-3,86
D	3,71	-2,17	-2,94	0,67
N	20,06	20,23	18,00	20,27
R	4,99	5,63	0,98	1,98
F	24,39	21,11	27,22	34,74



Obrázek 5.20 Nejvýraznější průměrné změny *Jitteru* u žen při intoxikaci.

Tabulka 5.14 Změny průměrných hodnot shimmerů v procentech mezi 0‰ a 1‰ u ženských nahrávek

n/p	$\Delta shimmerLOC$ [%]	$\Delta shimmerdB$ [%]	$\Delta shimmerAPQ3$ [%]	$\Delta shimmerAPQ5$ [%]
A	-28,47	-13,29	-28,77	-9,31
Á	6,32	1,01	4,52	215,02
E	9,37	22,72	5,75	102,85
É	-22,65	-17,42	-26,29	1404,16
I	-22,29	-20,66	-16,29	-71,89
Í	30,40	45,86	40,24	651,95
O	-11,16	-18,71	-14,13	-68,72
Ó	-15,41	-33,03	-4,94	-71,33
U	-23,36	-25,44	-31,81	-91,65
Ú	-39,65	-42,87	-44,76	209,28
T	17,36	17,28	19,60	683,85
B	3,23	-6,03	-2,73	-95,40
S	1,05	-2,10	2,38	65,21
D	-21,62	-27,95	-23,80	-90,42
N	11,96	29,00	16,39	96,13
R	-3,08	20,92	-5,67	2502,76
F	34,84	12,39	33,11	-50,99



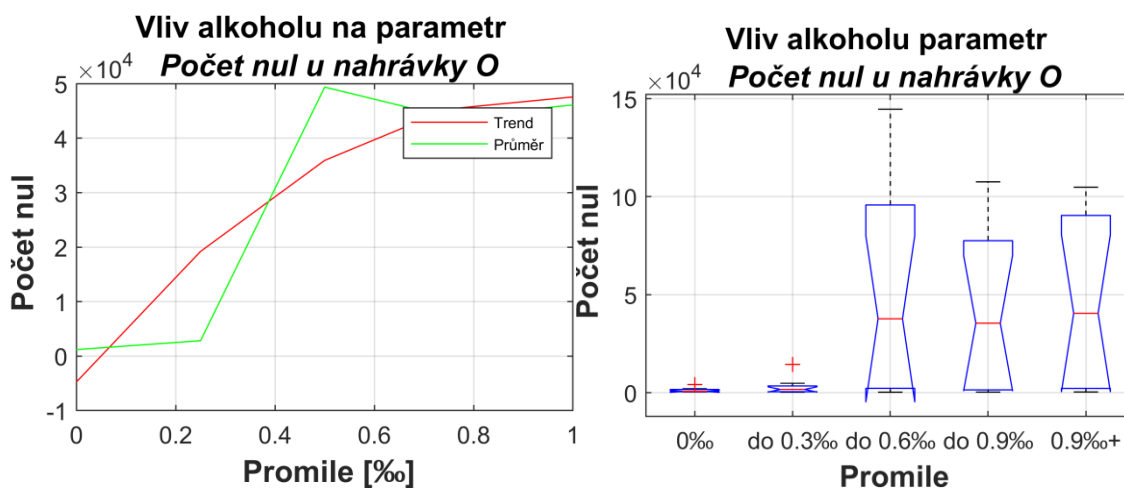
Obrázek 5.21 Nejvýraznější průměrné změny Shimmeru u žen při intoxikaci.

5.3 Počet průchodů nulou

Počet nul značně závisí na znělé oblasti nahrávky a samotné délce nahrávky. Při vyšších frekvencích dochází k většímu počtu průchodu nulou. Podle tabulky 5.15 je patrné, že jak u mužských osob, tak u ženských osob dochází až k vysokému navýšení počtu průchodů nulou u samohlásek O a U, což může být způsobeno zvýšením frekvencí v hlase, ale i prodlužováním nahrávek při intoxikaci. Snížení počtu průchodů nulou v nahrávce dochází u samohlásek A, Á, I, Í a Ú. Předpokládáme, že došlo k rozdílům v časových úsecích nahrávek u řeči střízlivé a intoxikované.

Tabulka 5.15 Změny průměrných hodnot počtu průchodů nulou v procentech mezi 0‰ a 1‰

n\p	Δ pocetnul(M) [%]	Δ pocetnul(Z) [%]	n\p	Δ pocetnul(M) [%]	Δ pocetnul(Z) [%]
A	-14,7	-27,83	Ú	-15,85	-14,84
Á	-82,42	-74,06	T	-26,41	-58,51
E	65,76	-57,33	B	202,81	69,1
É	101,78	-10,78	S	157,82	331,37
I	-8,54	-4,9	D	-9,9	4,1
Í	-82,88	-90,76	N	88,05	-52,57
O	3884,01	3563,03	R	-18,49	-63,31
Ó	20,47	135,22	F	17,28	42,75
U	1106,86	1902,61			



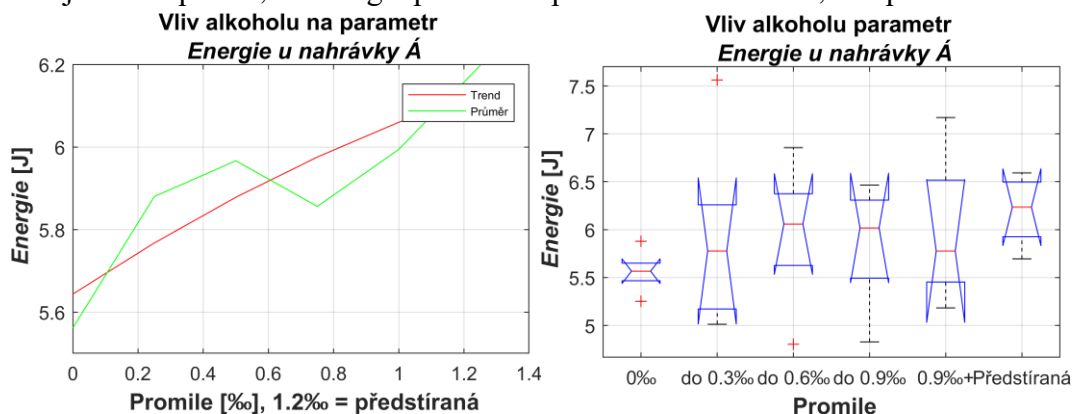
Obrázek 5.22 Nejvýraznější změna průměrného počtu průchodů nulou u samohlásky O u žen v závislosti na intoxikaci

5.4 Předstíraná opilá řeč

Profesionálně předstíraná opilá řeč byla vytvořena za pomoci 6 mužských subjektů s hereckým vzděláním. V tabulce 4.1 jsou to subjekty číslo 1, 6, 9, 10, 27 a 36. Předstíraná opilá řeč byla vytvořena v zápětí po řeči střízlivé.

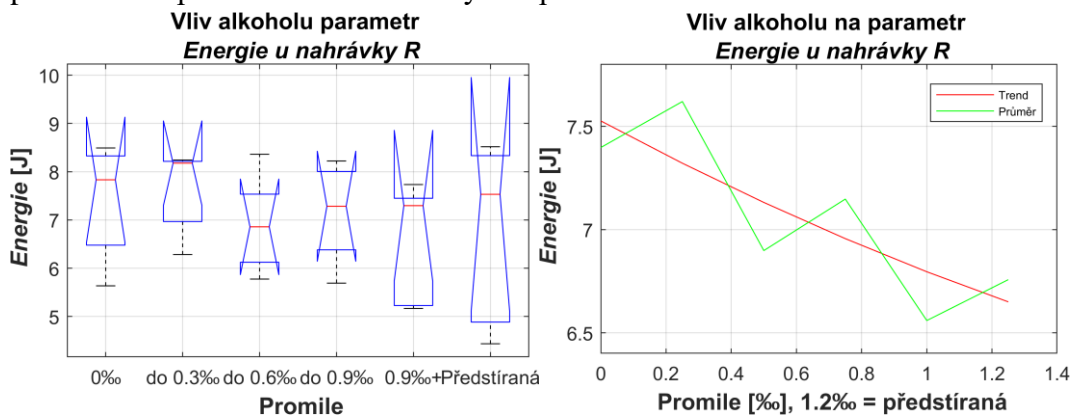
Energie

U předstírané intoxikace se parametr energie podle grafů 5.23 zvyšuje u testované samohlásky *Á*. Rozpětí hodnot u předstírané opilé řeči je nejbližší podobné řeči střízlivé. Energie má nejbližší k intoxikaci do 0.6‰. Směrodatné odchylky nepřesahují jak u střízlivé, tak u předstírané opilé řeči hodnotu 1. Při intoxikaci je směrodatná odchylka více jak 1. Je patrné, že energie při hrané opilé řeči vzrůstá více, než při intoxikaci.

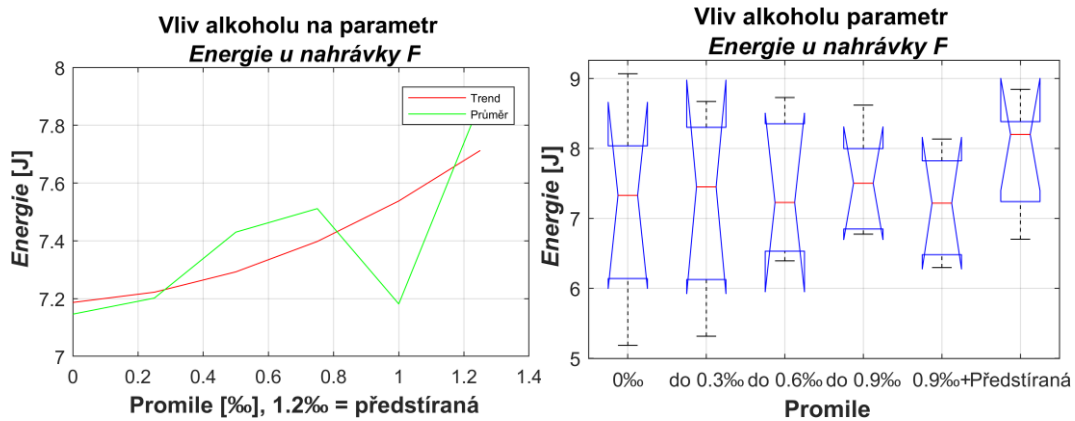


Obrázek 5.23 Změny u parametru *Energie* při předstírané intoxikaci u samohlásky *Á*

Pro další srovnání je vybrána čtená a volná řeč. Z grafů 5.24 a 5.25 se dá odvodit, že energie předstírané řeči vzrůstá a její rozptyl se zvětší z hodnoty 1 u střízlivé řeči na hodnotu 3. Je zjevné, že předstíraná opilá čtená řeč podněcuje k větším emočním projevům a tím dochází k velkému rozptylu a rozpětí hodnot energie. Dochází ale i k velkému snížení energie, což může být způsobeno předstíráním unylosti řeči. U volné předstírané opilé řeči dochází ke zvýšení průměru hodnot.



Obrázek 5.24 Změny u parametru *Energie* při předstírané intoxikaci u čtené řeči.

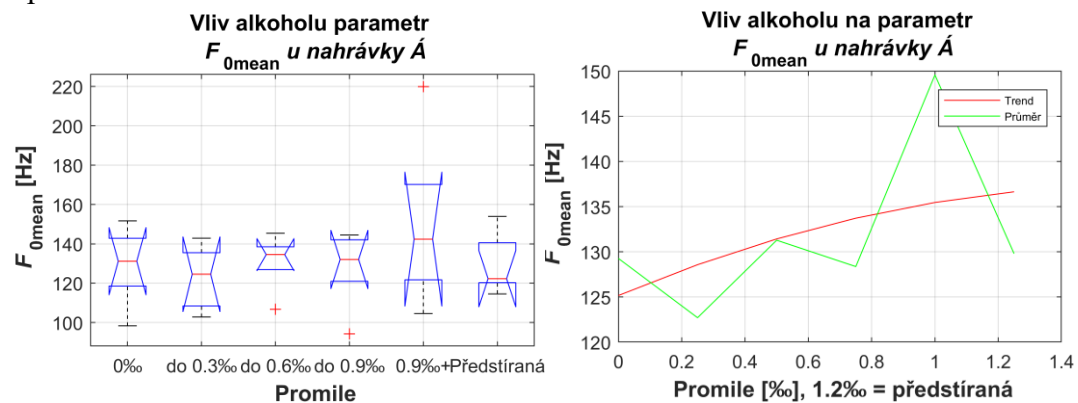


Obrázek 5.25 Změny u parametru $Energie$ při předstírané intoxikaci u volné řeči.

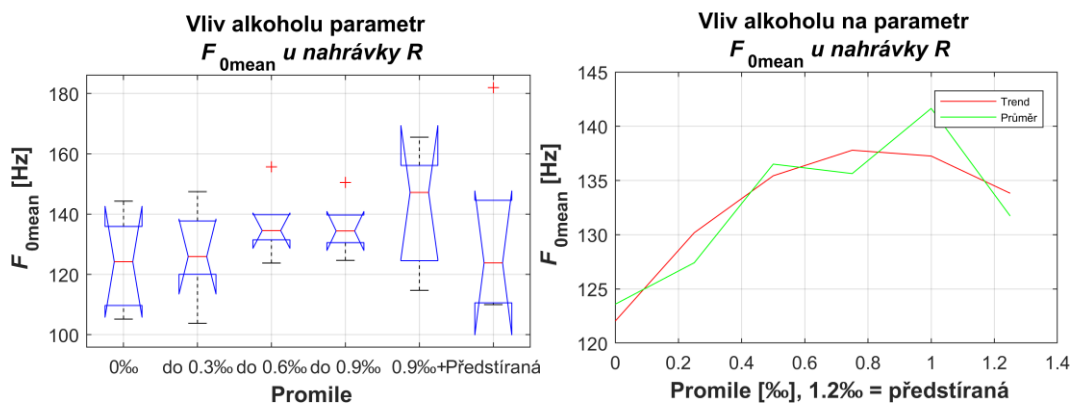
Základní hlasový tón F_0

Základní hlasový tón u předstírané opilé řeči u samohlásky Á je podobný jako u řeči střízlivé. Jeho průměr je zvýšen o 1Hz a je snížen rozptyl hodnot. U intoxikované interpretaci samohlásky Á dochází k postupnému zvýšení průměrné hodnoty, jak je patrné u grafu na obrázku 5.26.

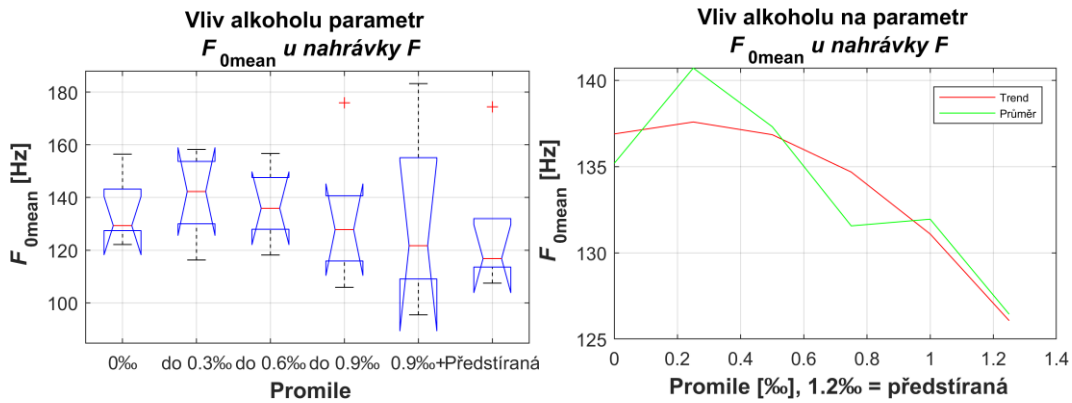
U čteného a volného textu dochází k opačným změnám. U čteného textu dochází ke zvýšení o 8Hz průměrné hodnoty. Median zůstává na stejné hodnotě 124Hz. Při intoxikaci se median zvětšuje na hodnotu 147. U volného textu se základní hlasový tón snižuje při hrané opilé řeči. Snížení se dá přisoudit uvolnění mluvčího při utváření hrané opilosti bez nutnosti se soustředit na text.



Obrázek 5.26 Změny u parametru F_0 při předstírané intoxikaci u samohlásky Á .



Obrázek 5.27 Změny u parametru F_0 při předstírané opilé čtené řeči.

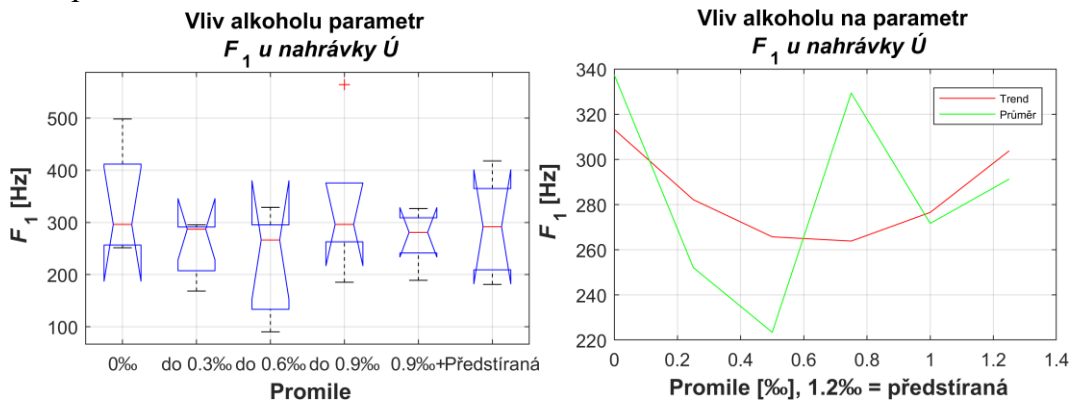


Obrázek 5.28 Změny u parametru F_0 při předstírané opilé volné řeči.

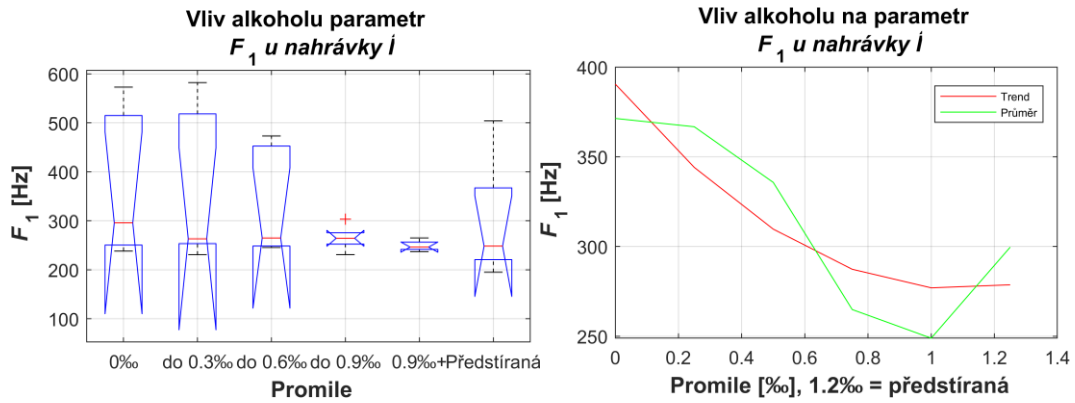
Formantové příznaky

U pozorovaných změn při intoxikaci podle tabulek 5.3 a 5.4. dochází ke změnám u samohlásky Ú a Í u formantu F_1 , u samohlásky Ú formantu F_2 a větší změně u šířky pásma formantu F_2 u samohlásky E.

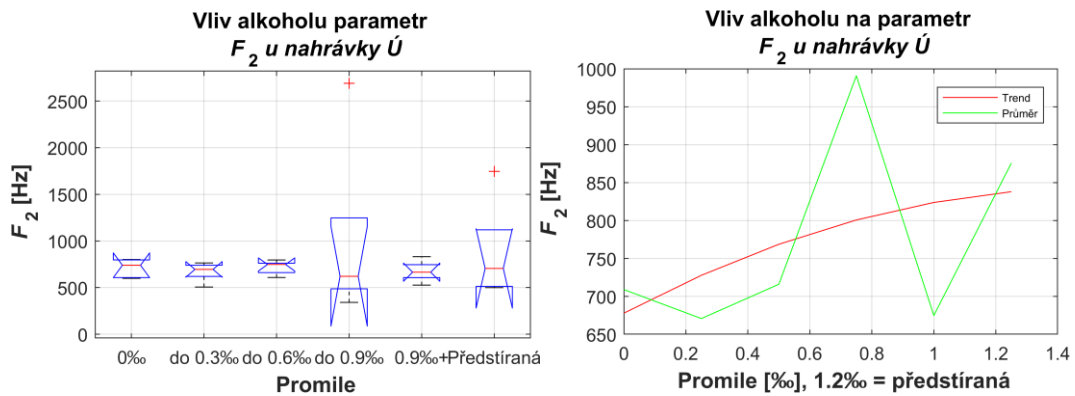
Při pozorování předstírané opilé řeči u samohlásky Ú dochází ke snížení průměrné frekvence. U samohlásky Í u formantu F_1 dochází při intoxikaci ke snížení a zúžení rozptylu hodnot. U předstírané opilé řeči jsou hodnoty mírně sníženy. U formantu F_2 záleží na nastavení dutiny ústní při vytváření samohlásek. Vyšší frekvence vznikají při menších rozměrech dutiny ústní. U samohlásky Ú se lze domnívat, že je dutina ústní více přivřena.



Obrázek 5.29 Změny u parametru F_1 při předstírané intoxikaci u samohlásky Ú.



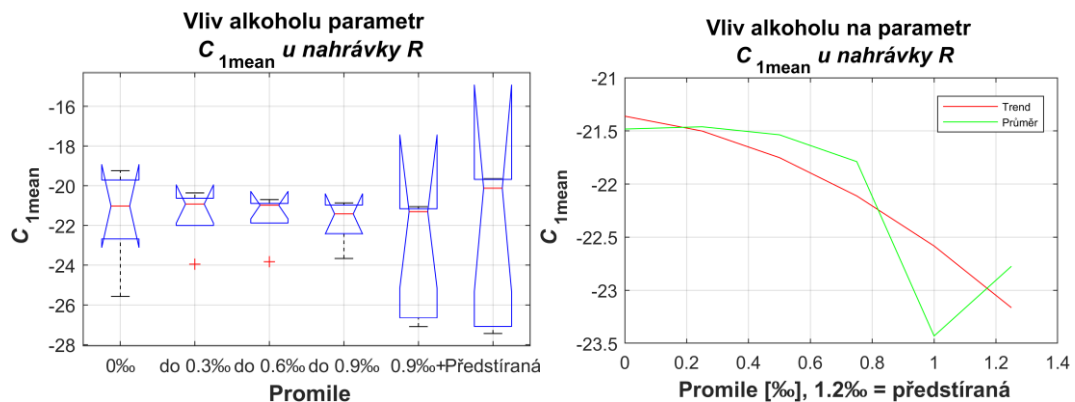
Obrázek 5.30 Změny u parametru F_1 při předstírané intoxikaci u samohlásky \dot{I} .



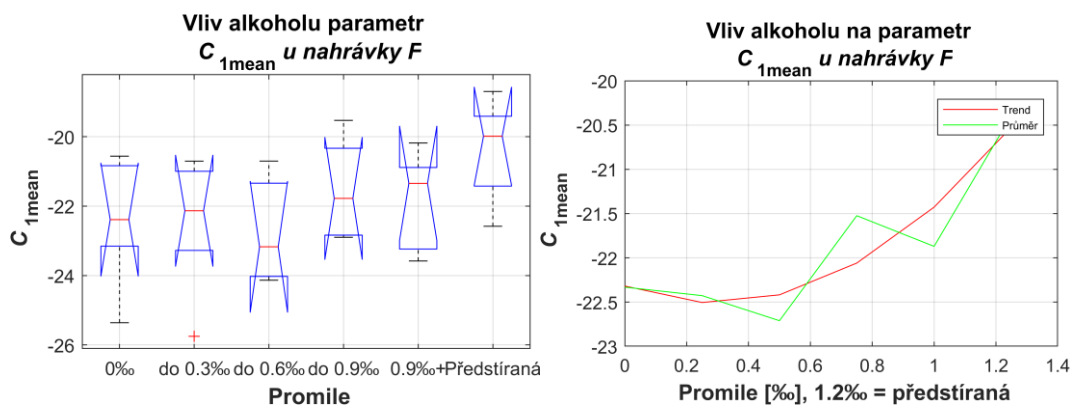
Obrázek 5.31 Změny u parametru F_2 při předstírané intoxikaci u samohlásky \dot{U} .

Kepstrální koeficienty

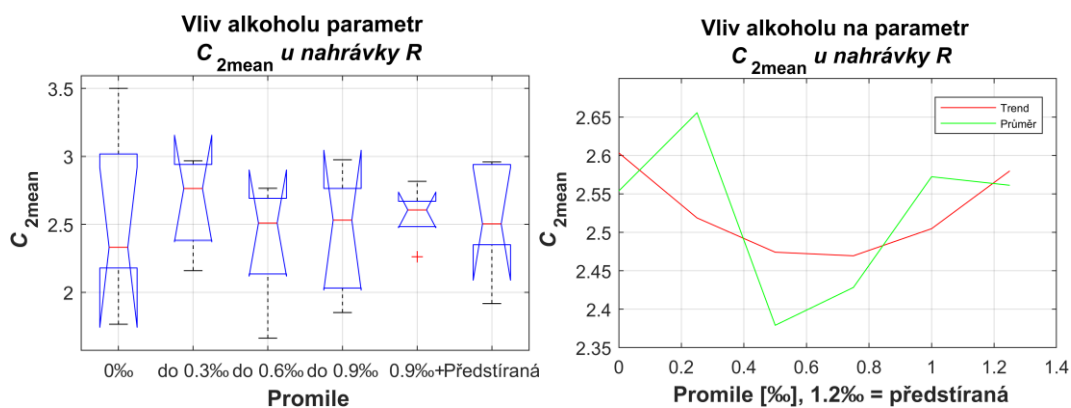
U kepstrálních koeficientů porovnáváme průměrné hodnoty prvních dvou parametrů u čteného a volného textu. Při čteném textu se kepstrální koeficient C_1 sníží a značně se zvětší rozptyl. U volného textu se hodnota zvyšuje, rozpětí se neznatelně sníží. U kepstrálního koeficientu C_2 dochází k méně rozeznatelným vlivům a hodnoty se zdají být prakticky totožné s minimálními rozdíly. U intoxikace dochází u koeficientu C_2 u volné řeči k znatelnějšímu poklesu.



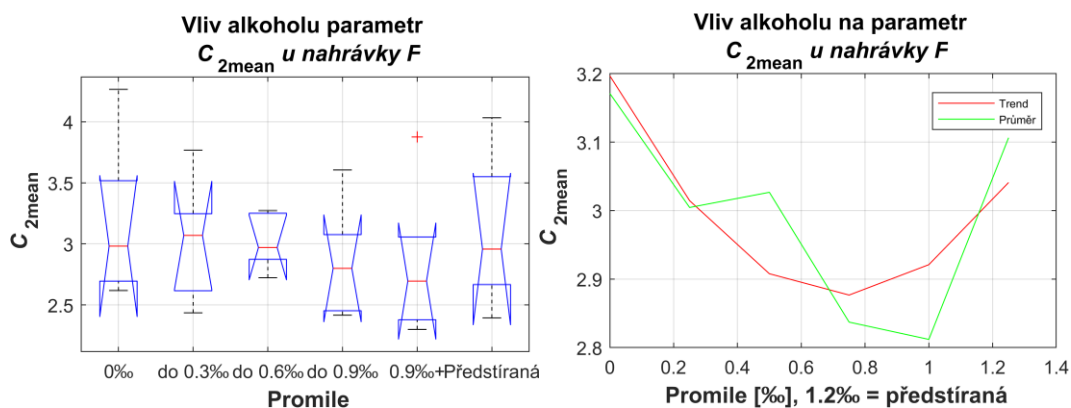
Obrázek 5.32 Změny u parametru C_1 při předstírané intoxikaci u čteného textu.



Obrázek 5.33 Změny u parametru C_1 při předstírané intoxikaci u volné řeči.



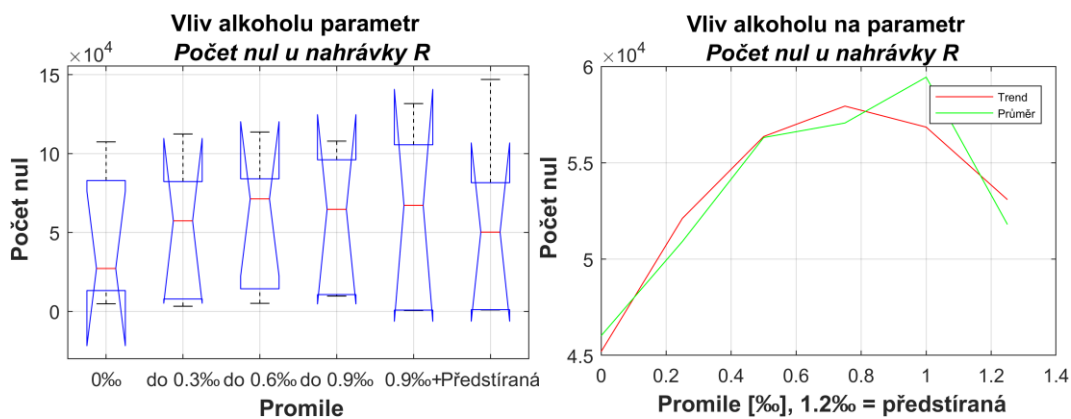
Obrázek 5.34 Změny u parametru C_2 při předstírané intoxikaci u čteného textu.



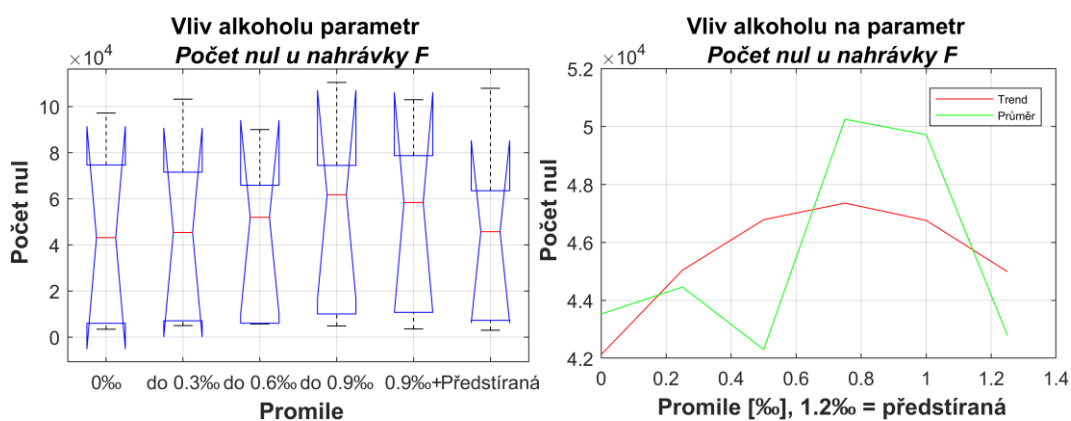
Obrázek 5.35 Změny u parametru C_2 při předstírané intoxikaci u volné řeči.

Počet průchodů nulou

U parametru počtu průchodů nulou dochází při intoxikaci u čteného textu k navýšení hodnot. U předstírané opilé řeči se mírně zvýší průměrná hodnota. Nejvíce se změni rozptyl hodnot. U volné řeči se hodnoty mění u předstírané intoxikaci jen mírně.



Obrázek 5.36 Změny u parametru *Počet nul* při předstírané intoxikaci u čtené řeči.



Obrázek 5.37 Změny u parametru *Počet nul* při předstírané intoxikaci u volné řeči.

Popis testu intoxikace užitý v aplikaci ALCTest

Z hodnot všech parametrů bychom mohli získat užitečné informace o změnách hodnot při intoxikaci. V aplikaci ALTest jsou u každého parametru střízlivé řeči vypočítány jeho maximální a minimální hodnoty a ty jsou použity jako limity pro porovnání parametrů z nové nahrávky. Následně je bez dalšího váhování spočteno procento hodnot mimo tyto limity. Hodnoty parametrů mají ale značně velké rozdíly a dochází k velmi nepřesnému vyhodnocení.

```
Jméno načteného souboru je: Á12MBDE.wav
Energie je: mimo limit
F1 je v limitech a bw1 v limitech
F2 je v limitech a bw2 v limitech
F3 je v limitech a bw3 v limitech
F13 je v limitech
F12 je v limitech
F0 je: v limitech
F0Max je v limitech F0Min je v limitech
M1 je v limitech
M2 je v limitech
M3 je v limitech
C1Max je v limitech C1Mean je v limitech
C2Max je v limitech C2Mean je v limitech
C3Max je v limitech C3Mean je v limitech
C4Max je v limitech C4Mean je v limitech
C5Max je v limitech C5Mean je v limitech
C6Max je v limitech C6Mean je v limitech
C7Max je v limitech C7Mean je v limitech
C8Max je v limitech C8mean je v limitech
JitterAbs je v limitech JitterRel je v limitech JitterRAP je v limitech JitterAPQ5 je v limitech
ShimmerLocal je v limitech ShimmerdB je v limitech ShimmerAPQ3 je v limitech ShimmerAPQ5 je v limitech
Počet průchodů nulou je v limitech
Procent parametrů mimo limit: 2.5%
```

Obrázek 5.38 Test na intoxikaci v aplikaci ALCTest.

6. ZÁVĚR

Diplomová práce seznamuje s problematikou intoxikace jedince a zkoumá možnosti detekce alkoholu v řečovém signálu. Zkoumá, jakým způsobem se dá získat informace o intoxikaci jedince bez užití dechové zkoušky a odběru tělních tekutin. Práce se zaměřuje na získání informace o intoxikaci z lidského hlasu. Z lidského hlasu se dají získat informace, které se mohou zobrazit v závislosti na čase a frekvenci. V časové oblasti se při intoxikaci objevuje možnost zkoumání změn v tempu řeči, energii, jitteru, shimmeru a fonetických změn. Ve frekvenční oblasti je možné zkoumat změny formantových příznaků F_1 až F_n , které obsahují důležité informace o charakteru samohlásek a znělých souhlásek, šířky pásem jednotlivých formantů a antiformantů. Hlavní parametr, který určuje charakteristiku mluvčího, je základní tón hlasu F_0 . Získává se pomocí kepstrální analýzy. Jeho střední hodnota je vnímána jako výška hlasu. Intoxikace alkoholem má podle dostupných studií značný vliv na tyto parametry. Práce nachází několik patentů, které úspěšně využívají porovnávání změn parametrů hlasu k samotné detekci alkoholu v řeči. Úspěšná detekce je již při 0,5‰ alkoholu v krvi.

Diplomová práce vytvořila hlasovou databázi opilé řeči v rozsahu od 0‰ do 1‰ získáním nahrávek řeči od 40 subjektů. V práci jsou zahrnuty i nahrávky profesionálně předstírané opilé řeči. Aplikací ALCTest, vytvořenou v programu MATLAB, lze vytvořit nahrávku, která se dále analyzuje. Analýza zahrnuje výpočet LPC spektra, získání frekvencí prvních tří formantů, jejich střední hodnoty a šířky pásem. Spočte základní hlasový tón, vypočte kepstrum a prvních 8 kepstrálních koeficientů. Dále se věnuje počtu průchodu nulou v nahrávce a výpočtem jitterů a shimmerů.

Všechny parametry jsou vypočítány u každé z pěti úrovní intoxikace a následně porovnány. Výsledkem jsou hodnoty největších změn parametrů při intoxikaci alkoholem.

U delších samohlásek dochází při intoxikaci ke snižování energie. Stejně tak u čteného delšího textu, naopak u volné řeči se celková energie zvětšuje.

U průměrného základního hlasového tónu $F_{0\text{mean}}$ u většiny nahrávek dochází ke zvýšení průměrné hodnoty. K významnějšímu poklesu dochází u obou pohlaví u samohlásky Ó, rozmlouvadla S a volné řeči.

U formantových příznaků sledujeme poklesy i nárůsty průměrných hodnot. Snížení u frekvence F_1 , která nám popisuje frekvenci vytvářenou v dutině hrdelní, dochází u samohlásek A, É, Í a Ú. U žen jen u samohlásek A, É, Ú. U formantu F_2 , který popisuje frekvenci dutiny ústní, dochází k průměrným poklesům u samohlásek Í, Ó a Ú. U žen u samohlásek É a Í. U formantu F_3 , který popisuje frekvenci dutiny nosní, se výrazněji snižuje u samohlásek É a Ó, stejně tak u žen. Průměr prvních tří formantů se výrazněji snižuje u samohlásky Ó a Ú. U žen u samohlásky É a Í.

Výraznější změny v parametrech jsou změna šířky pásma bw_2 u formantu F_2 u samohlásky E, kde dochází k značnému zvětšení pásma bw_2 a u samohlásky O, kde

průměrné zvětšení šířky pásma dosahuje 100% původní průměrné šířky. U žen se zvětšuje šířka pásma bw_2 o 200% průměrné šířky.

U průměrných hodnot kepstrálních koeficientů dochází k největším pozorovatelným poklesům u volné a čtené řeči u kepstrálního koeficientu C_5 . Změny v kepstrální oblasti jsou značně proměnlivé a zasloužily by si hlubší výzkum, vzhledem k již fungujícímu patentu [29]. U čtené řeči maximální hodnota všech kepstrálních koeficientů klesá jak u mužů tak u žen.

U parametrů Jitteru dochází u mužů při intoxikaci ke snižování průměrné hodnoty skoro u všech získaných nahrávek. Výjimku tvoří samohlásky Í a Ú . Ke snížení průměrných hodnot dochází i u parametrů Shimmeru s výjimkou samohlásky Ú a rozmlouvadla N . U žen dochází k výraznějšímu zvětšení hodnot parametrů Jitteru u volného i čteného textu a rozmlouvadel s výjimkou rozmlouvadla S . U Shimmerů jsou hodnoty značně rozdílné.

U parametru „Počet průchodů nulou“ dochází k nejvýraznějšímu navýšení u krátkých samohlásek O a U . U čteného textu dochází ke snížení počtu průchodu nulou a naopak u volného textu ke zvýšení. Dá se říct, že při čteném textu se při intoxikaci zrychluje čtení a snižují frekvence hlasu a při volné promluvě se naopak text prodlužuje a dochází k navýšení vyšších frekvencí.

Dále práce porovnává opilou předstíranou řeč se střízlivou a opilou řečí. Základní hlasový tón se u předstírané řeči výrazně nemění, při čtené řeči se mírně zvýší a u volné řeči mírně sníží oproti střízlivé řeči. Nejvýraznější změny pozorujeme u rozptylu hodnot kepstrálního koeficientu C_1 , kde dochází ke zvětšení rozptylu hodnot. U formantů dochází ke zvětšení formantu F_2 u samohlásky Ú .

V aplikaci ALCTest je použito srovnání všech maximálních a minimálních hodnot parametrů pro limity testu opilé řeči. U nové neznámé nahrávky jsou vytvořeny všechny uvedené analýzy a parametry jsou porovnány. Pokud jsou nějaké parametry mimo limit, je nahrávka vyhodnocena v procentech parametrů mimo limit.

Metoda porovnání všech parametrů ze všech nahrávek a od všech respondentů se nejvíce jeví jako příliš přesná. Pro další zpřesnění je možné použít metodu na jedince nebo pro porovnání všech parametrů metody Deep Learning.

LITERATURA

- [1] ŘÍZENÍ POD VLIVEM ALKOHOLU NEBO DROG. VLIV, STATISTIKY, ÚČINKY, ODBOURÁVÁNÍ, TESTY. *Bezpecnostprace.info* [online]. Magazín BezpečnostPráce.info, 2018 [cit. 2020-12-03]. Dostupné z: <https://www.bezpecnostprace.info/doprava/rizeni-pod-vlivem-alkoholu-drog/>
- [2] Popis výpočtu zbytkového alkoholu v krvi. *Bezpečné cesty* [online]. Tábor: Simopt, 2019 [cit. 2020-12-03]. Dostupné z: <https://www.bezpecnecesty.cz/cz/alkohol-kalkulacka/algorithmus-vypoctu>
- [3] Stanovení alkoholu v organismu člověka. Ing. Jaroslav Zikmund [online]. Praha, 2019 [cit. 2020-12-03]. Dostupné z: <http://www.zikmund.org/alkohol/stanoveni.htm>
- [4] PSUTKA, Josef. *Mluvíme s počítačem česky*. Praha: Academia, 2006. Česká matice technická (Academia). ISBN 80-200-1309-1.
- [5] PAVLÍČEK, Tomáš. *Zpracování řečových signálů ve spektrální oblasti*. Brno, 2012. Semestrální práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Prof. Ing. Milan Sigmund, CSc.
- [6] SIGMUND, Milan. *Rozpoznávání řečových signálů: přednášky*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2007. ISBN 978-80-214-3526-1.
- [7] SIGMUND, Milan. *Analýza řečových signálů: přednášky*. Brno: Vysoké učení technické, 2000. ISBN 80-214-1783-8.
- [8] KORTUS, Jiří. *Laboratorní úloha Skrývání dat ve zvukovém záznamu*. Brno, 2014. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce doc. Ing. Karel Burda, CSc.
- [9] WAV. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, [cit. 2020-12-03]. Dostupné z: <https://en.wikipedia.org/wiki/WAV>
- [10] SOBELL, L.C., M.B. SOBELL a R.F. COLEMAN. Alcohol-Induced Dysfluency in Nonalcoholics. *Folia phoniatrica* [online]. 1982, (34), 316–323 [cit. 2020-12-10]. Dostupné z: doi:10.1159/000265672
- [11] KÜNZEL, H.J. a A. BRAUN. The Effect of Alcohol on Speech Prosody. In: *Proc. of International Congress of Phonetic Sciences (ICPhS)*. Barcelona: UAB, 2003, s. 2645–2648. ISBN 1-876346-48-5.
- [12] HOLLIEN, H., JONG DE, G., MARTIN, C.A., R. SCHWARTZ, a K. LILJGREN. Effects of Ethanol Intoxication on Speech suprasegmentals. In: *Journal of the Acoustical Society of America*. 2001, 110 (6), 3198–3206.
- [13] COONEY, O.M., MCGUIGAN, K., P. MURPHY a R. CONROY. Acoustic Analysis of the Effects of Alcohol on the Human Voice. In: *Journal of the Acoustical Society of America*. 1998, 103 (5), 2895.

- [14] KLINGHOLZ, F., R. PENNING. a E. LIEBHARDT. Recognition of Low-Level Alcohol Intoxication from Speech Signal. In: *Journal of the Acoustical Society of America*. 1988, 84 (3), 929–935.
- [15] JOHNSON, K., D.B PISONI a R.H. BERNACKI. Do Voice Recordings Reveal Whether a Person is Intoxicated? A case study. In: *Phonetica*. 1990, 41, 215–237.
- [16] SHIEL, F., Ch. HEINRICH a V. NEUMEYER. Rhythm and Formant Features for Automatic Alcohol Detection. In: *Phonetik.uni-muenchen.de* [online]. 2010, 4 [cit. 2020-12-04]. Dostupné z: <https://www.phonetik.uni-muenchen.de/forschung/publikationen/Schiel-IS2010.pdf>
- [17] LEVIT, M., R. HUBER, A. BATLINER a E. NOETH. Use of Prosodic Speech Characteristics for Automated Detection of Alcohol Intoxication. In: *Prosody in Speech Recognition and Understanding* [online]. NJ, 2001, October 22-24 2001, s. 103–106 [cit. 2020-12-01]. Dostupné z: https://www.isca-speech.org/archive_open/prosody_2001/prsr_019.html
- [18] SIGMUND, M., A. PROKES a P. ZELINKA. Detection of Alcohol in Speech Signal using LF Model. In: *Artificial Intelligence and Applications* [online]. Calgary, AB, Canada: ACTAPRESS, 2010, 2010, s. - [cit. 2020-12-01]. ISBN 978-0-88986-817-5. Dostupné z: doi:10.2316/P.2010.674-062
- [19] SCHIEL, F. a C. HEINRICH. Laying the Foundation for In-car Alcohol Detection by Speech. In: *Interspeech 2009* [online]. Brighton, 2009, s. 983–986 [cit. 2020-11-30]. Dostupné z: https://www.isca-speech.org/archive/archive_papers/interspeech_2009/papers/i09_0983.pdf
- [20] SCHIEL, F., C. HEINRICH a S. BARFÜSSER. Alcohol language corpus: the first public corpus of alcoholized German speech. *Lang Resources & Evaluation* [online]. 2012, (46), 503–521 [cit. 2020-10-10]. Dostupné z: <https://doi.org/10.1007/s10579-011-9139-y>
- [21] SCHULLER, B., STEIDL, S., BATLINER, A., F. SCHIEL. a J. Krajewski. The Interspeech 2011 Speaker State Challenge. *Proc. of Interspeech* [online]. 2011 Florence, 3201–3204 [cit. 2020-10-10]. Dostupné z: <https://www.phonetik.uni-muenchen.de/forschung/publikationen/Schuller-IS2011.pdf>
- [22] HEINRICH, C. a F. SCHIEL, CHS. (2014) The Influence of Alcoholic Intoxication on the Short-time Energy Function of Speech. *The Journal of the Acoustical Society of America* [online]. 2014, 135:5, 2942-2951 [cit. 2020-11-11]. Dostupné z: <https://doi.org/10.1121/1.4870705>
- [23] European Parliament Proceedings Parallel Corpus 1996-2011. www.statmt.org [online]. UK: EuroMatrixPlus, 2011 [cit. 2020-12-04]. Dostupné z: <https://www.statmt.org/europarl/>
- [24] Český národní korpus. *Český národní korpus* [online]. Ústav Českého národního korpusu Panská 890/7 110 00 Praha 1: Ústav Českého národního korpusu, 2020 [cit. 2020-12-05]. Dostupné z: <https://www.korpus.cz/>

- [25] SCHIEL, F., C. HEINRICH a V. NEUMEYER. Rhythm and Formant Features for Automatic Alcohol Detection. Proc. {{Interspeech}} 2010 [online]. 2010, 458-461 [cit. 2020-12-11]. Dostupné z: <https://www.phonetik.uni-muenchen.de/forschung/publikationen/Schiel-IS2010.pdf>
- [26] BAUMEISTER, B., C. HEINRICH a F. SHIEL. The Influence of Alcoholic Intoxication on the Fundamental Frequency of Female and Male Speakers. *The Journal of the Acoustical Society of America* [online]. 2012, 132, 442. [cit. 2020-12-05]. Dostupné z <https://doi.org/10.1121/1.4726017>
- [27] STEENEKEN, H.J.M. a J.H.L. HANSEN. Speech under Stress Conditions: Overview of the Effect on Speech Production and on System Performance. *ICASSP '99* [online]. 2012, 4, 2079–2082. . [cit. 2020-12-01]. Dostupné z <https://ieeexplore.ieee.org/document/758342>
- [28] SCHERER, Klaus R. Speech and emotional states. *Speech evaluation in psychiatry*. 1981, 189-220.
- [29] LEE, So-Woo. SAMSUNG ELECTRONICS CO LTD. *Controller for starting vehicle by means of sound*. KR. CN1110790C. Uděleno 2003-06-04. Zapsáno 1998-03-04.
- [30] BAE, Myung-Jin; BAE, Seong-Geun; LEE, Sang-Gil. SOONGSIL UNIVERSITY FOUNDATION INDUSTRY-ACADEMIC COOPERATION FOUNDATION. *Controller for starting vehicle by means of sound*. 2016. KR. KR101621780B1. Uděleno 2016-05-17. Zapsáno. 2015-02-10
- [31] HIDEAKI, Hirano. HIRANO ZANAVY INFORMATICS KK. *DRUNKEN STATE DETECTION DEVICE*. JP. JP2009096327A. Uděleno 2009-05-07. Zapsáno 2016-05-17.
- [32] LEE, See Woo. *A guide system of drinking condition using speech signal and communication network of wireless or wire*. 2002. KR. KR100497837B1. Uděleno 2005-06-28. Zapsáno 2002-10-16.
- [33] Audacity. *Cs.wikipedia.org* [online]. Wikimedia Foundation, 7. 12. 2019n. 1. [cit. 2020-12-07]. Dostupné z: <https://cs.wikipedia.org/wiki/Audacity>
- [34] KRÁČALA, Martin. *Fonetické zpracování řeči pomocí software Praat*. Brno, 2014. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce prof. Ing. Milan Sigmund, CSc.
- [35] Matlab-gui [online]. [cit. 2020-12-07]. Dostupné z: <https://www.mathworks.com/discovery/matlab-gui.html>
- [36] BLOOD ALCOHOL LEVEL. *Www.notredamecollege.edu* [online]. South Euclid, Ohio: Phoenix House, 2020 [cit. 2020-12-11]. Dostupné z: <https://www.notredamecollege.edu/sites/default/files/-Blood-Alcohol-Chart.pdf>
- [37] *214219_datasheet* [online]. In: . Japan: Audio-Technica U.S., 2004, s. 1 [cit. 2021-5-14]. ISBN P51680-EN. Dostupné z: https://images.static-thomann.de/pics/atg/atgdata/document/specs/214219_datasheet.pdf