

Univerzita Karlova v Praze

Filozofická fakulta

Fonetický ústav

Diplomová práce

Lenka Vyhnálková

Vliv vzdělání na schopnost maskovat svůj hlas

The effect of education on the ability to disguise one's voice

Praha 2013

Vedoucí práce: Doc. Mgr. Radek Skarnitzl, Ph.D.

Ráda bych poděkovala Doc. Mgr. Radku Skarnitzlovi, Ph.D., za odborné vedení této práce a za cenné rady a připomínky při jejím zpracovávání. Také bych na tomto místě chtěla poděkovat všem participantům, bez nichž by tato práce nemohla vzniknout.

Prohlášení:

Prohlašuji, že jsem diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 16. ledna 2013

.....

Jméno a příjmení

Abstrakt (česky)

Maskování hlasu se může potenciálně objevit v každém řečovém projevu spojeným s kriminálním případem. Aby bylo možné pachatele správně identifikovat, je třeba analyzovat řečový materiál a porozumět tomu, jakým způsobem mohou jednotlivé typy maskování ovlivňovat hlasové kvality mluvčího. Tato práce se zabývá schopností maskování hlasu u tří skupin mluvčích rozdělených podle úrovně nejvyššího dosaženého vzdělání. Jejím cílem bylo zjistit, jaké strategie mluvčí ve snaze o zakrytí identity používají a zda se ve volbě těchto technik a v jejich úspěšnosti jednotlivé skupiny liší. Výzkum je založen na 86 nahrávkách 43 mladých lidí ve věku od 20 do 31 let, které byly pořízeny v Praze a Plzni. Od každého participanta byly získány dvě čtené promluvy, nemaskovaná a maskovaná, které byly mezi sebou porovnávány. Výsledky ukazují, že preferovanými technikami používanými k maskování hlasu byly fonační změny, zejména snížení a zvýšení základní frekvence hlasu. Mezi třemi zkoumanými skupinami se sice objevily rozdíly nejen v povaze strategií maskování, ale i v jejich úspěšnosti, ale tyto rozdíly byly tak malé, že k jejich potvrzení by bylo třeba dalšího výzkumu.

Klíčová slova (česky)

hlas, identita mluvčího, maskování hlasu, forenzní fonetika, sociofonetika

Abstract (in English):

Voice disguise can potentially occur in every utterance that is associated with any criminal case. In order to identify the perpetrator it is necessary to analyze the speech and understand how the different types of voice disguise can affect the speaker's voice qualities. This thesis focuses on the ability of voice disguise, portraying three groups of speakers in relation to their educational background. The aim of this work is to determine the strategies adopted by the speaker to conceal his/her identity and furthermore it poses the question whether differences among the three groups of speakers, their choice of strategy and its inherent success can be found. The basis for this research stems from 86 recordings which were undertaken in Pilsen and Prague with 43 young people aged 20 to 31. Two read utterances, one undisguised and the other freely disguised, were obtained from each of the participants and were compared with each other. The results show that the preferred forms of voice disguise appeared to involve changes in phonation - especially decrease or increase of fundamental frequency of the speaker's voice. Among the three groups of speakers, their choice and the success of the chosen strategy only minor differences could be found, yet for a final confirmation of this observation more research is needed.

Klíčová slova (anglicky):

voice, speaker identity, voice disguise, forensic phonetics, sociophonetics

Obsah

1 ÚVOD.....	9
2 SOCIOFONETIKA.....	12
2.1 Historie.....	12
2.2 Vymezení pojmu	12
2.3 Oblasti zkoumání.....	14
2.3.1 <i>Produkce řeči</i>	14
2.3.2 <i>Percepce řeči</i>	17
2.3.3 <i>Úrovně zkoumání</i>	19
2.4 Používané metody	20
2.4.1 <i>Sběr dat</i>	20
2.4.2 <i>Analýza dat</i>	20
2.4.3 <i>Percepční studie</i>	21
2.5 Epizodický model řečové percepce	22
2.6 Možnosti využití sociofonetiky	23
3 FORENZNÍ FONETIKA	24
3.1 Oblasti využití	24
3.1.1 <i>Sporné promluvy</i>	25
3.1.2 <i>Určování autenticity nahrávek</i>	25
3.1.3 <i>Identifikace pachatele svědkem</i>	25
3.1.4 <i>Rozpoznání mluvčího</i>	30
3.2 Používané metody	34
3.2.1 <i>Poslechová analýza</i>	34
3.2.2 <i>Akustická analýza</i>	35
3.2.3 <i>Voiceprint</i>	36
3.2.4 <i>Automatické rozpoznávání</i>	36
3.2.5 <i>Komerční versus forenzní rozpoznání mluvčího</i>	38
3.3 Variabilita, klasifikace mluvčího	38
3.3.1 <i>Zdroje variability mluvčích</i>	39
3.3.2 <i>Klasifikační charakteristiky mluvčího v současné forenzní praxi</i>	41
4 MASKOVÁNÍ HLASU	44
4.1 Typy záměrného maskování hlasu	45
4.1.1 <i>Elektronické maskování hlasu</i>	45

4.1.2 Neelektronické maskování hlasu	46
4.2 Maskování hlasu ve forenzní praxi.....	54
5 METODA.....	56
5.1. Charakteristika výzkumného vzorku	56
5.2 Nahrávání a výzkumný materiál.....	56
5.3. Způsob zpracování dat	57
5.3.1 Poslechová analýza	57
5.3.2 Akustická analýza.....	57
5.4 Poslechový test	60
6 VÝSLEDKY	62
6.1 Poslechová analýza.....	62
6.1.1 Přehled použitých strategií maskování hlasu.....	62
6.1.2 Některé forenzně relevantní rysy nalezené v projevech.....	73
6.2 Akustická analýza.....	74
6.2.1 Základní frekvence	74
6.2.2 Vokalické formanty.....	82
6.2.3 HNR, jitter, shimmer	90
6.3 Ilustrace konkrétních případů maskování hlasu	97
6.3.1 Fonační změny	97
6.3.2 Denazalizace	100
6.3.3 Změny v artikulaci hlásek.....	101
6.4 Poslechový test	106
7 DISKUSE	110
7.1 Poslechová analýza.....	110
7.2 Akustická analýza.....	111
7.3 Poslechový test	113
7.4 Limity práce	113
7.5 Návrhy k dalšímu výzkumu	114
8 ZÁVĚR	116
SEZNAM POUŽITÝCH ZDROJŮ	118
PŘÍLOHY	126
Příloha 1 Informovaný souhlas	126
Příloha 2 Text 1 – normální čtený projev	127

Příloha 3 Text 2 – maskovaný projev.....	128
Příloha 4 Věty použité v poslechovém testu.....	129
Příloha 5 Instrukce k poslechovému testu	132
Příloha 6 Histogramy.....	133
Fonetici.....	133
Nelingvisti	138
Mluvčí bez maturity.....	143

1 Úvod

Většina lidí je během krátkého časového okamžiku schopná rozpoznat známé lidi bez očního kontaktu, pouze na základě jejich hlasového projevu, například po telefonu. V případě, že člověk takto slyší neznámého mluvčího, bývá zpravidla poměrně jednoduché určit, zda se jedná o ženu či muže, dítě, dospělého nebo starého člověka, rodilého mluvčího či cizince. Někdy lze i na základě nářečních rysů rozpoznat to, z jaké části země mluvčí pochází, či na základě jeho ústního projevu odhadnout úroveň jeho vzdělání. Význam těchto schopností narůstá ve forenzním kontextu, v němž často bývá nutné identifikovat neznámé hlasy, a to jak při rozpoznávání pachatele trestného činu jeho obětí či svědkem, tak i forenzními experty.

Práce forenzních expertů bývá ztížena nejen variacemi uvnitř řeči jednoho mluvčího (např. vlivem emocí), ale také kvalitou dostupného zvukového materiálu a rizikem maskování hlasu mluvčího. Toto maskování představuje záměrnou změnu hlasu za účelem skrytí identity mluvčího, potenciálně je jím ohrožen každý řečový materiál v kriminálním případě, jež expert za účelem analýzy obdrží. Pravděpodobnost toho, že se pachatel pokusí zakrýt svoji identitu prostřednictvím maskování svého hlasu, je závislá na typu kriminálního činu, jehož se dopouští. Tato pravděpodobnost stoupá zejména tehdy, když si je pachatel vědom možnosti nahrávání svého projevu za účelem policejního šetření (Braun, 2006: 178; de Figueiredo & de Souza Britto, 1996: 168; Künzel, Gonzales-Rodriguez, & Ortega-García, 2004; Masthoff, 1996: 160; Zhang & Tan, 2008: 118).

Mluvčí má k dispozici velké množství technik, jež může k zamaskování svého hlasu použít. Nejen že existují elektronické nástroje schopné změnit některé z hlasových parametrů, ale zároveň i mluvčí sám může ve větší či menší míře svoji řečovou produkci záměrně ovlivnit. Ve forenzní praxi je pak důležité správně určit způsob maskování hlasu, aby bylo možné dále hodnotit, jakých fonetických (a případně lingvistických) aspektů řeči se snahy o maskování hlasu týkají a v jakém rozsahu byly charakteristické vlastnosti hlasového projevu mluvčího modifikovány. Díky těmto poznatkům je pak možné pachatelův normální hlasový projev do značné míry rekonstruovat.

Maskování hlasu se v zahraničí (zejména v Německu) začalo zkoumat již během sedmdesátých let, v posledních desetiletích se vědci snaží vyvinout i automatické systémy, jež by byly schopny maskování hlasu rozpoznat, a i přes jeho použití mluvčího správně

identifikovat (Perrot & Chollet, 2008: 5682). V českém kontextu nebylo dosud maskování hlasu jako takové detailněji zkoumáno.

Tato diplomová práce se zabývá schopností maskování hlasu u tří skupin mluvčích rozdělených podle úrovně a typu dosaženého (nebo případně ještě neukončeného) vzdělání. První z nich sestává z absolventů a studentů vyšších ročníků oboru Fonetika na Filozofické fakultě Univerzity Karlovy, druhou představují absolventi a studenti nelingvistických oborů na vysokých školách v Praze a Plzni, do třetí skupiny byli zahrnuti mluvčí, jejichž nejvyšší dosažené vzdělání bylo nižší než středoškolské s maturitou. Cílem této práce je na základě pořízených nahrávek zjistit, jaké techniky si mluvčí ve snaze o maskování hlasu nejčastěji osvojují, zda využívají i jejich kombinací a jak úspěšné tyto techniky ve skutečnosti jsou. Dále se zaměřuje na to, zda se jednotlivé skupiny mluvčích budou mezi sebou významně lišit jak ve výběru maskovacích technik, tak i ve schopnostech maskování svého hlasu.

Teoretická část je členěna do tří kapitol, které se k výzkumu provedenému v této práci vztahují. První z nich se věnuje oblasti sociofonetiky, její historii, vymezení jejího pojmu, jí zkoumaným oblastem, metodám, jež používá, i možnostem jejího využití. Na ni navazuje kapitola, která představuje forenzní fonetiku jako disciplínu a oblasti, v nichž se využívá - sporné promluvy, určování autenticity nahrávek, identifikaci pachatele svědkem a rozpoznání mluvčího. V rámci této kapitoly jsou představeny také metody používané ve forenzním výzkumu včetně již metody zastaralé, tzv. voiceprintu. Další část se pak zabývá zdroji variability mezi mluvčími i v rámci projevů jednoho mluvčího a klasifikačními charakteristikami mluvčího v současné forenzní praxi. Poslední kapitola teoretické části se věnuje maskování hlasu. Představuje jeho typy a dále rozebírá neelektronické maskování hlasu, jeho nejčastěji se vyskytující realizace a výsledky již dříve provedených výzkumů.

Metodologická část práce charakterizuje výzkumný vzorek a nastiňuje kritéria, která byla při jeho výběru zohledňována, popisuje výzkumný materiál i průběh nahrávání participantů. Je v ní nastíněn způsob zpracování získaných dat i sestavení poslechového testu.

Výsledky vlastního výzkumu jsou prezentovány ve čtyřech kapitolách, první představuje výsledky poslechové analýzy prostřednictvím přehledu použitých strategií maskování hlasu, druhá popisuje výsledky akustické analýzy týkající se hlasových parametrů, jako je základní frekvence, vokalické formanty a jitter, shimmer a harmonicita. Předposlední kapitola této části práce ilustruje některé konkrétní případy maskování hlasu, jejichž realizace

jsou dobře viditelné v oscilogramech a spektrogramech. Závěrečná kapitola se zabývá výsledky provedeného poslechového testu.

2 SOCIOFONETIKA

2.1 Historie

Termín sociofonetika se poprvé objevil v disertační práci Deshaie-Lafontainové v roce 1974 (Foulkes, Scobbie, & Watt, 2010: 703) a již od roku 1979 se sociofonetika objevuje jako samostatná tematická sekce na ICPHS¹. V jejím rámci zde bylo předneseno devět příspěvků, ale pod tímto pojmem se ještě stejně jako u Deshaie-Lafontainové rozuměly centrální otázky sociolingvistiky s ohledem na fonetické proměnné (Foulkes a kol., 2010: 703). Názory na to, kdy byly realizovány první vyloženě sociofonetické výzkumy, se liší. Podle Thomase (2011: 8) byly prováděny ještě před vznikem samotného termínu v rámci sociolingvistiky a za jednoho z jejich prvních zástupců lze podle něj považovat Williama Labova. Naopak Foulkes a kol. (2010: 703) za první opravdu sociofonetickou práci vydanou ve významném časopise považují až studii o vídeňské němčině napsanou Dresslerem a Wodakovou roku 1982.

Až do poloviny devadesátých let byl pojem sociofonetika poměrně neznámý a jen zřídkakdy užívaný (Thomas, 2011: 1). V polovině devadesátých let zájem o sociofonetiku vzrůstá, s čímž souvisí i postupné rozšíření jejího záběru (Foulkes, 2005: 496; Foulkes a kol., 2010: 703). Vzestup její oblíby se projevuje i na konferencích, například v roce 2003 bylo na ICPHS předneseno již kolem devadesáti příspěvků dotýkajících se sociolingvistiky (Foulkes a kol., 2010: 703), sociofoneticky zaměřené příspěvky se objevují také na Meeting of the Acoustical Society of America (Thomas, 2011: 9).

2.2 Vymezení pojmu

Přestože byl termín sociofonetika poprvé použit již v sedmdesátých letech, její jednoznačné vymezení chybí. Sociofonetika se nachází na pomezí sociolingvistiky a fonetiky (Baranowski, v tisku; Thomas, 2011: 1), představuje jejich vzájemnou integraci, integraci jejich principů, technik a teoretických rámců (Foulkes a kol., 2010: 703). Avšak náhledy na to, co přesně lze pod pojmem sociofonetika rozumět, se značně rozcházejí. Dosud je mnoha fonetiky a sociolingvisty dokonce nahlížena jen jako metodologický konstrukt, zatímco jiní ji již chápou i ve smyslu určitého teoretického rámce (Thomas, 2011: 1). Za samostatnou disciplínu se svými vlastními otázkami, technikami a způsoby vědeckého dokazování zatím sociofonetika považována nebývá a otázkou zůstává, zda tímto směrem bude v budoucnosti směřovat (Baranowski, v tisku; Labov, 2006: 501).

¹ The International Congress of Phonetic Sciences.

Pojem sociofonetika je nejrozšířenější mezi fonetiky, kteří jej užívají k popisu řečové produkce ve spojitosti s dialekty, skupinami mluvčích a řečovými styly (Foulkes, 2005: 495; Foulkes & Docherty, 2006: 410). Používá se ale také ve foneticky orientovaném výzkumu pod záštitou variační sociolingvistiky. Takto zaměřené studie se sice věnují vztahu fonetických (fonologických) forem a sociálních faktorů, ale vždy se zvláštním zřetelem k postižení jazykové změny (Foulkes, 2005: 495). Podle Thomase (2011: 1) se fonetický aspekt sociofonetiky omezuje převážně na fonetiku moderní - akustickou a artikulační analýzu a percepční experimenty, ze sociolingvistických témat pak přejímá rozdíly založené na stylu, geografickém rozdělení, sociální třídě, genderu, příslušnosti ke generaci, etnické příslušnosti a lidech se stejnými zájmy.

Sociofonetika předpokládá, že se mluvčí svému prostředí přizpůsobuje i ve své výslovnosti, proto nahlíží na odchylky a změny jako na základní vlastnosti jazyka (Thomas, 2011: 2). Variabilita je jedním z charakteristických rysů řeči, neexistují dva stejné hlasy ani dvě stejné promluvy. Tato variabilita řeči je způsobena mnohými faktory, biologickými (artikulační omezení, aerodynamické jevy, akustické jevy spojené s vokálním traktem), jazykově specifickými (fonologické znalosti mluvčích daného jazyka) a sociálními (Foulkes & Docherty, 2006: 409-410).

Právě sociofonetická variace tvoří jeden ze základních pojmů, jimiž se sociofonetika zabývá. Sociofonetické variace představují vzorce chování naučené skrze zkušenost mluvčích s užíváním jazyka v sociální interakci (Foulkes & Docherty, 2006: 411). Odkazují „k proměnlivým aspektům fonetické nebo fonologické struktury, v níž alternativní formy korelují se sociálními faktory“ (tamtéž). Avšak určení toho, které řečové variace jsou pouhými sociálními produkty nezávislými na biologických vlastnostech vokálního traktu, nebývá vždy jednoznačné (Foulkes a kol., 2010: 705-706). Často proto za sociofonetickou variaci bývají považovány ty variace, které jsou alespoň částečně produktem sociální konstrukce (Foulkes & Docherty, 2006: 412; Foulkes a kol., 2010: 706).

Sociofonetické variace nejčastěji odkazují na sociální faktory jako gender, věk, etnicita, sociální třída mluvčího, příslušnost ke skupinám nebo zeměpisný původ (Foulkes & Docherty, 2006: 411) a mohou být spojeny nejen s produkcí, ale i percepcí řeči (Smith, Hall, & Munson, 2010: 2). Fonetická a fonologická forma může být ovlivněna také kontextem interakce, například stylem, tématem diskuse či adresátem (Foulkes, 2005: 496; Hay

& Drager, 2007: 91). Stejně tak může variace v řeči přispívat k vyjádření solidarity s druhým účastníkem pomluvy nebo odstupu od něj (Foulkes, 2005: 496).

Základním úkolem sociolingvistiky je zjistit, které sociální kategorie jsou v řeči zachyceny a jak tyto kategorie definovat (Foulkes & Docherty, 2006: 411). Ve spojitosti se sociálním významem v řeči se v sociolingvistice používají dva základní distinktivní pojmy, jazykové styly a jazykové stereotypy. Styl bývá chápán jako soubor lingvistických rysů sloužící určitému komunikačnímu cíli nebo odrážející určitou sociální identitu. Stereotypy jsou pak představy lidí o tom, jakým způsobem různé sociální skupiny produkují zvuky, užití stereotypů může být i součástí jazykového stylu. Toto rozdělení variací na styly a stereotypy může být užitečným prostředkem pro porozumění některým rozdílům ve významu sociofonetických variant (Smith a kol., 2010: 3-4).

Sociolingvistika se snaží přijít na to, jak jsou sociálně strukturované odchylky ve zvukovém systému osvojeny, jak jsou kognitivně uchovávány, subjektivně hodnoceny a zpracovávány v řeči a percepci. Sociofonetika pak přispívá k rozvoji fonetických a sociolingvistických teoretických modelů, její metody a data přispívají k teoretickým fonologickým modelům, modelům osvojování jazyka a dlouhodobého uchovávání lingvistických znalostí (Foulkes a kol., 2010: 704).

2.3 Oblasti zkoumání

Vymezení pole zkoumání sociofonetiky je problematické. Sociofonetika nečerpá pouze z fonetiky a sociolingvistiky, ale také z psycholingvistiky, klinické lingvistiky, osvojování prvního a druhého jazyka, teoretické fonologie a počítačové lingvistiky (Foulkes a kol., 2010: 704). Většina prací zkoumajících vztah mezi sociálními faktory a fonetickou variací se zaměřuje na produkci řeči (Drager, 2010: 473; Foulkes, 2005: 496; Foulkes & Docherty, 2006: 410; Foulkes a kol., 2010: 706; Labov, 2006: 500; Thomas, 2002: 115), k expanzi sociofonetiky v dalších oblastech (jako například percepci řeči, osvojování jazyka či forenzní fonetiky) dochází až v polovině devadesátých let (Foulkes & Docherty, 2006: 410-411).

2.3.1 Produkce řeči

Jak již bylo výše naznačeno, produkce řeči je nejčastějším zdrojem sociofonetického výzkumu. Ten byl donedávna spojován zejména s akustickou analýzou vokálů (Baranowski, v tisku; Foulkes, 2005: 496; Hay & Drager, 2007: 92; Thomas, 2011: 9), a přestože se

v současnosti věnuje pozornost i instrumentální analýze jiných typů zvuků řeči, analýza vokálů a jejich změn zaujímá stále vedoucí postavení (Baranowski, v tisku). Vokály byly nejprve analyzovány hlavně z hlediska formantů, akustická analýza se ale postupně rozšířila i na další aspekty vokalické kvality, jako trvání či rychlost formantových tranzientů (Foulkes, 2005: 496-497). Právě měření vokálů na základě prvních dvou formantů a zanedbávání dalších percepčně relevantních jevů jako trvání samohlásky, dynamiky formantů, šířky vlnového pásma formantů a dalších formantů (F3 a vyšší), které přispívají ke kvalitě vokálu, představují nejproblematičtější jevy metodologické stránky analýzy vokálů (Foulkes a kol., 2010: 718).

Analýza konsonantů a prozodických jevů se objevuje znatelně později (Thomas, 2011: 9-10). Na rozdíl od analýzy vokálů užívá velká většina sociofonetických studií zabývajících se konsonanty především poslechovou analýzu, protože je za prvé často méně časově náročná než analýza instrumentální a za druhé je někdy obtížné najít jednoznačné akustické koreláty sluchového dojmu (Baranowski, v tisku).

Současné sociofonetické zkoumání řečové produkce lze rozdělit do několika podoblastí, a to zkoumání a) sociální třídy, komunit a sítí, b) věku a životních fází, c) pohlaví a genderu, d) regionálních variací, e) etnicity, rasy a bilingvismu a f) variace v rámci promluv jednoho mluvčího (Foulkes a kol., 2010: 708-716).

Zkoumání sociální třídy, komunit a sítí poukazuje na to, že jazykový výběr je jedním ze symbolických zdrojů v utváření a udržování identity jedince. V oblasti řečové produkce bylo provedeno poměrně velké množství výzkumů variability v řeči ve spojitosti se sexuální orientací mluvčího (Foulkes a kol., 2010: 709). Například Pierrehumbertová a kol. (2004) analyzovali vokály ve čtyřech krátkých větách nahraných 103 mluvčími a zjistili, že lidé s homosexuální či bisexuální orientací se ve výslovnosti vokálů liší od heterosexuálních jedinců. Jejich odlišné řečové vzorce podle autorů nejsou odrazem biologických rozdílů, ale jsou spíše naučené, jejich vokalická výslovnost je posunuta směrem k výslovnosti opačného pohlaví s heterosexuální orientací. Podobným tématem se zabýval například i Podesva (2007), jenž zkoumal užívání falzetu jednoho homosexuálního mluvčího ve třech různých situacích. Zjistil, že míra výskytu falzetu závisí na konkrétní situaci, v níž se mluvčí nachází, a že se tento typ kvality hlasu u něj objevuje zejména v expresivních promluvách. Tento expresivní význam falzetu představuje nejen způsob utváření jeho femininního vystupování, ale podle autora by mohl být i jedním z prostředků konstrukce jeho gay identity.

Sociofonetický výzkum věku a životních fází se nejčastěji zaměřuje na dětství a dospívání, protože se specifické rysy těchto dvou stádií odrážejí v řeči nejvýrazněji. Jde zejména o změny v řeči spojené se změnami referenční skupiny mluvčích (Foulkes a kol., 2010: 710-711).

Často se provádí sociofonetické studie i s ohledem na pohlaví a gender mluvčích. Tato oblast zkoumání se však soustředí zejména na „dělání genderu“², protože pohlaví lze na základě mluveného projevu zpravidla na první poslech rozlišit. Specifickou otázkou je zde především to, jakým způsobem spolu pohlaví a gender interagují (Foulkes a kol., 2010: 711-712).

Geografické rozdíly v řečové produkci mají poměrně dlouhou historii v rámci dialektologie. Důležité centrum zájmu sociofonetiků v této oblasti ztělesňuje sociální mobilita lidí, procesy národní a „supralokální“ standardizace a to, jak jsou lidé vystaveni odlišným regionálním varietám v řeči a jak si je uvědomují (Foulkes a kol., 2010: 712-713). Na regionální rozdíly na subsegmentální úrovni řeči poukazuje například Thomas (2000, viz dále).

Etnicita byla jedním z prvních témat, jimiž se sociolingvistika od časů prvních Labovových prací zabývala, byla však studována především na základě anglického jazyka. Zkoumání bilingvismu a učení se druhému jazyku pak mohou poukazovat na strategické výběry mluvčích, jež mohou přispívat k sociální identitě mluvčího (Foulkes a kol., 2010: 715).

Variace v řečových projevech jednoho mluvčího lze zkoumat prostřednictvím řečových stylů v různých komunikačních situacích nebo na základě longitudinálních výzkumů. Jeden z takových výzkumů prováděli Harrington, Palerthorpe a Watson (2000). Ti zjišťovali, zda se v průběhu času může měnit výslovnost jednoho dospělého člověka přejímáním jazykových změn vyskytujících se ve společnosti. Z analýz vánočních proslovů královny Alžběty II. odvysílaných na BBC během let 1950-1989 zjistili, že výslovnost některých jejích vokálů podléhá trendům, které jsou typické pro mladší mluvčí nižšího společenského postavení. V další studii Harrington (2006) poukazuje na dlouhodobé změny v královnině šva (snížení F1, F2, F4 a zvýšení F3) a na napjatější výslovnost vokálu [i:], která

² Pojem dělání genderu (doing gender) odkazuje na to, že náš gender (rod) není daný, ale je neustále znovu a znovu vytvářen v průběhu interakcí.

poukazuje na nepatrný posun její výslovnosti směrem k modernějším trendům³. Harringtonův výzkum byl důležitý zejména v tom smyslu, že poukázal na to, že projev dospělého člověka není stabilní, ale podléhá změně nesouvisející s procesem stárnutí (Labov, 2006: 501-502).

2.3.2 Percepce řeči

Stejně jako mluvčí produkují fonetické proměnné se sociálními významy, posluchači vnímají řeč různě v závislosti na své vlastní produkci, na své předchozí zkušenosti s dialekty a na sociálních charakteristikách, jež mluvčím připisují. Sociální informace získaná z poslechového materiálu může být poměrně konzistentní, posluchači se do určité míry shodují v identifikaci etnicity mluvčího, jeho genderu a sexuality a některých charakterových vlastností. Na druhou stranu vztah mezi fonetickou a sociální informací není jednosměrný. Fonetické varianty, které posluchač vnímá, sice mohou ovlivnit charakteristiky připsané mluvčímu, ale naopak i tyto charakteristiky mohou ovlivnit způsob vnímání těchto zvuků (Drager, 2010: 473-476).

Jak již bylo naznačeno, percepci řeči se věnuje mnohem méně studií než její produkci. Vyžaduje totiž pozornost experimentálnímu designu, velké množství příprav a často i využití syntetizéru. Do popředí mohou sociofonetický výzkum percepce posunout dva faktory. Za prvé je toto pole stále poměrně neprozkoumané a vyznačuje se tak velkým potenciálem pro další výzkum, za druhé je analýza dat mnohem méně časově náročná než analýza řečové produkce (Thomas, 2002: 115).

Oblasti současného sociofonetického zkoumání percepce řeči rozčleňují Foulkes a kol. (2010: 723-726) na čtyři základní kategorie, a to a) identifikaci indexikálních jevů, b) hodnocení indexikálních jevů, c) percepci probíhající změny a d) dopad sociální a regionální variace na percepci a zpracování.

Hlavním cílem studií týkajících se identifikace indexikálních jevů je určení toho, jaké rysy posluchači v identifikačním procesu využívají, a zda to souvisí s rysy zjištěnými ze studií řečové produkce (Foulkes a kol., 2010: 724). Jednu ze studií, kterou je možné zařadit do této kategorie, prováděli Munson a kol. (2006). Ti zkoumali, na základě jakých faktorů posluchači vyvozují sexuální orientaci mluvčích ze čtené řeči. Nejprve zjistili jemné rozdíly ve výslovnosti tří hlásek v závislosti na sexuální orientaci mluvčího. V percepčním experimentu se ukázalo, že se tyto zjištěné rozdíly částečně překrývají s prediktory, jichž posluchači pro

³ Tyto změny tak nejsou ovlivněny procesem stárnutí.

určení sexuální orientace mluvčího využívali. Jako nejsilnější prediktory vnímané sexuální orientace se ukázaly frekvence samohláskových formantů. Posluchači byli úspěšnější v odhadu sexuální orientace mužů, což podle autorů může být způsobeno i větší obeznámeností s řečovými styly homosexuálních a bisexuálních mužů než žen. Ve třetím jimi provedeném experimentu výsledky ukázaly, že vnímaná sexuální orientace silně souvisí s vnímanou výškou a srozumitelností řeči. Tím naznačují, že hodnocení sexuální orientace z jediného slova nemusí být skutečným hodnocením sexuální orientace, ale hodnocením nějakého jiného parametru, jenž je v řeči zakódován.

Studie zaměřené na hodnocení indexikálních jevů se snaží zjistit, jak se řečové jevy odráží v posluchačově představě o osobě mluvčího (Foulkes a kol., 2010: 724-725). Například Walker (2007) zkoumala percepci dvou variant (inovativní a konzervativní) konsonantů /t/ a /r/ ve spojení s vnímaným věkem a sociální třídou mluvčího. Zjistila, že použití konzervativní varianty souhlásky ve větě se signifikantně projevilo ve vnímání vyššího věku i vyšší sociální třídy mluvčích, naopak při použití varianty inovativní byla jejich vnímaná sociální třída i věk klasifikovány jako nižší.

Třetí oblast výzkumu, studium percepce probíhající změny, se soustředí na to, zda jsou posluchači schopni zachytit řečovou změnu ještě před tím, než si ji sami osvojí. Tyto tři zmíněné oblasti zkoumání založeného na percepci testují reakce posluchače na rysy ze sociolingvistické perspektivy. Poslední oblast se zaměřuje na vybavování lexikálních jednotek a rozpoznání fonémů (Foulkes a kol., 2010: 724-726).

Thomas (2002: 117) pro sociofonetický výzkum percepce používá jiné dělení. Rozděluje je do pěti základních tematických oblastí, mezi něž řadí a) schopnost posluchačů identifikovat regionální dialekt, etnicitu nebo socioekonomický status mluvčích, b) jak mohou stereotypy ovlivnit percepci zvuků, c) percepční splynutí vokálů nebo nesoulad v percepci, d) jak nářeční rozdíly ovlivňují kategorizaci fonémů, e) jak variace ovlivňují posluchačovo mínění o osobnosti mluvčího.

Studie produkce a percepce řeči se mohou navzájem doplňovat, a tím pomáhat k „porozumění tomu, jak se objevuje jazyková změna, jak se utváří stereotypy, jak je lingvistická variace uložena v mysli a jaké kognitivní procesy jsou základem užívání jazyka“ (Drager, 2010: 473).

Sociolingvistika se svými studii produkce a percepce řeči poskytuje důkaz pro to, že sociolingvistické variace mají k lingvistickým informacím vztah, který musí být uchováván v mysli mluvčích (Drager, 2010: 474).

2.3.3 Úrovně zkoumání

Korelace mezi řečovou formou a sociálními faktory lze najít na segmentální, suprasegmentální i subsegmentální úrovni (Foulkes & Docherty, 2006: 412; Foulkes a kol., 2010: 722), avšak dominantním předmětem zájmu sociofonetiky jsou zejména variace segmentální (Foulkes, 2005: 496; Foulkes & Docherty, 2006: 412; Foulkes a kol., 2010: 717). Na této úrovni je možné najít rozdíly vycházející z fonetického systému (často dialekty, ale i distribuce hlásek v závislosti na věku mluvčího), fonotaktické distribuce fonémů a lexikální distribuce fonémů a alofonické realizace (Foulkes & Docherty, 2006: 412).

Suprasegmentální jevy zkoumá sociofonetika v mnohem menší míře než fonetika. Zájem se obrací především na regionální a sociální variabilitu intonace a rytmické jevy ve srovnání jednotlivých jazyků (Foulkes a kol., 2010: 722). Rozdíly mohou být nalezeny také v realizaci přízvuku nebo jeho umístění (Foulkes & Docherty, 2006: 415). Poměrně málo pozornosti je věnováno nastavení a kvalitě hlasu (Foulkes a kol., 2010: 722), což může být způsobeno i tím, že s kvalitou hlasu může interagovat velké množství různých parametrů (Foulkes, 2002; Foulkes a kol., 2010: 722). Porozumění nastavení hlasu je však důležité nejen pro popisné účely, ale i proto, že s jinými sociofonetickými proměnnými může souviset i na segmentální úrovni (Foulkes, 2002).

Díky rozvoji instrumentálních technik je možné zkoumat sociofonetické variace i na úrovni subsegmentální. Ty se týkají relativního trvání a intenzity nebo temporální koordinace artikulačních pohybů (Foulkes & Docherty, 2006: 415). Thomas (2000) prováděl dva experimenty zaměřené na subsegmentální úroveň produkce a percepce řeči. Porovnával mluvčí z Lareda a Ohia a zjistil, že v závěrečné fázi realizace diftongu /ai/ před /-d/ nebo /-t/ se u obou skupin objevily rozdíly v jeho realizaci před znělým a neznělým konsonantem. Skupiny se však mezi sebou lišily v míře změny formantů. Na základě tohoto výsledku byl sestaven percepční test, aby ověřil, zda jsou tyto lidé schopni zjištěné rozdíly využít k rozlišení závěrečné znělé, neznělé nebo chybějící hlásky. Z výsledků vyplynulo, že skupina z Lareda, přestože vykazovala malé spektrální odlišnosti ve výslovnosti zkoumaného diftongu před znělou a neznělou, v percepčním testu pro rozpoznání slova tyto spektrální rozdíly také užívala.

2.4 Používané metody

Sociofonetika čerpá z mnoha metodologií, které byly původně vytvořeny v jiných disciplínách (Foulkes a kol., 2010: 727). Nejdůležitější body v otázce sociofonetické metodologie představují sběr dat, jejich analýza a percepční studie.

2.4.1 Sběr dat

Největší metodologické rozpětí sociofonetiky lze najít asi v metodách sběru dat, které odráží rozdíly ve výzkumných otázkách, jež si sociolingvistika na jedné a fonetika na druhé straně kladou. V sociolingvistice Labovovského typu jsou data obvykle sbírána od heterogenního vzorku mluvčích, který často zahrnuje řadu řečových stylů. Sběr dat většinou probíhá přímo v terénu. Experimentální fonetické studie nejčastěji získávají data prostřednictvím randomizovaných seznamů slov (někdy nesmyslných, avšak dodržujících fonotaktiku daného jazyka). Tato data jsou zpravidla shromážděna v ideálních akustických podmínkách od menších a homogenních skupin mluvčích, často studentů univerzit. Oba uvedené přístupy mají své silné i slabé stránky. Materiál získaný v laboratorních podmínkách podléhá lepší kontrole, usnadňuje analýzu i srovnání materiálu. Naopak analýza materiálu získaného v terénu bývá obtížnější kvůli technické kvalitě nahrávek či spontánnosti projevu. Laboratorní materiály se však zpravidla nedostanou do hloubky fonetického repertoáru mluvčího a je těžké s jistotou potvrdit, zda výslovnost v materiálu koresponduje s výslovností v přirozených situacích - celkově mohou postrádat přirozenost. Sociolingvistické metody představují problémy jiné - pro výběr vhodných dat je třeba zvážit mnoho vzorků a terénních záležitostí, jako např. definice řečové komunity, relevantní sociální a demografické rozčlenění apod. (Foulkes a kol., 2010: 727-729).

2.4.2 Analýza dat

Velké množství sociofonetického výzkumu je založeno na poslechové analýze, pro niž je charakteristické využití binárních kategorií proměnných. To je způsobeno jednak historicky (neexistence vhodného softwaru), jednak metodologicky - statistická analýza pracuje s proměnnými rozřazenými do kategorií. Nevýhodou poslechové analýzy je to, že právě kvůli práci s kategoriemi není schopna zachytit celý rozsah kontinua, na němž se fonetické proměnné nacházejí (Hay & Drager, 2007: 91-92). Využití akustické analýzy však nutně neznamená, že je tento způsob analýzy lepší než pečlivá poslechová analýza. Ta je sice vystavena rizikům lidské chyby a subjektivity, ale i akustická analýza se může stát předmětem

podobných chyb. Je totiž ovlivněna výběrem nahrávacího zařízení, softwarem, nastavením analýzy, kritérii měření apod. (Foulkes a kol., 2010: 730).

Instrumentální artikulační fonetika se v sociofonetice využívá minimálně. Artikulační data by však mohla být nápomocná ke zjišťování nepředvídatelných vztahů mezi zvukem řeči, vokálním traktem a sociálními proměnnými. Pro získání tohoto typu dat se jako vhodné jeví zejména ultrazvukové zobrazení jazyka - Ultrasound Tongue Imaging (UTI) a elektropalatograf (EPG). Kromě těchto způsobů analýz se samozřejmě v sociofonetickém výzkumu provádí i analýza statistická (Foulkes a kol., 2010: 730-733).

Drager a Hay (2012) poukazují na možnosti využití kombinace kvantitativních a kvalitativních metod v sociolingvistickém výzkumu. V případové studii dívek ze střední školy ukazují, jak lze prostřednictvím kvalitativních dat vysvětlit data kvantitativní. Akustická analýza se soustředila na slovo „like“, přičemž bylo zjištěno, že výslovnost konsonantu /k/ má spojitost s dívčí identitou a příslušností mluvčích k jedné ze dvou skupin studentek. Odchyly ve výslovnosti v porovnání s příslušností ke skupině pak byly vysvětleny prostřednictvím kvalitativních dat.

2.4.3 Percepční studie

Nejnáročnější částí percepční studie je její příprava. Je třeba pečlivě vybrat mluvčí (častěji se používají reálné hlasy než uměle vytvořené) a typ projevu (čtený nebo mluvený), který bude produkován. Důležitá je i kvalita nahrávky, zejména pokud chce výzkumník užít syntetizátor řeči. Pokud se řeč bude syntetizovat, musí se zvolit typ syntézy i vhodný syntetizátor (nejčastěji se využívá LPC a Klattova syntetizátoru). Po přípravě a nahrání řečového materiálu dochází k samotné přípravě poslechového testu. Nejprve je nutné se rozhodnout, zda budou k testu použita sluchátka. Protože sluchátka nepřispívají k přirozené povaze prostředí, často se od nich výzkumníci odklánějí. V tom případě se však musí maximálně zamezit výskytu rušivých zvuků z okolí. Výzkumník by měl být informován i o případných sluchových vadách posluchačů. Důležité je také to, jakým způsobem je stimul posluchačům prezentován a jakým způsobem jsou odpovědi zaznamenávány (Thomas, 2002: 130-135).

Samotné percepční experimenty je možné rozdělit na pět nejčastějších oblastí zkoumání. Posluchači mají za úkol a) rozhodnout, zda dva stimuly zní stejně, či odlišně, b) rozhodnout, které dva stimuly ze tří či více zní stejně, c) vybrat, který foném, slovo, etnická

skupina apod. ze dvou nebo tří výběrů stimulu se spolu nejvíce shodují, d) identifikovat stimulus a zapsat jej, e) posoudit, jak realisticky, přirozeně či typicky zní daný zvuk pro určitou skupinu (Thomas, 2002: 135).

2.5 Epizodický model řečové percepce⁴

Se sociofonetikou bývají nejčastěji spojovány epizodické modely řečové percepce. Vychází (na rozdíl od generativních teorií) z toho, že obsah paměti není omezen, a tudíž se v ní nenachází pouze lingvistické entity, ale i promluvy jednotlivých lidí. Samozřejmě, že si posluchači nepamatují všechny promluvy, které kdy slyšeli, ale podle této teorie z nich vytváří statistické porozumění řeči, jíž jsou vystaveni (Thomas, 2011: 265). Epizodický model představuje jediný teoretický rámec, který začleňuje indexikalitu do fonologické znalosti. Znalost lingvistické struktury je zde založena na paměti všech jazykových zkušeností jedince, každý exemplář obsahuje současně lingvistickou i nelingvistickou informaci - akustický obraz totiž poskytuje informace o tom, kdo mluvil a jak zněl hlas mluvčího. Předpokládá tedy, že v paměti nejsou lingvistické informace uloženy bez sociálních faktorů, na něž poukazují, a že vztahy mezi nimi jsou v paměti vytvořeny automaticky (Foulkes & Docherty, 2006: 426).

Epizodické modely řečové percepce byly vyvinuty za účelem objasnit určité aspekty lingvistického chování lépe než generativní teorie. Jsou postaveny na foneticky i sociálně motivované variaci, zohledňují frekvenční vlivy a vysvětlují, proč variace nečiní mluvčím obtíže (Thomas, 2011: 266).

Tento model bývá aplikován na několik oblastí. Byl navržen kvůli osvojování jazyka, ale využívá se i jako model procesů u dospělých mluvčích, zejména jako mechanismus percepce řeči a normalizace, ale i jako model produkce řeči. Tato teorie se aplikuje také na lingvistické proměnné, které odkazují na sociální identitu (Thomas, 2011: 266-267).

Například osvojování jazyka vysvětluje epizodický model na základě předpokladu, že se dítě učí jazyk skrze promluvy ostatních jedinců. V raných fázích, kdy má dítě pouze omezenou zkušenost s malým množstvím exemplářů, jsou pro něj důležité tři skupiny lidí - dospělí muži, dospělé ženy a děti -, jejichž fonetické realizace promluv jsou založeny zejména na biologických rozdílech. Ve chvíli, kdy se sociální svět dítěte rozšíří, se dá předpokládat, že

⁴ Exemplar Theory.

exempláře ostatních dětí ovlivní vlastní řečovou produkci dítěte (Foulkes & Docherty, 2006: 428-430).

Epizodický model percepce řeči přináší řadu zajímavých hypotéz do sociofonetického výzkumu, avšak tyto hypotézy nebyly dosud ani explicitně vysloveny ani testovány (Hay & Drager, 2007: 97). Jedná se tedy o teorii nepotvrzenou, k jejímu potvrzení je třeba dalšího sociofonetického výzkumu (Thomas, 2011: 271). Podle Labova (2006: 512-513) jsou epizodické modely problematické, přestože nepopírá, že některé abstraktní struktury se v jisté míře komplexity v paměti uchovávají. Podle něj ale slova a zvuky nejsou sociálně stigmatizované a k zodpovězení otázek týkajících se uchovávání sociálních informací je třeba se posunout na mnohem abstraktnější úroveň.

2.6 Možnosti využití sociofonetiky

Sociofonetika může sloužit jako hodnotný zdroj pro další aplikované vědy. Variace je jedním z hlavních témat, jimiž se zabývají řečové technologie, a její poznání pomáhá zdokonalovat nejen systémy pro rozpoznání řeči, ale i její syntézu. Stejně tak přispívá sociofonetický výzkum k řečové a jazykové terapii, protože objevování variací pomáhá odhalovat normální a patologické jevy v dětské řeči a jazykovém vývoji. Přesnější určení normálních vzorců mluvených variací pak může být využito i v pedagogice. V neposlední řadě hraje sociofonetika zásadní roli ve forenzním zkoumání. Pro porovnání mluvčích je znalost variací mezi mluvčími (i variací v rámci řečových projevů mluvčího jednoho) klíčová, důležité může být i poznání procesu stárnutí (zejména pokud se mezi nahrávkou v době činu a pozdější nahrávkou podezřelého nachází delší časové rozmezí) apod. (Foulkes a kol., 2010: 736-737).

3 FORENZNÍ FONETIKA

„Forenzní fonetika je aplikací znalostí, teorií a metod obecné fonetiky na praktické úkoly, které vyplývají z kontextu práce policie nebo prezentace důkazů u soudu, stejně jako vývoje nových specificky forenzně fonetických znalostí, teorií a metod“ (Jessen, 2008: 671). Svůj oficiální status získal termín forenzní fonetika roku 1991, kdy byla založena Mezinárodní asociace pro forenzní fonetiku⁵, jež vydává časopis *International Journal of Speech, Language and the Law*, dříve nazývaný *Forensic Linguistics*. V roce 2004 rozšířila organizace svůj název na Mezinárodní asociace pro forenzní fonetiku a akustiku⁶ (Jessen, 2008: 672). Kromě ní v Evropě funguje v současnosti ještě pracovní skupina *Forensic speech and audio analysis* v rámci Evropské sítě institutů forenzních věd⁷ (Jessen & Jessen, 2008: 136).

Forenzní zkoumání hlasu má však delší historii, objevuje se již během druhé světové války, i když nejrelevantnější výzkumy se provádí až od konce padesátých a začátku šedesátých let (Hollien, 2002: 19-24). V poválečném období se moderní forenzní přístupy vyvinuly zejména v USA, Velké Británii a Německu. V USA se výzkum zaměřoval spíše na segmentální rovinu, ve Velké Británii naopak na rovinu suprasegmentální (Hollien, 2002: 70). V Německu se forenzní výzkum objevil o něco později než v prvních dvou jmenovaných zemích (Hollien, 2002: 71), jeho Bundeskriminalamt⁸ používá forenzní metody zkoumání hlasu od 70. let (Jessen & Jessen, 2008: 136).

3.1 Oblasti využití

Forenzní fonetika hraje zásadní roli ve stále větším množství soudních procesů. Centrální aspekt forenzní fonetiky představuje identifikace mluvčího, někdy označovaná jako rozpoznání mluvčího. Dalším jejím aspektem je analýza obsahu určitých úseků v řeči, jimž je těžké porozumět, obvykle nazývaná jako analýza či prozkoumání sporných promluv. Některé úkoly forenzní fonetiky jsou úzce spojeny s řečovými technologiemi či akustikou, jako třeba zlepšování kvality nahrávek, určování jejich autenticity nebo automatické rozpoznávání řeči (Jessen, 2008: 671). Forenzní fonetika se také využívá při sestavování poslechových testů

⁵ International Association for Forensic Phonetics.

⁶ International Association for Forensic Phonetics and Acoustic.

⁷ European Network of Forensic Science Institutes.

⁸ Dále BKA.

určených k identifikaci pachatele svědkem trestného činu (Butcher, 2002: 2). Někdy může být forenzní fonetik požádán i o analýzu nejazykových zvuků (Jessen & Jessen, 2008: 137).

3.1.1 Sporné promluvy

Sporné promluvy se vyskytují nejčastěji v nahrávkách se sníženou kvalitou. Forenzní fonetici mohou být požádáni o zhodnocení kvality předložené nahrávky, srozumitelnosti řeči v ní obsažené, ale i o transkripci textu (Butcher, 2002: 1-2). Snížená kvalita může být způsobena například velkou mírou hluku v pozadí či šumem telefonní linky (Baldwin & French, 1990: 59). Rušivé zvuky okolí je sice možné částečně odstranit pomocí filtrů, při těchto úpravách však společně s odstraněním šumu může dojít i ke zkreslení řečového signálu (Svobodová in Rak a kol., 2008: 463). Úspěšnost těchto technik je závislá i na povaze hluku v nahrávce, tj. především na tom, v jakých frekvenčních pásmech se nachází (Baldwin & French, 1990: 60). Přestože tedy nahrávky narušené hlukem mohou být různými technikami zlepšeny, fonetická a sociolingvistická analýza nahrávek může být v určení řečeného klíčová (Foulkes & French, 2001: 1).

3.1.2 Určování autenticity nahrávek

Forenzní fonetici mohou také rozhodovat o tom, zda bylo s nahrávkou nějakým způsobem manipulováno (Butcher, 2002: 2; Jessen & Jessen, 2008: 137), zda do ní nebyly nějaké pasáže vloženy, nebo v ní změněny, či dokonce smazány (Baldwin & French, 1990: 61; Jessen & Jessen, 2008: 137). Dříve se určování autenticity týkalo zejména kazetových nahrávek, kdy se hojně využívalo vizualizace a měření nahraných stop na elektromagnetických páskách (Jessen & Jessen, 2008: 137). V současnosti se díky dostupnému softwaru úpravy stávají snadno „neviditelnými“ (Butcher, 2002: 2), proto se autenticita nahrávky určuje zejména na základě diskontinuit v intonačních konturách, nepatřičných sekvencí intonačních vzorů, segmentálního krácení či prodloužení na neočekávaných místech a dalších neobvyklých fonetických jevů (Nolan, 1999: 15). Stejně tak se sledují i diskontinuitní jevy na jazykové úrovni, které nejsou vysvětlitelné gramatikou jazyka (Jessen & Jessen, 2008: 137).

3.1.3 Identifikace pachatele svědkem

V situacích, kdy nahrávka pachatele i očitý svědek trestného činu chybí, může být využito svědků a obětí kriminálního činu k identifikaci podezřelého podle hlasu. V tomto

případě je důležité rozlišit, zda se jedná o známý či cizí hlas vzhledem ke svědkovi či oběti (Jessen, 2008: 676; Jessen, 2010: 380; Jessen & Jessen, 2009: 30-31).

Pokud se jedná o známý hlas, zpravidla je svědek schopen identifikovat mluvčího i jmenovitě. Pokud je hlas pachatele svědkovi či oběti před trestným činem cizí, je možné provést test hlasové konfrontace, jenž je inspirovaný vizuální identifikací podezřelých (Broeders & Amelsvoort, 2001: 238; Jessen, 2008: 678; Jessen & Jessen, 2009: 31; Künzel, 1994: 136). Místo osob se však do řady staví řečové projevy, které nejsou prezentovány simultánně, ale sekvenčně (Broeders & Amelsvoort, 2001: 239; Jessen, 2008: 678). Sekvenční prezentace snižuje možnost toho, že by svědek porovnával jednotlivé lidi a vybral z nich mluvčího pachateli nejvíce podobného. V tomto typu testu je však třeba předejít tendenci vybírat mluvčí nacházející se na jeho konci. Tomu lze zamezit tím, že svědkovi nebude sdělen počet mluvčích do poslechového testu zahrnutých (Broeders & Amelsvoort, 2001: 239). Stejně jako ve vizuální identifikaci se totiž v souboru hlas podezřelého nalézá společně s dalšími hlasy mluvčích, kteří s případem nijak nesouvisí. Sestavení poslechového testu je však mnohem náročnější než sestavení jeho vizuálního ekvivalentu (Jessen, 2008: 678).

3.1.3.1 Voice line-ups⁹

Poslechové testy tohoto typu by se měly provádět až ve chvíli, kdy neexistuje dostatek důkazů proti podezřelému a ani není možné svědka s podezřelým konfrontovat vizuálně - svědek mu neviděl do obličeje (Broeders & Amelsvoort, 1999: 1373; Butcher, 2002: 2). K sestavení testu je nutné dostatečné množství reprezentativního řečového materiálu podezřelého a dostupnost dalších mluvčích (figurantů), kteří do testu budou zahrnuti. Svědek by měl také slyšet hlas pachatele v době činu po dobu minimálně 30 sekund a měl by s ním být v interakci (Broeders & Amelsvoort, 1999: 1373). Protože doba, po kterou je člověk schopný pamatovat si nový hlas, rapidně klesá, musí být svědek vyslechnut co nejdříve po kriminálním činu, aby uvedl charakteristiky hlasu pachatele. Stejně tak i poslechový test musí být proveden co nejrychleji (Jessen, 2008: 679), incident by se neměl odehrát před více než třemi měsíci před aplikací testu. Identifikační procedura by měla být organizována policií, uskutečněná trénovanými policisty a asistovaná forenzními lingvisty či fonetiky (Broeders & Amelsvoort, 1999: 1373).

⁹ Dále hlasová konfrontace.

Do poslechového testu jsou tedy kromě podezřelého zahrnuty i nahrávky hlasů figurantů. Požadavky na minimální počet těchto dalších mluvčích se mezi autory liší, zpravidla se hovoří o čtyřech až šesti dalších hlasech mimo podezřelého (např. Broeders & Amelsvoort, 1999: 1373; Broeders & Amelsvoort, 2001: 238; Jessen, 2008: 680). Tyto hlasy by se měly hlasu podezřelého svými charakteristikami podobat (Broeders & Amelsvoort, 2001: 241; Broeders & Amelsvoort, 2001: 238; Butcher, 1996: 97; Jessen, 2008: 680; Jessen, 2007b: 182; Jessen & Jessen, 2009: 31; Künzel, 1994: 136), hlasu podezřelého by měly odpovídat v pohlaví, věku, typu přízvuku či dialektu, základní frekvenci, řečovém tempu, hlasové kvalitě, vnímané úrovni dosaženého vzdělání a socioekonomického zázemí (Broeders & Amelsvoort, 1999: 1374). Mimo to jsou mluvčí zahrnutí do poslechového testu vybíráni i na základě údajů o povaze hlasu neznámého pachatele získaných svědkem (Broeders & Amelsvoort, 1999: 1373-1374; Jessen, 2007b: 181-182). Takto sestavený poslechový test je pak spolehlivější než pouhá konfrontace s hlasem podezřelého, neboť zvyšuje jeho průkazní hodnotu (Broeders & Amelsvoort, 2001:239).

Protože by jedinec v testu mohl být posluchačem vybrán na základě řečových charakteristik odlišných od zbytku mluvčích v testu, neměl by se ani podezřelý (Jessen, 2008: 680) ani nikdo z ostatních mluvčích svým hlasem výrazně odlišovat (Butcher, 1996: 97; Butcher, 2002: 2). Zároveň je však třeba věnovat pozornost tomu, aby se díky nastaveným kritériím hlasy nepodobaly až příliš (Broeders & Amelsvoort, 2001: 240).

Nahrávky v poslechovém testu by měly být interakčně podobné (Nolan, 2003: 278) a realizované za podobných podmínek (Butcher, 1996: 97). Komunikační kontext nahrávek (telefonický hovor, spontánní vs. čtený text, křik apod.) i jejich řečový styl by také měly co možná nejvíce odpovídat situaci při kriminálnímu činu. Všechny vzorky řeči by měly být přibližně stejně dlouhé a trvat déle než 20 sekund, hlasitost by měla být pro všechny hlasy stejná (Broeders & Amelsvoort, 1999: 1374).

To, jestli se výběr vzorků a nahrávání zdařily, musí být testováno na posluchačích, kteří nejsou nijak zapojeni do případu (Broeders & Amelsvoort, 1999: 1374; Jessen, 2008: 680). Tito posluchači mají obvykle se svědkem společné rysy (jazyk, dialekt, pohlaví apod.), nesmí být s ním ani s dalšími mluvčími v testu příbuzní, ani udržovat se svědkem a dalším testovacím posluchačem kontakt (Broeders & Amelsvoort, 1999: 1375). Měli by být dopředu seznámeni s charakterem trestného činu a stejně tak i s popisem řeči pachatele udané svědkem. Tito posluchači se zaměřují na to, zda se v poslechovém testu nenachází mluvčí,

kteří by se výrazněji odlišovali od ostatních (Broeders & Amelsvoort, 2001: 241-242). Pokud nějakého z nich (kromě podezřelého) kvůli jeho specifické charakteristice označí, musí být tento mluvčí z testu vyřazen nebo nahrazen někým jiným. Test je pak třeba znovu testovat na nových posluchačích, dokud všechny hlasy nebudou mít stejnou pravděpodobnost výběru (Broeders & Amelsvoort, 1999: 1375; Broeders & Amelsvoort, 2001: 242; Jessen, 2008: 680). V každém testu by se měl nacházet pouze jeden podezřelý, jednotlivé hlasy by za sebou měly následovat v náhodném pořadí (Broeders & Amelsvoort, 1999: 1373).

Ještě před samotným zahájením testu musí být svědkovi položena otázka, zda se s hlasem osoby zapojené do kriminálního činu od té doby neseťkal. Pokud svědek slyšel hlas podezřelého ve spojitosti s policií či v kontextu případu, nemůže být test dále prováděn (Broeders & Amelsvoort, 1999: 1375). Člověk, který bude přítomen testu identifikace, by neměl vědět, který hlas patří podezřelému (Broeders & Amelsvoort, 1999: 1375; Kreiman & Sidtis, 2011: 257). Svědek je pak po každém vzorku tázán, jestli se jedná o hlas, který slyšel během incidentu (Jessen, 2008: 680).

Test nesmí být prováděn dvakrát, protože by svědek mohl identifikovat nějaký z hlasů z prvního testu, a ne z kriminálního činu (Broeders & Amelsvoort, 1999: 1375; Jessen, 2008: 681; Jessen & Jessen, 2009: 31). Na svědkovu žádost je ale možné jeden nebo více hlasů v průběhu testu zopakovat (Broeders & Amelsvoort, 1999: 1375).

V případě existence více svědků je třeba zabránit jejich kontaktu během identifikační procedury. Stejně tak nesmí být informováni o výsledku procedury provedené s jiným svědkem. Každému svědkovi musí asistovat jiný policista, který nebyl přítomen předchozí identifikační proceduře. Všichni svědci musí být dotazováni, zda byli v kontaktu s jiným svědkem, který podstoupil poslechový test před nimi. Pokud ano, je nutné zjistit, jakou informaci obdrželi (Broeders & Amelsvoort, 1999: 1375-1376).

3.1.3.2 Problémy naivní identifikace mluvčího

Hodnotit svědectví posluchače je náročné, protože schopnost naivních posluchačů identifikovat či rozlišit hlasy je ovlivněna mnoha různými faktory, o jejichž interakci je známo jen velmi málo (Rose, 2002). Spolehlivost identifikace pachatele svědkem ovlivňují zejména tři faktory - komunikační kanál, mluvčí a posluchač. Komunikační kanál může být omezen například vzdáleností mezi mluvčím a posluchačem, hlukem a různými šumy či telefonickým zařízením (Jessen, 2008: 677). Schopnost rozpoznání důvěrně známých hlasů

v telefonním hovoru testovali Foulkes a Barron (2000). V experimentu bylo zjištěno, že i přes blízký vztah a častost kontaktu nebyli participanti vždy schopni hlasy přiřadit ke správnému mluvčímu, v jednom případě dokonce posluchač nepoznal svůj vlastní hlas. Ve spolehlivosti identifikace se může odrazit i délka pronesené promluvy nebo například křik (Jessen, 2008: 677). Křikem ve spojitosti s délkou promluvy se zabývali Blatchford a Foulkes (2006), kteří mimo jiné také potvrdili, že délka promluvy představuje významnou proměnnou v úspěšné identifikaci mluvčího. Stejně tak se doba trvání hlasového projevu mluvčího jako jedna z klíčových proměnných správné identifikace objevila i ve výzkumu, který prováděli Orchard a Yarmey (1995).

Druhý důležitý faktor spolehlivosti identifikace představují mluvčí, protože někteří bývají snadněji identifikováni než jiní (Blatchford & Foulkes, 2006: 252; Foulkes & Barron, 2000: 195; Jessen, 2008: 678; Rose, 2002). Hlasy mohou obsahovat různé množství distinktivních rysů, které je umožňují identifikovat od zbytku populace (Hollien, 2002: 46-47). Těmito distinktivními rysy často bývají výrazné sociolingvistické a regionální jevy (Blatchford & Foulkes, 2006: 251), hodnota základní frekvence (Foulkes & Barron, 2000: 194; Lavner, Gath, & Rosenhouse, 2000: 25) či řečové patologie (Foulkes & Barron, 2000: 189; Hollien, 2002: 46-47). Přesto zatím existuje jen poměrně málo informací o spojení mezi akustickými charakteristikami hlasu mluvčího a procesy jeho identifikace posluchačem. Jako jeden ze slibných přístupů, jak tato spojení objevit, se jeví analýza a resyntéza řečového signálu (Lavner a kol., 2000: 9).

Další ze základních problémů identifikace mluvčího svědky trestného činu je to, že se lidé mezi sebou ve schopnosti identifikovat hlasy liší (Blatchford & Foulkes, 2006: 252; Jessen, 2008: 678; Kreiman & Sidtis, 2011: 251; Rose, 2002; Watt, 2009: 10). Posluchači mohou mít snížené sluchové schopnosti, důležitou roli hraje jejich paměť (Jessen, 2008: 677-678), ale i další faktory jako věk (Blatchford & Foulkes, 2006: 243; Eriksson, Sullivan, Zetterholm, Czigler, Skagerstrand, & van Doorn, 2010: 40), doba mezi prvním poslechem hlasu a identifikační procedurou a povaha a rozsah maskování hlasu (Blatchford & Foulkes, 2006: 243). Významnou roli hraje i aktuální emoční stav mluvčího i jazyk promluvy (Hollien, 2002: 46; Rose, 2002). V neposlední řadě se na míře úspěšnosti rozpoznání mluvčího podílí i stupeň obeznamení posluchače s jeho hlasem (Blatchford & Foulkes, 2006: 243; Rose, 2002; Yarmey, Yarmey, Yarmey, & Parliament, 2001: 296) a míra pozornosti, jakou svědek hlasu pachatele v době incidentu věnoval (Kreiman & Sidtis, 2011: 252).

Při sestavování testu hlasové konfrontace je třeba pamatovat ještě na několik věcí. Zetterholm, Sarwar a Allwood (2009) ve svém experimentu poukázali na několik tendencí posluchačů, které by se v testu mohly objevit. Posluchači zpravidla tíhnou k tomu nevybírat prvního mluvčího z řady, ale spíše volit hlasy nacházející se uprostřed či ten úplně poslední. Ukazují také, že posluchači kladou důraz na pouhou podobu některých řečových charakteristik (např. podobná F_0^{10}), které pachatele s náhodným mluvčím spojují, a že míra jistoty není spolehlivým ukazatelem správnosti identifikace.

3.1.4 Rozpoznání mluvčího

Ve forenzním výzkumu je třeba odlišovat naivní a technické rozpoznávání mluvčího. Naivní rozpoznávání mluvčího je prováděno posluchači bez lingvistického či fonetického tréninku, jejichž úsudek je založen na sluchovém dojmu bez použití dalších speciálních technik (Nolan, 1999: 1; Watt, 2009: 1). Technické rozpoznávání mluvčích vykonávají trénovaní fonetici, kteří mají k dispozici analytické metody jako přepis nahrávky prostřednictvím fonetické transkripce (IPA) nebo akustickou analýzu za pomoci softwaru na zpracování řeči (Watt, 2009: 2).

Pod kategorií rozpoznání mluvčího spadají dva typy úkolů, a to verifikace a identifikace mluvčího. Nejdůležitější rozdíl mezi nimi spočívá v tom, že při verifikaci je k dispozici známý a uzavřený soubor mluvčích (Rose, 2002). Navíc mluvčí, jehož identita má být potvrzena, většinou spolupracuje. Identifikace mluvčího se používá zejména v případech s kriminálním obsahem (Hollien, 2002: 5-6; Nolan, 1999: 1), kdy obvykle uzavřený soubor podezřelých není k dispozici (Rose, 2002). Může se jednat o špionáž, vydírání, hrozby a výhrůžky, podezřelé teroristické komunikace i sexuální harašení (Rodman, McAllister, Bitzer, Cepeda, & Abbitt, 2002: 24). Jessen (2008: 673) identifikaci mluvčího rozčleňuje na tři podoblasti - porovnávání hlasu, profilování hlasu a analýzu identifikace mluvčího oběťmi a svědky.

Většina práce forenzních fonetiků se týká profilování a porovnávání mluvčích. Profilování se provádí v případě, kdy není k dispozici žádný podezřelý. Jeho účelem je snížit populaci možných podezřelých identifikováním lingvistických rysů typických pro určité geografické a sociální skupiny a neobvyklých výslovností, jež mohou odkazovat na výjimečné anatomické či patologické charakteristiky mluvčího (Watt, 2009: 12).

¹⁰ Základní frekvence.

Nejběžnějším úkolem forenzního fonetika však bývá srovnání dvou a více vzorků neznámého hlasu (zpravidla pachatele) s jedním nebo více vzorky hlasu známého (zpravidla podezřelých). Úkolem fonetika je pak určit, zda vzorky pochází od stejné osoby (Nolan, 1999: 2; Rose, 2002). Někdy se stává, že forenzní případy začínají profilováním mluvčího a po nalezení podezřelého končí porovnáváním (Jessen, 2010: 379).

3.1.4.1 Profilování hlasu

Profilování hlasu a klasifikace mluvčího se provádí na základě nahrávky pachatele ve chvíli, kdy neexistuje podezřelý. Často se využívá v počáteční fázi policejního vyšetřování, kdy hlasový projev může být důležitým vodítkem k určení identity pachatele (Baldwin & French, 1990: 64; Foulkes & French, 2001: 1; Jessen, 2008: 673; Jessen, 2007b: 180-181; Jessen, 2010: 379). Expert je požádán o vytvoření profilu mluvčího, který bude pomáhat při snaze o jeho dopadení. Profil může obsahovat informace o regionu, v němž mluvčí vyrostl, jeho věku, pohlaví, úrovni vzdělání, sociálním zázemí, mateřském jazyku, zdravotním stavu ovlivňujícím jeho hlas apod. (Jessen, 2008: 673-674). Informace získané z řečového materiálu se mohou lišit v závislosti na délce a kvalitě nahrávky, na tom, zda a jak pachatel maskuje svůj hlas, i na míře nářečních a sociolingvistických informací obsažených v nahrávkách. Utvoření profilu mluvčího pomáhá definovat cílovou populaci, a tím zúžit okruh hledání podezřelých (Foulkes & French, 2001: 1). Někdy se jako alternativa nebo souběžně s touto analýzou hlasu může dát nahrávka pachatele do masmédií, například rádia či televize (Jessen, 2007b: 181; Jessen & Jessen, 2008: 137; Jessen, 2010: 379).

Podle Jessena (2008: 674; 2010: 381) je vhodné termíny profilování hlasu a klasifikace mluvčího oddělovat. Profilování hlasu je podle něj praktičtější zaměřené - zpravidla na poskytování informací o mluvčím za účelem dopadení podezřelého. Klasifikace mluvčího je teoretičtější, je to snaha o zařazení mluvčího do kategorií obecně (věk, pohlaví, sociální skupina, region apod.). Přestože profilování hlasu spadá do kategorie klasifikace mluvčího, obsahuje ještě identifikaci i méně charakteristických rysů řeči mluvčího sloužících k jeho nalezení. Klasifikace mluvčího se oproti tomu využívá i pro účely konstruování testů hlasové konfrontace.

3.1.4.2 Porovnávání hlasů

Jakmile jsou k dispozici nahrávky pachatele i podezřelého (příp. podezřelých)¹¹, může být uskutečněno porovnávání hlasu. To je založeno na podrobném srovnání obou řečových vzorků za účelem odhadu pravděpodobnosti, zda byly produkovány stejnou osobou (Foulkes & French, 2001: 2; Watt, 2009: 14). Výsledkem analýzy tak není tvrzení o ne/totožnosti mluvčích na obou nahrávkách, ale názor pohybující se na škále míry jistoty (Foulkes & French, 2001: 2). Toto porovnávání bývá také jinak nazýváno jako identifikace mluvčího expertem či technická identifikace mluvčího, čímž se odlišuje od naivní identifikace mluvčího svědkem (Jessen, 2008: 681).

Od podezřelého by se mělo zajistit více typů řečových vzorků, ideálně čtenou a improvizovanou promluvu. Čtený text je vhodné získat jak zopakováním některých pasáží z důkazního materiálu, tak z běžného textu, například novinového článku. Tyto úkoly je pak vhodné pro jistotu zopakovat (Hollien, 2002: 77).

Pro porovnání hlasů jsou důležité dva aspekty, a to podobnost hlasů a typičnost jejich charakteristik. Některé charakteristiky analyzované v řeči neznámého mluvčího a podezřelého mohou být totiž poměrně typické pro celou populaci, zatímco jiné mohou být naopak vzácné (na což je třeba brát ohledy). Koncepce pracující s podobností a typičností se nazývá Bayesovský přístup, jeho jádrovým konceptem je koeficient pravděpodobnosti (Butcher, 2002: 13; Jessen, 2008: 682), který kvantifikuje sílu důkazního materiálu (Rose, 2002). Pokud je koeficient pravděpodobnosti vyšší než jedna, vzrůstá pravděpodobnost, že jsou oba hlasy ze stejného zdroje, pokud je koeficient menší než jedna, tak naopak (Butcher, 2002: 14; Jessen, 2008: 683). Bayesovský přístup se široce akceptuje ve forenzních vědách od devadesátých let, ve forenzní analýze řeči se využívá zejména v automatickém rozpoznávání (Jessen, 2008: 683).

3.1.4.2.1 Omezení při porovnávání hlasu

Jeden z problémů při porovnávání dvou (a více) hlasů představuje variace v rámci řeči jednoho mluvčího, jejíž extrémní formu ztělesňuje snaha o maskování hlasu. Omezení plynou i z technického pozadí, jedna z nahrávek často bývá pořízena během telefonního hovoru,

¹¹ Svobodová (in Rak a kol., 2008: 456) nazývá tyto dva typy nahrávek jako sporné (anonymní) a srovnávací (odborně pořízené).

v němž je signál omezen na pásmo zhruba od 200 do 3500 Hz. To znamená, že charakteristiky mluvího objevující se mimo tento rozsah nemohou být využity (Jessen, 2008: 684-685). V telefonním hovoru jsou vynechány komponenty řeči s nízkou frekvencí, jako základní frekvence nebo i první formanty některých řečových zvuků. Stejně tak některé specifické rysy mluvího jako zvláštní či defektivní výslovnost vysokofrekvenčních obstruentů se může stát nerozpoznatelnou (Künzel, 1994: 136). V kontextu telefonních hovorů může často docházet i k vysokému stupni narušení signálu okolním hlukem, zejména pokud je užito telefonů na veřejných místech a rušných ulicích. Některé řečové charakteristiky se zdají být dokonce přímo ovlivněny faktem, že je ke komunikaci využito telefonního spojení. Studie naznačují, že mluvíci mají tendenci do telefonu mluvit hlasitěji a zvyšovat svoji základní frekvenci (Foulkes & Barron, 2000: 181-182). Za tohoto předpokladu pak vyšší nebo stejná základní frekvence v nahrávce pořízené v běžné konverzaci může vést k vyloučení podezřelého (Rose, 2002).

Další problémy při porovnávání mluvích se objevují, pokud je řeč narušená ozvěnou, hlukem nebo překrývajícími se hlasy. Stejně tak délka nahrávky ovlivňuje výsledné rozhodnutí. Čím je nahrávka kratší, tím méně specifických řečových vzorců mluvího vykazuje. Přestože se neuvádí žádná minimální hranice, měla by nahrávka obsahovat minimálně osm sekund řeči od anonymního mluvího a dvojnásobek od podezřelého (Jessen, 2008: 685-686).

Komplikované může být i získání řečového materiálu od podezřelých, protože ti vždy nemusí být ochotni spolupracovat. Často se stává, že podezřelí i na otázky ohledně svého jména, data narození či adresy odpovídají „Bez komentáře“. Opakování této fráze pak nemůže poskytnout dostatek řečového materiálu k porovnání hlasu podezřelého s hlasem pachatele (Baldwin & French, 1990: 74). V těchto případech závisí získání materiálu na právním systému daného státu a na tom, zda mohou být k účelu porovnání použity nahrávky z policejních rozhovorů či telefonních odposlechů (Jessen, 2010: 379).

Občas se vyskytují pochybnosti, zda jsou forenzní fonetici ve skutečnosti schopni rozpoznat mluvího či nějaké z charakteristik jeho hlasu lépe než naivní posluchači. Na toto téma bylo provedeno několik výzkumů, například Köster, Jessen, Khairi a Eckert (2007) se zabývali tím, zda se forenzní experti a laici budou lišit v rozpoznávání hlasové kvality mluvího. Přestože všichni participanti tohoto experimentu byli seznámeni se zkoumanými typy hlasové kvality, výsledky poukázaly na signifikantně lepší skóre forezních

fonetiků oproti naivním posluchačům. I mezi naivními posluchači se však našlo několik participantů, kteří měli výsledky srovnatelné s fonetiky. To naznačuje, že i talent a percepční schopnosti hrají určitou roli v identifikaci hlasové kvality.

3.2 Používané metody

Forenzní fonetika je tradičně založena spíše na poslechových než na akustických metodách. Podle Jessena (2007b: 191) existují kromě jiných důvodů i tři praktické aspekty této skutečnosti. Analýzu hlasu bývá totiž často potřeba provést velice rychle kvůli informacím, které mohou pomoci ve vyšetřování. Předběžný profil mluvčího je často očekáván již v průběhu jednoho dne a akustická analýza může být časově nákladnější než analýza poslechová. Někdy se také požaduje uskutečnění analýzy mimo laboratoř na nějakém konkrétním místě - ne vždy je možné brát všechno vybavení s sebou. Nezřídka je také kvalita materiálu k analýze snižena a pak je jednodušší z něj informace získat cestou poslechové analýzy.

V současnosti se při analýze hlasu většinou upřednostňuje kombinace akustických a poslechových metod (Nolan, 1999: 14; Rose, 2002; Watt, 2009: 14), kdy je opakované a pozorné poslouchání doprovázeno detailním instrumentálním zkoumáním pomocí akustického fonetického softwaru (Watt, 2009: 14). Poslechová analýza je totiž nejefektivnější v odhalování podobností a rozdílů vyplývajících z jazykového systému, zatímco akustická analýza poskytuje kvantitativní vyjádření slyšeného a může objevit rozdíly, jež z poslechu nejsou patrné (Nolan, 1999: 14).

3.2.1 Poslechová analýza

Poslechová analýza zahrnuje pozorný a opakovaný poslech se zaznamenáváním charakteristik hlasu dotyčného (Butcher, 2002: 4). Její první krok zpravidla představuje transkripce segmentálních rysů a zaznamenání hezitací, dysfluencí, nestandardní gramatiky, neobvyklých slov apod. (Watt, 2009: 14). Poslechová analýza se zabývá především jazykovými jevy a měla by poskytovat shrnutí údajů o mluvčím na základě jazykového systému a způsobu jeho realizace (Rose, 2002). Jazykově-fonetická transkripce a deskripce je důležitá pro klasifikaci mluvčího, zejména v identifikování dialektů, sociolektu či cizineckého přízvuku. Využívá se i při popisu hlasové kvality a hezitačních projevů (Jessen, 2008: 688-690).

Butcher (2002: 4-5) rozděluje zkoumané rysy na základě poslechu do čtyř základních kategorií: a) rysy kvality hlasu, b) nelingvistické charakteristiky řeči, které nejsou produkovány hrtanem (artikulační nastavení), c) aspekty artikulace spojené s geografickým a sociálním zázemím mluvčího a d) identifikace osobitých výslovnostních rysů (řečové vady a různé druhy dysfluencí). V českém kontextu se využívá skupin jevů, mezi něž patří kvalita hlasu a způsob dýchání, způsob artikulace hlásek a hláskových skupin, prozodie (frázování, melodie, přízvuk, pauzy, tempo řeči, rytmus, dynamika) a celková úroveň projevu (Svobodová in Rak a kol., 2008: 464).

3.2.2 Akustická analýza

Jako doplnění analýzy poslechové se používá analýza akustická, protože umožňuje kvantifikaci slyšených aspektů řeči spojených s mluvčím (Nolan, 1999: 12). Akustická analýza se v první řadě zabývá základní frekvencí hlasu mluvčího, dlouhodobým průměrným spektrem¹² a frekvencí vokálních formantů (Butcher, 2002: 5-8). Její pomocí se ale také měří artikulační tempo, kdy se využívá kombinace poslechového a akustického zpracování (Jessen, 2008: 693).

Průměrná základní frekvence hlasu mluvčího získaná akustickou analýzou bývá často jedním ze základních faktorů porovnávání více mluvčích. Ve dvou promluvách jednoho mluvčího se rozmezí průměrné základní frekvence většinou neliší o více než 5 Hz. V případě, že hodnota průměrné základní frekvence pachatele i podezřelého je výrazně nižší nebo vyšší, než je průměr v populaci, zvyšuje se pravděpodobnost, že oba vzorky pochází od stejného mluvčího. Stejně se akustické analýzy využívá i ke zjištění standardní odchylky základní frekvence od průměru. Tato odchylka indikuje, do jaké míry je řeč mluvčího monotónní (Baldwin & French, 1990: 46-48). Právě to, že mnoho jedinců používá jisté charakteristické struktury a vzorce, může usnadnit identifikaci pachatele (Hollien, 2002: 60).

Některým informacím přidává akustická analýza přesnost a objektivitu. Své využití nachází i v případech, kdy se objevuje nesoulad mezi základní frekvencí hlasu mluvčího a hodnotami formantů jeho vokálů. Ty se objevují nejen ve snahách o maskování hlasu, ale například i u transsexuálů podstoupivších hormonální léčení a operaci hlasivek. Délka vokálního traktu totiž operačně modifikovaná být nemůže, proto její akustické koreláty odhalí původní maskulinní identitu mluvčího (Jessen, 2007b: 191-195).

¹² LTAS (z long-term average spectrum).

Akustická analýza může také odhalit jevy, které se poslechově zdají totožné, ačkoli tomu tak objektivně není. Například samohlásky znějící stejně mohou mít naprosto odlišné formantové vzorce. Sluchové ústrojí totiž může integrovat dva formanty ležící ve vzájemné blízkosti a zacházet s nimi jako s jedním formantovým vrcholem (Nolan, 1999: 12).

3.2.3 Voiceprint

V metodách identifikace mluvčího je třeba odlišovat tzv. otisk hlasu¹³. Na konci šedesátých a začátku sedmdesátých let se tato technika stala hojně využívanou, zejména v USA, některých zemích Evropy a v Izraeli. Mělo se jednat o objektivní metodu forenzního porovnávání hlasu na základě vizuální interpretace spektrogramů (Künzel, 1994: 138). Technika otisku hlasu měla být analogická otiskům prstů nebo analýze DNA. Tato metoda však nemůže být využívána jako adekvátní technika forenzního rozpoznávání mluvčího (Baldwin & French, 1990: 53; Butcher, 2002: 3; Künzel, 1994: 139), a to z několika důvodů. Jedním z nich představuje fakt, že neexistuje rys hlasu, který by byl pro každého mluvčího jedinečný. Navíc většina měřitelných charakteristik hlasu může být mluvčím vědomě měněna (např. základní frekvence, aspekty kvality hlasu, artikulace konsonantů i vokálů). Pro většinu hlasových charakteristik navíc neexistují data, která by popisovala jejich rozvržení v populaci. Zároveň se akustické parametry mění v závislosti na podmínkách nahrávání i na změnách hlasu samotného (Butcher, 2002: 3). Z těchto důvodů mohou dva spektrogramy stejného slova vyslovené stejným mluvčím vypadat naprosto odlišně a zároveň spektrogramy stejného slova vyslovené rozdílnými mluvčími si mohou být velmi podobné (Baldwin & French, 1990: 55).

3.2.4 Automatické rozpoznávání

Automatické rozpoznávání řeči mimo forenzní výzkum sahá až do šedesátých let. Bylo vyvinuto za účelem omezení přístupu k citlivým informacím a chráněným objektům a jiným biometrickým aplikacím. Tento způsob rozpoznávání řeči je jednodušší než rozpoznávání řeči pro forenzní účely, které se rozvinulo přiblížením se tří faktorů: zefektivněné metody obecně používaného automatického rozpoznávání, pokročilého Bayesovského způsobu myšlení a lepší počítačové technologie (Jessen, 2008: 698-699).

Automatické rozpoznání mluvčího bývá členěno na verifikaci a identifikaci. Verifikace potvrzuje či vyvrací identitu mluvčího a využívá se nejčastěji v komerčních službách (Hennebert in Rak a kol., 2008: 471). Při automatické verifikaci ve

¹³ Někdy do češtiny překládáno jako fonoskopie, autorka považuje tento název za zavádějící.

své nejjednodušší verzi se nahraje promluva či fráze mluvčího, který tvrdí, že je určitou osobou, a porovná se s uloženou (referenční) verzí projevu této osoby. Verifikace představuje spíše omezený úkol než identifikaci, protože je pro srovnání k dispozici ekvivalentní jazykový materiál (Nolan, 1999: 9). V automatické textově závislé verifikaci mluvčího se promluvy porovnávají s parametry, jako je například kontura základní frekvence, sekvence lineárně predikčních koeficientů či keprální koeficienty (Nolan, 1999: 9).

Automatická identifikace mluvčího se obvykle snaží spojit neznámý hlas s hlasem známým z uzavřené množiny mluvčích. Používá se zejména v oblastech odposlechu (Hennebert in Rak a kol., 2008: 471). Tato identifikace mluvčího může být provedena třemi způsoby. První metoda užívá dlouhodobé průměry akustických jevů. Nevýhodou této metody je, že vyřazuje některá data užitečná k rozlišení mluvčích a že může požadovat dlouhé řečové promluvy pro stabilní statistiku. Druhý přístup využívá neuronové sítě, jeho nevýhodou je požadavek na velké množství dat k „vytrénování“ modelů mluvčího. Třetí přístup představuje segmentační metodu, která srovnává mluvčí na základě podobných promluv nebo podobných fonetických sekvencí (Rodman a kol., 2002: 25).

Systémy rozpoznání mluvčího je možné dále dělit na textově závislé, systémy s textovou výzvou a textově nezávislé. Textově závislé systémy využívají stejnou promluvu jak pro registraci, tak pro autentizaci mluvčího. Systémy s textovou výzvou jsou založeny na tom, že sekvence slov, jež mají být při autentizaci vysloveny, nejsou mluvčímu předem známy. K potvrzení uživatele systém požaduje, aby mluvčí vyslovil určité sekvence slov, což zajišťuje lepší úroveň bezpečnosti než systémy předchozí. Výhodou je také to, že si uživatel nemusí pamatovat určité heslo či text. Textově nezávislé systémy k rozpoznání nepoužívají konkrétní texty, ale spíše řečový signál z globálního pohledu. Tento druh systémů však není tak přesný jako systémy předchozí, protože množství registračních dat je samozřejmě limitované (Hennebert in Rak a kol., 2008: 476-478). Účelům forenzní fonetiky jsou textově nezávislé systémy bližší, protože nelze předpokládat, že obě nahrávky určené k porovnávání budou obsahovat dostatek stejných slov a frází (Rose, 2002).

Nevýhodu automatického rozpoznávání představuje střední přesnost oproti jiným biometrickým metodám. To je způsobené zejména variabilitou řeči jednoho mluvčího, ale i variabilitou mezi jednotlivými mluvčími (někteří lidé mají podobné řečové charakteristiky). Stejně tak šum okolí může automatické rozpoznávání komplikovat. Další nevýhodou je i častá

nutnost opakovaných registračních sezení, aby registrace mluvího byla dostatečně dlouhá pro potřeby rozpoznání (Hennebert in Rak a kol., 2008: 472).

Užití automatického rozpoznávání mluvího je v typických forenzních situacích omezeno. Často totiž bývá využíváno srovnávání souborů nahrávaných přes telefon nebo odposlouchávací zařízení s nahrávkami pořízenými ve výslechové místnosti (Butcher, 2002: 3).

3.2.5 Komerční versus forezní rozpoznání mluvího

Mezi komerčním a forezním rozpoznáváním řeči existuje několik rozdílů. V komerčních systémech si mluvíci zpravidla přeje být rozpoznán, zatímco ve forezní sféře tomu bývá přesně naopak - podezřelé osoby často odmítají spolupracovat a nechtějí produkovat žádný řečový vzorek. Komerční systémy mají také jistou výhodu v omezeném počtu možných slov a frází, které mohou být vybrány tak, aby obsahovaly hlásky obsahující vysoce specifické informace o mluvčím, jako například nosové konsonanty. Ve forezním případě naopak neexistuje „předvybraný“ text, který bude produkován, aby mohl být využit pro srovnání slov. Ve většině forenzních případů se pracuje s materiálem získaným prostřednictvím telefonního hovoru, v komerčních systémech bývá akustické prostředí pozorně kontrolováno. Velké množství komerčních systémů je založeno na konečném množství promluv, zatímco v typických forenzních situacích se nedá předpokládat, zda je pachatel mezi dvěma či dvaceti podezřelými. Pokud komerční systém rozpoznávání řeči selže, nemá to zpravidla tak fatální důsledky jako v případě forezního rozpoznávání řeči (Künzel, 1994: 135-136).

3.3 Variabilita, klasifikace mluvího

Klasifikace mluvího ve forezní fonetice je zpravidla spojena se dvěma různými praktickými úkoly, profilováním mluvího a porovnáváním hlasů. Jak již bylo řečeno, rozdíl mezi nimi se nachází v ne/dostupnosti nahraného řečového materiálu a ne/dostupnosti osoby podezřelé ze spáchání trestného činu (Jessen, 2007b: 180).

Na tomto místě bych ráda zmínila dvě věci, jež s klasifikací mluvích souvisí, a to jednak zdroje variability mluvích, jednak klasifikační charakteristiky mluvího používané v současné forezní praxi. Zdroje variability mluvích bývají v literatuře nejčastěji spojovány

s porovnáváním dvou nahrávek, zatímco klasifikační charakteristiky spíše s profilováním hlasu (např. Jessen & Jessen, 2008).

3.3.1 Zdroje variability mluvcích

Je zřejmé, že se mezi mluvčími musí vyskytovat určitá variabilita, aby od sebe mohli být podle hlasu rozeznáni. Existuje několik zdrojů této variability, některé odlišnosti jsou dány anatomicky, některé jsou spojeny s funkcí vokálního traktu, některé mohou být patologické a některé dané sociolektem (Watt, 2009: 5-6).

Výslednou řeč člověka ovlivňuje několik biologicky daných vlastností vokálního traktu. Pravděpodobně největší vliv má délka a tvar vokálního traktu a frekvenční rozsah kmitání hlasivek (způsobený jejich délkou a tloušťkou). Tyto faktory se nejvíce podílejí na základní frekvenci a frekvenčním rozsahu formantů mluvcího (Jessen & Jessen, 2008: 139; Nolan, 1999: 3). Ve forenzním výzkumu je také důležitá nosní dutina, protože její tvar a velikost se mezi jednotlivými mluvčími poměrně liší a zároveň je omezena možnost jejího ovlivnění vlastní vůlí (Amino & Arai, 2009: 27; Nolan, 1999: 5; Rose, 2002). Na produkci nazál se podílí i činnost véla, jež je také pro každého mluvcího specifická a stejně jako nosní dutina nemůže být mluvcím libovolně ovládána (Amino & Arai, 2009: 27). Na druhou stranu rozpoznání mluvcího na základě vlastností nosní dutiny může být komplikované. Rezonanční vlivy nosní dutiny totiž nejsou izolovaným jevem, ale kombinují se s dalšími vlastnostmi vokálního traktu. Zároveň jsou tyto rezonance nejvíce ovlivnitelné zdravotním stavem mluvcího (Nolan, 1999: 5). Přesto je možné nazály považovat za forenzně relevantní (Amino & Arai, 2009: 27; Rose, 2002), zejména pak koronální /n/ a /ŋ/ (Amino & Arai, 2009: 23). Některé zvuky mohou být ovlivněny i dalšími charakteristikami, například výslovnost hlásky /s/ velikostí a tvarem zubů (Nolan, 1999: 5). K biologicky podmíněným faktorům lze přiřadit ještě kvalitu hlasu. K ní přispívají nejen organické komponenty (anatomie a fyziologie vokálního traktu, délka vokálního traktu, nazální dutina), ale i artikulační nastavení, to znamená habituální¹⁴ nastavení, která jedinci při řeči používají (Rose, 2002).

Výslovnost mluvcího se liší zpravidla také v závislosti na jeho sociálním a geografickém zázemí, mezi jednotlivými mluvčími existuje sociolingvistická variace (blíže viz předchozí kap.). K této „osvojené“ kategorii lze přidat ještě jazykové zvláštnosti například v podobě výplňkových, v řeči často se opakujících slov (Jessen & Jessen, 2008: 139).

¹⁴ Ustálená.

Jessen a Jessen (2009: 30) odlišují jako zdroje variability v řeči mluvčích ještě habituální zvyky, které nejsou spojené s gramatickým systémem jazyka. Do nich řadí mluvní tempo, variace v rámci řečové melodie či pauzy a jejich častost, trvání a umístování a jejich případné vyplnění hezitačními zvuky. Mluvčí se liší i ve frekvenci použitých hezitací v projevu a v kontextu jejich použití. Pauzy mohou být vyplněny i slyšitelnými nádechy, které také spadají do kategorie habituálního řečového jednání.

Mluvčí se nemusí lišit jen mezi sebou, variace existuje i v rámci promluv jednoho mluvčího. Jeden ze zásadních zdrojů variace v řeči jednoho mluvčího představuje jeho emoční stav (Hansen & Patil, 2007: 110; Jessen & Jessen, 2008: 32). Například stres může být v řeči indikován koktáním, opakováním či přeříkáváním, může vést ke zvýšení počtu nádechů a následně tak působit i na tempo řeči. Stres může také způsobit vzestup základní frekvence a ovlivnit systém produkce řeči ve výběru slov, gramatické a větné struktury i v produkci fonémů (Hansen & Patil, 2007: 110-112). Nejvyšší odchylky v základní frekvenci a intenzitě oproti neutrálnímu stavu mluvčího vykazuje z emočních reakcí hněv (Hansen & Patil, 2007: 114-117; Rose, 2002). Emoce mohou být vyvolány nespočtem faktorů, bylo dokázáno, že například i lhaní vede ke zvýšení základní frekvence, rozšíření jejího rejstříku a vzniku krátkých pauz v řeči (Anolli & Ciceri, 1997: 278). Celkově lze říci, že prudké emoční reakce mají zpravidla za následek nápadné změny hlasu, často se objevují různé fonační typy nebo široký frekvenční rozsah v promluvě (Rose, 2002).

I typ promluvy se odráží ve výsledné podobě řeči mluvčího, třeba v podobě artikulačního tempa. Tyto jevy zkoumal Jessen (2007a) ve třech typech promluv - čtené, spontánní v přímé interakci a přes telefon. Pro čtenou řeč bylo artikulační tempo signifikantně vyšší a směrodatná odchylka signifikantně menší než pro řeč spontánní. Pro forenzní potřeby je tyto rozdíly dobré znát, protože tyto znalosti mohou být použity při porovnávání nahrávek pocházejících z odlišných zdrojů (čtený/spontánní text). Přesto je dobré mít na paměti, že se mezi mluvčími objevují výjimky, které tento trend nepotvrzují. Mimo to se na variabilitě v projevech jednoho mluvčího může podílet řada dalších faktorů, například i užití návykové látky (Foulkes & French, 2001: 2) či aktuální zdravotní stav mluvčího (Nolan, 1999: 5).

Variabilita v rámci určité vlastnosti řeči jednoho mluvčího se může překrývat s variabilitou mezi dvěma různými mluvčími. Není možné určit hodnotu rozsahu rozdílů mezi vzorky, které by vždy odlišily stejné a odlišné mluvčí. Hlasy jednotlivců se liší v závislosti na

situaci, což může vést k tomu, že za určitých okolností mohou být dvě promluvy od jednoho mluvčího odlišné či naopak dvě promluvy od různých mluvčích velmi podobné (Rose, 2002).

3.3.2 Klasifikační charakteristiky mluvčího v současné forenzní praxi

3.3.2.1 Pohlaví

Určit pohlaví mluvčího bývá zpravidla poměrně jednoduchý úkol. Jeho nejvýraznějším fonetickým korelátorem je základní frekvence hlasu mluvčího. Objevují se však i případy, kdy ze základní frekvence pohlaví mluvčího není možné jednoznačně určit. Tyto situace se zpravidla objevují při maskování hlasu, kdy základní frekvence chybí (šepot) nebo se objevují neobvyklé fonační typy (např. falzet). Další problémy mohou vyvstat, pokud je základní frekvence mluvčího neobyčejně vysoká (v případě muže) či nízká (v případě ženy) (Jessen, 2007b: 183-184; Jessen & Jessen, 2008: 137). Jeden z takových nejasných příkladů z forenzní praxe uvádějí Jessen a Jessen (2008: 137). Jednalo se o sérii vyhrožování bombovými útoky na Düsseldorfském letišti v září 2003. Pohlaví volajícího nebylo jednoznačné, přestože se zdálo, že se jedná o ženu maskující svůj hlas. Tato domněnka byla nakonec potvrzena, volajícím byla žena, která zaokrouhlováním rtů svůj hlas posunula do blízkosti mužských frekvencí¹⁵.

V těchto nejednoznačných případech se zpravidla k identifikaci pohlaví mluvčího používají formantové frekvence, další vodítka může poskytovat morfologie, syntax, lexikologie i pragmatika (Jessen, 2007b: 184).

3.3.2.2 Věk

Podle hlasu není zpravidla obtížné rozdělit mluvčí do čtyř kategorií - dětství, dospívání, dospělost a stáří. Toto dělení však zpravidla nebývá forenzně relevantní, protože většinu trestných činů páchají dospělí lidé (Jessen & Jessen, 2008: 137). Pachateli zpravidla bývají lidé ve věku mezi 20 a 40 lety, někdy do 50 let, ale jen zřídka starší. Většina změn v řečových vzorcích se však objevuje právě mimo tyto věkové kategorie (Jessen, 2010: 383). Určit přesný věk mluvčího tak bývá obvykle složitější, nejedná se totiž o rozhodnutí v rámci binárních kategorií a neexistuje pro něj jednotné vodítko. Navíc biologické stáří a skutečný věk si nemusí odpovídat, záleží na tom, jakým způsobem mluvčí o svoje hlasové ústrojí

¹⁵ Zaokrouhlením rtů dojde k prodloužení vokálního traktu – délka vokálního traktu je jednou z charakteristik, jež ovlivňují rozpoznání mužských a ženských hlasů.

pečují (Jessen, 2007b: 184-185). Stárnutí hlasového ústrojí je ovlivněno cigaretami, alkoholem, častým hlasitým necvičeným mluvením, ale i různými chorobami (Jessen & Jessen, 2008: 137). Klasifikace věku mluvčího se opírá většinou o celkový poslechový dojem z hlasu mluvčího bez další analýzy. Jeho odhad je sice informativní, ale není příliš přesný, obvykle se věk určuje v rozmezí deseti až patnácti let (Jessen, 2007b: 185; Jessen & Jessen, 2008: 137).

3.3.2.3 Dialekt

Klasifikace dialektu ve forenzní fonetice se týká odhadnutí regionu, v němž mluvčí strávil většinu svého života do dospělosti. Zakládá se na předpokladu, že po dosažení dospělého věku již mluvčí svůj dialekt příliš nemění, i když se přestěhuje do jiné oblasti (Jessen, 2007b: 185). Dialekt se odráží ve všech jazykových úrovních, nejen ve výslovnosti, ale také v syntaxi a lexiku (Jessen & Jessen, 2008: 137).

3.3.2.4 Cizinecký přízvuk

Cizinecký přízvuk se ve forezních vědách stává s postupem času stále důležitějším. Obvykle není těžké poznat, že řeč mluvčího je ovlivněna jiným mateřským jazykem. Obtížnější je však již přesně určit mateřský jazyk mluvčího. Tato determinace může být například provedena na základě hlásek, s nimiž mluvčí nemá problémy – tím totiž vzniká pravděpodobnost, že tyto hlásky obsahuje i jeho mateřský jazyk (Jessen, 2007b: 187).

3.3.2.5 Sociolekt

Sociolekt je založen na sociolingvistických proměnných (blíže viz předchozí kap.). Jednou z dalších (ještě nejmenovaných) sociolingvistických proměnných je vzdělání, které se odráží v jazykových doménách, jako je slovník, syntax a stylistika. Stejně tak může svoji roli vzdělání hrát i ve výslovnosti, člověk s vyšším vzděláním bude patrně mluvit s menšími nářečními rysy. Další takovou sociolingvistickou proměnnou je profese mluvčího (Jessen, 2007b: 188).

3.3.2.6 Zdravotní stav

Patologie řeči může být při forezním zkoumání velmi nápomocná, protože při porovnávání hlasu může snížit počet podezřelých přicházejících v úvahu (Jessen, 2007b: 189). Je zřejmé, že dlouhodobé odchylky v řeči jsou užitečnější než ty krátkodobé (Jessen,

2007b: 189; Jessen & Jessen, 2008: 138). Jednou z takových dlouhodobých odchylek je koktání, které se vyskytuje zhruba u jednoho procenta populace, díky čemuž se může stát zásadní charakteristikou vedoucí k identifikaci podezřelého (Jessen, 2007b: 189-190). Koktání se může projevat třemi způsoby - opakováním, zastavením se (časté u explozív) či prodlužováním (zejména u frikativ). Protože se tedy balbutici¹⁶ ve způsobu koktání dále rozdělují, lze předpokládat, že osob s konkrétním typem koktání se v populaci objeví méně než 1 % (Baldwin & French, 1990: 50-51). Další forenzně relevantní řečové odchylky ztělesňují například vady řeči (Jessen, 2007b: 190).

3.3.2.7 Další zvláštnosti hlasu mluvčího

Kromě výše zmíněných kategorií by měla být při klasifikaci mluvčího věnována pozornost i dalším zvláštnostem hlasu mluvčího, jež se v populaci příliš často nevyskytují. Mezi ně patří například velmi hluboké nebo naopak vysoké hlasy, nezvyklá kvalita hlasu, silná aspirace či hodně rychlé nebo naopak pomalé tempo řeči (Jessen & Jessen, 2008: 138).

¹⁶ Lidé s poruchou plynulosti řeči.

4 MASKOVÁNÍ HLASU

Maskování hlasu bývá obvykle definováno jako záměrná změna charakteristických rysů hlasu člověka za účelem skrytí své skutečné identity (např. Künzel, 2000: 150; Perrot, Aversano, & Chollet, 2007a: 101; Perrot & Chollet, 2008: 5682; Zhang & Tan, 2008: 118). Někdy však pod pojem „voice disguise“ mohou být zařazeny i nedobrovolné změny hlasu způsobené emocionálním stavem mluvčího, nemocí a užitím drog či alkoholu nebo změny spojené s komunikačním kanálem, jako je například omezení šířky pásma v telefonech, telefonních systémech a nahrávacích zařízeních (Rodman, 1998; Rodman & Powell, 2000).

Člověk je schopen produkovat promluvu výrazně odlišnou od své normální hlasové kvality, změnou některých kvalitativních rysů hlasu se dokonce dokáže posunout do blízkosti populačního průměru (Kreiman & Sidtis, 2011: 243). Maskování tak zdůrazňuje flexibilitu produkce v rámci řečového projevu jednoho mluvčího (Jessen, 2008: 684), která umožňuje mluvčím zakrýt jejich základní hlasové vzorce (Kreiman & Sidtis, 2011: 243). Tím se stává jedním z největších problémů forenzního rozpoznání mluvčího.

Výzkum maskování hlasu se začal objevovat v sedmdesátých letech, kdy se jím zabývali třeba Künzel nebo Köster (Perrot & Chollet, 2008: 5682). Zkoumání prováděná v této době lze rozdělit do dvou hlavních proudů, elektronického a neelektronického. Neelektronický zkoumal především schopnost laických posluchačů identifikovat osoby maskující svůj hlas, elektronický často zahrnoval analýzu spektrogramů řeči, tzv. otisků hlasu (Rodman, 1998). V současnosti se pak již více než deset let vědci také snaží o vyvinutí automatických systémů rozpoznání mluvčího detekujících maskování hlasu (Perrot & Chollet, 2008: 5682).

Celkové procento snah o modifikaci přirozeného hlasu v kriminálních případech není vysoké, na základě údajů německého BKA se během posledních dvaceti let 20. století maskování hlasu objevilo mezi 15 % a 25 % forenzních případů (Künzel, 2000: 149). Podobný údaj získaný na základě komunikace se třemi německými forezními experty (Künzelem, Kösterem a Masthoffem) uvádí i Braun (2006: 179) - ve 175 analyzovaných případech došlo k maskování hlasu u 22,9 % mluvčích. Nejmenší procento maskování hlasu ve forenzní praxi udává Masthoff (1996: 161), a to pouze v 5 % ze všech případů zpracovaných na univerzitě v Trieru v letech 1989-94.

Frekvence výskytu maskování hlasu závisí i na typu trestného činu, k němuž se řečový projev váže. V případech, kdy si pachatel možnost nahrávání svého hlasu za účelem policejního vyšetřování uvědomuje, pravděpodobnost maskování hlasu znatelně stoupá (Braun, 2006: 178; de Figueiredo & de Souza Britto, 1996: 168; Künzel, Gonzales-Rodriguez, & Ortega-García, 2004; Masthoff, 1996: 160; Zhang & Tan, 2008: 118). Například v německém BKA se během let 1989-94 maskování hlasu objevilo v 52 % případů, kdy si pachatel mohl myslet, že bude během kriminálního jednání nahráván (Masthoff, 1996: 160-161). Vyšší pravděpodobnost úsilí o změnu hlasu se vyskytuje i ve chvíli, kdy se pachatel s volaným osobně zná (Braun, 2006: 180).

Nejčastěji dochází k maskování hlasu mluvčího za účelem skrytí své identity u únosů (Braun, 2006: 178; de Figueiredo & de Souza Britto, 1996: 168; Künzel, 2000: 150; Künzel a kol., 2004; Masthoff, 1996: 160; Zhang & Tan, 2008: 118), vydírání (Braun, 2006: 178; Künzel, 2000: 150; Künzel a kol., 2004; Masthoff, 1996: 160; Zhang & Tan, 2008: 118), sexuálního harašení, podvodných telefonátů (Künzel, 2000: 150) a volání na tísňovou policejní linku (Zhang & Tan, 2008: 118). U vydírání dokonce Masthoff (1996: 160-161) udává (na základě údajů z Trieru) maskování hlasu v 69 % případů. Naopak u přestupků a vražd se maskování nevyskytuje téměř vůbec (Braun, 2006: 180).

4.1 Typy záměrného maskování hlasu

Maskování hlasu může nabývat mnoha různých forem. Zpravidla se rozlišuje mezi dvěma základními typy, a to elektronickým a neelektronickým. Neelektronické maskování může být realizováno využitím mechanických prostředků (např. pera) nebo přirozenou modifikací hlasu (Perrot a kol., 2007a: 102), elektronické maskování hlasu pak představuje manipulaci hlasu prostřednictvím speciálního softwaru (Perrot, Morel, Razik & Chollet, 2009: 180).

4.1.1 Elektronické maskování hlasu

Přesto, že je použití elektronického zařízení k modifikaci hlasu asi nejúčinnější formou jeho maskování (Hollien, 2002: 49), v kriminálních případech se tyto elektronické formy vyskytují jen zřídka (Clark & Foulkes, 2007: 196; Künzel a kol., 2004; Masthoff, 1996: 161). Masthoff (1996: 161) dokonce uvádí, že se objevuje v méně než 1 % všech případů. Častěji než při páchání trestných činů se využívá například při televizním natáčení ve chvíli, kdy je odhalení identity dotazované osoby nežádoucí (Perrot a kol., 2007a: 103).

Prostřednictvím softwaru lze ovlivnit velké množství parametrů, jako například základní frekvenci, formanty nebo rytmus (Perrot a kol., 2009: 180). Hlavní techniku využití elektronického zařízení k pozměnění hlasu však představuje modifikace rejstříku F0 jednoduchým pohybem průměru či kontury F0. Nejeftivnější způsob změny hlasu pak sestává z upravení prozodie na základě modifikace F0 nebo trvání segmentů. Dokonce existují i metody, které tyto dvě charakteristiky modifikují nezávisle, aniž by došlo k posunu dalšího parametru. Výhodou softwaru je, že umožňuje výraznou transformaci základní frekvence nebo hlasového tónu mluvčího v reálném čase a zároveň produkci nekonečného množství nových hlasů (Perrot a kol., 2007a: 103-104).

Perrot a kol. (2007a: 102) rozlišují ve způsobu využití elektronických zařízení za účelem maskování hlasu konverzi a transformaci. Konverze se provádí přeměnou zdroje hlasu mluvčího směrem k určitému mluvčímu cílovému. Transformace představuje elektronickou modifikaci některých specifických parametrů, jako je frekvence, řečové tempo apod., aby bylo dosaženo změny zvuku hlasu.

4.1.2 Neelektronické maskování hlasu

U neelektronického maskování hlasu lze dále rozlišovat mezi imitací a vlastním maskováním (Perrot a kol., 2007a: 102). Přestože oba způsoby do značné míry závisí na schopnosti pachatele měnit hlas, imitace navíc vyžaduje i určitý talent, trénink a praxi (Kreiman & Sidtis, 2011: 246-247).

4.1.2.1 Imitace

„Imitace může být považována za extrémní formu maskování mluvčího, v níž se mluvčí nejen pokouší změnit své hlasy takovým způsobem, že již nemohou být rozpoznáni, ale dobrovolně manipulují své hlasy tak, že mohou být mylně pokládány za hlasy jiných lidí“ (Schlichting & Sullivan, 1997: 149). Imitace hlasu tak může na první pohled představovat výhodný způsob maskování identity právě odkazováním na nějakou další osobu. Imitující však musí řečové vlastnosti jiného mluvčího spolehlivě a konzistentně napodobovat po delší dobu (Braun, 2006: 177-178), proto imitace není jednoduchou, a tudíž ani univerzální metodou (Perrot a kol., 2009: 180). Pravděpodobně i z tohoto důvodu nebývá v kriminálních případech příliš využívána (Braun, 2006: 178; Kreiman & Sidtis, 2011: 247).

Imitace se ale nemusí vyskytovat pouze za účelem napodobení konkrétní osoby, ale také s cílem napodobit skupinovou identitu, například sociálním nebo regionálním dialektem

(Sjöström, Eriksson, Zetterholm, & Sullivan, 2008: 148; Zetterholm, 2007: 193). Tento typ napodobování však zpravidla k imitaci řazen nebývá.

K tomu, aby imitátor předstíral různé mluvčí, napodobuje mnoho různých jevů - laryngálních i supralaryngálních (Zetterholm, 2006: 70). Není možné, aby si imitátor osvojil všechny řečové charakteristiky určitého mluvčího, obvykle napodobí jen některé specifické charakteristiky, které k narušení rozpoznání stačí (Perrot a kol., 2007a: 107). Profesionální imitátor je cílovému mluvčímu schopen přizpůsobit dialekt, intonační vzorce, kvalitu hlasu, řečové tempo, základní frekvenci, ale také některé formantové frekvence. Dojem může být zesílen i užitím charakteristických slov a frází (Zetterholm, 2000: 148).

Vliv imitace hlasu na identifikaci mluvčího byl testován jak v poslechových testech (např. Schlichtling & Sullivan, 1997; Eriksson a kol., 2010), tak v úspěšnosti automatických rozpoznávacích systémů (např. Farrús, Wagner, Anguita, & Hernando, 2008). Ukázalo se, že imitace může snížit nejen přesnost a užitečnost testů hlasové konfrontace, ale i ovlivnit schopnosti identifikace automatických rozpoznávacích systémů. Imitace hlasu tak může představovat potenciální hrozbu bezpečnosti systémů, jež využívají automatické rozpoznání mluvčího (Farrús a kol., 2008).

4.1.2.2 Maskování

Mluvčí ve snaze o maskování své identity může využít variabilitu svého hlasu, a to změnou základní frekvence, supralaryngálních charakteristik (Jessen, 2008: 685; Perrot & Chollet, 2008: 5682), hlasové kvality, tempa či lingvistických aspektů řeči (Jessen, 2008: 685). K modifikaci řečových charakteristik mu mohou dopomoci i další, externí nástroje. Jednotlivé typy maskování hlasu bývají na základě společných charakteristik tříděny do několika kategorií. Ty však bývají v odborné literatuře pojímány různě. Tabulky 1-3 podávají příklady tří odlišných přístupů k rozdělení a typům maskování hlasu. Některé z typů maskování hlasu budou detailněji rozebrány níže.

Změna	maskování
zdroj hlasu	- zvýšená základní frekvence (s nebo beze změny rejstříku ¹⁷) - snížená F0

¹⁷ Pojem hlasový rejstřík označuje uměle stanovené rozložení hlasu člověka do několika částí podle jeho výšky (nejčastěji se rozlišuje mezi prsním, hlavovým a středním), tóny v jednom rejstříku charakterizuje jejich stejnorodost - podobnost zabarvení a podobné postavení hlasového ústrojí. Tyto tónové řady pak mezi sebou plynule přecházejí (Palková, 1997: 59).

	<ul style="list-style-type: none"> - třepená fonace - umělý chrapot - užití výchlípkových řas (= nepravých hlasivek) - šeptaná řeč
rezonanční jevy	<ul style="list-style-type: none"> - cizí objekt ve vokálním traktu (např. kuličkové pero) - dodatečný rezonátor (např. pivo může být drženo blízko úst) - hypo-/hypernazalita (nejčastěji hyponazalita při ucpání nosu) - tkanina držená před ústy (např. kapesník)
jazyk	<ul style="list-style-type: none"> - změna nářečního zabarvení - užití jiného dialektu (stejného jazyka) - předstírání cizineckého přízvuku
způsob mluvy	<ul style="list-style-type: none"> - redukce normální variace F0 (umělá monotónnost) - přehánění normální variace F0 - změna mluvního tempa (nejčastěji zpomalení) - změna přízvuchných vzorů („robotický hlas“)

Tabulka 1 Časté typy maskování ve forenzní analýze řeči (převzato z Künzel, 2000: 151)

<p>Změna zdroje</p> <ul style="list-style-type: none"> - změna F0 - šepot - užití třepené fonace, falzetu, zdvojnásobení periody <p>Změna tvaru vokálního traktu</p> <ul style="list-style-type: none"> - změna v nazalitě (ucpání nosu / snížení véla) - zkrácení vokálního traktu (zatažení rtů, zvýšení hrtanu) - prodloužení vokálního traktu (vyšpulení rtů / snížení hrtanu) - změna tvaru vokálního traktu (například rozšíření či stažení hltanu) <p>Změna dynamických aspektů řeči</p> <ul style="list-style-type: none"> - užití přízvuku - změna intonační kontury - změna tempa řeči nebo rytmu - objekt v ústech, sevřené zuby <p>Kombinace uvedených příkladů</p>
--

Tabulka 2 Některé způsoby maskování hlasu (převzato z Kreiman & Sidtis, 2011: 244)

Fonační	Fonemické	Prozodické	Deformace
Zvýšení základní frekvence	Užití dialektu	Intonace	Ucpané nosní dírky
Snížení základní frekvence	Cizinecký přízvuk	Umístění přízvuku	Sevřená čelist
Třepeň fonace	Řečová porucha	Prodłużování a krácení segmentů	Blokování skusu (řeč kuřáků fajfky)
Šepot	Imitování	Řečové tempo	Vyšpulení rtů
Inspirační hlas	Hypernazalita (snížené vélum)		Stažení tváří
Zdvižení nebo snížení hrtanu			Držení jazyka
			Objekt v ústech
			Objekt přes ústa

Tabulka 3 Dobrovolné neelektronické maskování hlasu (převzato z Rodman & Powell, 2000; Rodman, 1998)

Mezi nejoblíbenější typy maskování hlasu ve forenzních případech bývá řazen falzet a zvýšení základní frekvence (Braun, 2006: 179; Künzel, 2000: 149; Perrot a kol., 2009: 180; Zhang & Tan, 2008: 118), ucpaní nosních dírek (Jessen, 2008: 685; Künzel, 2000: 149; Perrot a kol., 2009: 180; Svobodová in Rak a kol., 2008: 463; Zhang & Tan, 2008: 118), kladení předmětů do úst (Jessen, 2008: 685; Perrot a kol., 2007a: 108; Svobodová in Rak a kol., 2008: 463; Zhang & Tan, 2008: 118) a třepeň fonace (Braun, 2006: 179). Oblíbenými metodami jsou také změny mluvního tempa (Jessen, 2008: 685; Svobodová in Rak a kol., 2008: 463; Zhang & Tan, 2008: 118), předstírání cizineckého přízvuku (Clark & Foulkes, 2007: 196; Jessen, 2008: 685; Künzel, 2000: 149; Svobodová in Rak a kol., 2008: 463; Zhang & Tan, 2008: 118), objekt před ústy (Perrot a kol., 2009: 180), snížená základní frekvence (Braun, 2006: 179; Perrot a kol., 2009: 180) a předstírání řečových vad (Svobodová in Rak a kol., 2008: 463). Do nejčastěji se objevujících metod maskování hlasu bývá řazen i šepot, ale v určení frekvence jeho výskytu se jednotliví autoři liší (srov. např. Braun, 2006: 179; Künzel, 2000: 149; Masthoff, 1996).

V České republice podle Svobodové (in Rak a kol., 2008: 463) používají pachatelé zatím jen velmi primitivní techniky maskování hlasu. Nejčastěji si volí ucpaný nos, kapesník na mikrofonu či změnu kvality hlasu. Poslední dobou narůstá snaha o řeč s ruským přízvukem.

Typy maskování hlasu přímo ovlivněné způsobem hlasivkové činnosti, jako je šepot, třepeň fonace, falzet, chraplavý hlas apod., produkují poměrně stabilní akustické efekty. Naopak typy jako vkládání předmětů do úst ovlivňují segmentální rysy řeči, které nezapadají

do žádného nářečního rámce, což může ztížit pokusy o identifikaci mluvčího (de Figueiredo & de Souza Britto, 1996: 169).

Mluvčí převážně využívají pouze jeden ze způsobů maskování hlasu, ale strategie se mohou objevit i v kombinaci (Braun, 2006: 178; Jessen, 2008: 685; Künzel, 2000: 150; Masthoff, 1996). V praxi se však tyto kombinace omezují na maximálně dvě nebo tři oblasti jazykově-řečového chování (Braun, 2006: 178).

Preferovanými formami maskování hlasu se ve své studii zabýval Masthoff (1996). Nahrávky dvaceti studentů ukázaly, že nejčastěji vybraným typem maskování byla změna způsobu fonace, nejčastěji v podobě šepotu, který byl následovaný zvýšením základní frekvence. Na dalších příčkách se umístilo její snížení a inspirační hlas (ten nebyl srozumitelný). V artikulační rovině nejčastěji docházelo k manipulaci vokálního traktu, jako vyšpulení rtů, držení nosu, zvýšení hrtanu a vytažení tváří. Mezi méně časté, ale stále nezanedbatelné techniky patřily imitace dialektu a zablokování pohybů jazyka, nejméně - 10 % lidí imitovalo cizinecký přízvuk a simulovalo řečové vady. Na úrovni způsobu mluvy docházelo k negramatické intonaci, pauzám a trvání. O několik let později zpracoval Masthoff další studii na podobné téma (Masthoff, 2000), 23 studentů mělo tentokrát cca 10 dní čas na to, aby si rozmysleli způsob maskování svého hlasu. V modifikaci hlasu na jediné úrovni se oproti předchozímu výzkumu nejčastěji objevovaly změny na úrovni artikulační, až poté následované změnou způsobu fonace. Změna způsobu fonace však tvořila největší část všech změn, společně s kombinacemi několika úrovní byla použita v 62 % všech případů. Nejoblíbenější bylo zvýšení základní frekvence, následované jejím snížením a aspirací, šepot použili pouze dva účastníci.

Ve výsledcích svého výzkumu z roku 1996 Masthoff navíc zjistil, že muži tíhnou spíše ke zvýšení základní frekvence, zatímco ženy ji snižují. Tento závěr však nesouhlasí s Künzelovým (2000) zjištěním, který nechal 50 žen a 50 mužů, aby si vybrali dva ze tří nabízených typů (snížení a zvýšení F₀, ucpání nosních dírek) maskování hlasu. Přestože muži mnohem častěji než ženy používali falzet, čímž dosáhli nápadnější změny hlasu než ženy, a ženy naopak vykazovaly větší odchylky od své normální základní frekvence při jejím snížení, vybírali si muži častěji ke svému maskování snížení F₀, zatímco ženy naopak. Celkově ale nejoblíbenější výběr žen představovala řeč se zacpaným nosem.

V souvislosti se změnou základní frekvence během snahy o maskování hlasu Künzel (2000) dále zjistil, že se u obou skupin mluvčích (u žen silněji) projevila tendence vybírat si

způsob maskování podle základní frekvence svého hlasu - lidé s nižší základní frekvencí často volili její snížení, zatímco ti s vyšší F0 preferovali spíše její zvýšení. Navíc se ukázalo, že spolu s vybraným typem maskování dochází ke změnám dalších řečových charakteristik, jako například ke snížení tempa společně se změnou F0, zvýšení hlasitosti společně se zvýšením F0 a mírnému zvýšení základní frekvence během řeči s ucpanými nosními dírkami.

Dalším z typů maskování založeného na změně fonace, které může mít silný vliv na identifikaci mluvčího, je šepot. Šepot totiž odstraňuje informaci o základní frekvenci (Perrot a kol., 2009: 180; Hollien, 2002: 49), snižuje dostupnost informací o intenzitě, kvalitě hlasu a částečně i prozodii nebo načasování řeči (Hollien, 2002: 49). Přesto je šeptaná řeč poměrně srozumitelná. Na časové rovině jsou vcelku dobře zachované prozodické jevy, spektrální jevy šeptané řeči jsou dobře zachované u neznělých konsonantů, ale přesné jsou dokonce i u znělých konsonantů a vokálů. Protože šeptání znělých konsonantů vyžaduje více artikulační snahy, je k výslovnosti těchto hlásek potřeba více času, aby se řeč udržela srozumitelnou (Jovičić & Šarić 2008: 264-273). Jovičić a Šarić (2008) zkoumali trvání segmentů v normální a šeptané řeči. Zjistili, že trvání konsonantů v šeptané řeči je zhruba o 10 % delší a že šeptané znělé konsonanty mají nižší intenzitu než ty realizované při normální fonaci. Nejvíce bylo trvání šeptaných konsonantů prodlouženo v iniciální pozici, na syntaktické rovině na konci vět. U znělých konsonantů se trvání šeptaného segmentu prodloužilo více než u neznělých, a to průměrně o 15,3 %. U neznělých šeptaných konsonantů šlo o prodloužení v průměru o 5,8 %.

Další poměrně často využívanou modifikaci fonačního ústrojí představuje třepená fonace. Hranice mezi normální a třepenou fonací však není vždy ostrá, protože mnoho mluvčích používá třepenou fonaci na konci frází a nádechových úseků. Třepená fonace je spojena se snížením proudu vzduchu a průměrné základní frekvence. Tento způsob fonace nejvíce ovlivňuje znělé segmenty, zatímco neznělé zůstávají relativně neovlivněné (Hirson & Duckworth, 1993: 193-194). Ve výzkumu Hirsona a Duckwortha (1993) vyšlo najevo, že přestože někteří mluvčí nebyli schopni vytvořit adekvátní třepenou fonaci, napodobovali rysy často doprovázející tento typ maskování hlasu, jako sníženou F0 a pomalejší tempo.

Poměrně často zkoumaným druhem maskování hlasu je napodobování cizineckého přízvuku nebo nářečí. Tento typ je lingvisty díky jejich znalosti fonologické struktury jazyka a jeho variet obvykle snadno rozpoznatelný, a tudíž jej nebývá obtížné odlišit od skutečného přízvuku mluvčího (Rose, 2002). V imitaci cizineckého přízvuku jsou mluvčí schopni využít

několik forem změn k zamaskování svého hlasu. Přestože se většina změn odehrává v artikulační rovině, výjimkou nejsou ani rozdíly v hodnotách výšky základní frekvence a jejím rejstříku. Tato imitace je charakteristická nejen fonetickými, ale i nefonetickými změnami, zvláštní prvky se mohou objevovat i na lexikální, syntaktické a morfologické rovině (Neuhauser, 2008: 154-155).

Použití externího objektu, v tomto případě vkládání tužky mezi zuby, zkoumali de Figueiredo a de Souza Britto (1996). Zjistili, že tento typ maskování ovlivňuje zejména tři artikulační charakteristiky – zaokrouhlování rtů, otevírání čelisti a zasunutí jazyka. Zaokrouhlování rtů se projevuje na výsledném charakteru zadních zaokrouhlených vokálů (zvýšení prvních tří formantů), míra otevřenosti čelisti se odráží v produkci nízkých, ale i vysokých vokálů. Při normální produkci nízkých vokálů je čelist otevřená, proto je ovlivněna i jejich artikulace (aby tužka nevypadla, není možné při řeči čelist tolik otevřít), a to snížením prvních tří formantů. Při produkci vysokých vokálů narůstá hodnota prvního formantu, hodnoty pro druhý a třetí klesají, na což má podle autorů pravděpodobně vliv nejen míra otevřenosti čelisti, ale také posun jazyka směrem dozadu. Přestože konsonanty nebyly v rámci této studie zkoumány, výsledky naznačují posteriorizaci alveolární frikativy /s/. Kromě těchto specifických jevů našli autoři další odlišnosti v závislosti na promluvách jednotlivých mluvčích.

4.1.2.3 Identifikace mluvčího při maskování hlasu

Výzkumy zabývající se vlivem jednotlivých typů maskování hlasu na míru úspěšné identifikace se v současnosti soustřeďují jak na posluchače, tak na automatické systémy rozpoznání mluvčího. Výzkumům identifikace mluvčího v sedmdesátých letech, mezi něž patřila například studie Reicha, Molla a Curtise (1976), se podařilo vyvrátit dříve používanou metodu otisku hlasu, když dokázaly, že pouze na základě spektrogramu není snadné určit, které z rysů by měly být ignorovány a které naopak považovány za relevantní.

Výsledky studií zaměřených na míru úspěšnosti identifikace mluvčího automatickými rozpoznávacími systémy se do jisté míry shodují. Zpravidla největší problém při správné identifikaci mluvčího představuje maskování ve formě ruky před ústy a pohybů F0, nejčastěji směrem dolů (např. Perrot a kol., 2009; Perrot, Preteux, Vassuer, & Chollet, 2007b; Zhang & Tan, 2008). Naopak nejvyšší efektivitu systémy vykazují u řeči s ucpanými nosními dírkami (Künzel a kol., 2004; Perrot a kol., 2007a). Systém vyvinutý na čínské univerzitě,

jehož účinnost byla zkoumaná na základě deseti druhů maskování hlasu, byl nejméně úspěšný při rozpoznání cizineckých přízvuků (Zhang & Tan, 2008).

Perrot a Chollet (2008) zkoumali v souvislosti s maskováním hlasu účinnost tří různých systémů automatického rozpoznání mluvího. Pro svůj výzkum si vybrali čtyři v kriminálních případech často používané druhy maskování hlasu - ruku před ústa, ucpané nosní dírky a zvýšenou a sníženou základní frekvenci. Po trénovací fázi autoři testovali spolehlivost tří klasifikačních metod - k-nejbližších sousedů, modelu Gaussovských hustotních směsí a vektorové kvantizace následované SVM¹⁸. Nejméně účinná byla poslední jmenovaná, přestože v identifikaci hlasu maskovaného rukou před ústy byla její výkonnost oproti Gaussovskému modelu horší.

Maskování hlasu se zásadně odráží i ve správné identifikaci mluvího posluchači. Reich a Duke (1979) zjišťovali vliv několika druhů maskování hlasu (napodobení stáří, chraptivý hlas, hypernazalizace, pomalé tempo řeči, volitelný způsob maskování) na poslechovou identifikaci mluvího. Posluchači byli rozděleni do dvou skupin, naivních a zkušených¹⁹, přičemž obě dvě skupiny byly stejně trénovány na rozpoznání mluvího. Z výsledků vyplynulo²⁰, že nejméně účinnou metodou maskování u naivních posluchačů byla nazalizovaná řeč a maskování podle výběru mluvího, u zkušených posluchačů se významně horší výsledek než u jiných způsobů maskování objevil pouze u nazalizovaných projevů. Celkově pak výsledky zkušených posluchačů byly lepší než výsledky naivních posluchačů. Zjevné problémy při identifikaci mluvího na základě poslechového testu se vyskytují nejen u nazalizovaných, ale i u denazalizovaných projevů (např. Perrot a kol.: 2007b).

Různou mírou elektronické modifikace základní frekvence a její vliv na poslechovou identifikaci mluvího zkoumali Clark a Foulkes (2007). Experiment ukázal, že s rostoucí mírou modifikace klesá úspěšnost rozpoznání mluvího. Co je však důležitější, výsledky odhalily, že i přes elektronickou modifikaci hlasu hrají při rozpoznání mluvího velkou roli jeho charakteristické znaky, stejně jako v případě nemaskované řeči – tzn., že někteří mluvíci jsou snadněji rozpoznatelní než jiní. Výsledky jednotlivých mluvích byly i se zvyšující se

¹⁸ Support vector machine.

¹⁹ Studenti doktorského studia a profesori oboru „Speech and Hearing Sciences“, tzn. věd zabývajících se řečí a sluchem.

²⁰ Úkolem posluchače bylo určit, zda dvě přehrávané věty byly vysloveny stejným, nebo rozdílným mluvícem.

zírou modifikace konzistentní, u některých z nich se maskování formou modifikace ukázalo jako méně úspěšné než u jiných.

Sjöström a kol. (2008) se zabývali vlivem dvou odlišných dialektů jednoho mluvčího na schopnost jeho identifikace posluchači. Na základě nahrávek mluvčího perfektně ovládajícího dva švédské dialekty sestavili testy hlasové konfrontace, které dokázaly, že i přepínání nářečí může být použito jako forma maskování hlasu, neboť posluchači měli větší problémy mluvčího hovořícího odlišným dialektem identifikovat. Dialekt jako formu maskování mluvčího zkoumal i Markham (1999). Chtěl vědět, zda jsou mluvčí schopni přesvědčivě změnit svoje nářečí tak, aniž by to kritičtí posluchači postřehli. Přestože se ukázalo, že se mluvčí ve schopnosti přesvědčivě a konzistentně imitovat přirozeně znějící dialekty liší, některé formy odkazující na nářečí byly posluchači přijaty jako základy přirozeného dialektu, ačkoliv signalizovaly vliv více než jednoho dialektu.

4.2 Maskování hlasu ve forenzní praxi

Identifikace normálních hlasů může být ztížena hlukem v pozadí, různými kanály přenosu, extrémními emocemi, nemocí apod. Pokud však dojde k maskování hlasu, identifikace se může stát ještě složitější, ba dokonce nemožnou. Proto je důležité maskování hlasu zkoumat (Zhang & Tan, 2008: 118). Jeho výzkum lze rozdělit do čtyř základních kategorií, a to posouzení potencionálních vlivů maskování hlasu na spektrografickou reprezentaci řeči, testování schopnosti posluchačů rozpoznat přítomnost maskování, posouzení schopnosti posluchačů identifikovat maskované hlasy a prozkoumání typů maskování vybraných mluvčími a fonetických důsledků těchto typů (Clark & Foulkes, 2007: 199).

Ve forenzních případech je detekce maskování hlasu jako takového problematická, protože obvykle není k dispozici vzorek normálního hlasového projevu mluvčího. Například při napodobování vad řeči nelze při absenci nahrávky normálního řečového projevu pachatele maskování hlasu poznat (Künzel, 2000: 150). Lepší případ nastává, pokud se podezřelý pokouší o maskování svého hlasu až ve chvíli, kdy je zadržen. V tomto případě je totiž možná určitá kontrola. Po podezřelém lze požadovat velké množství řeči za různých podmínek, protože pozměněný hlas není lehké dlouhodobě udržet (Hollien, 2002: 51).

Při srovnávání dvou typů nahrávek a odhadování ne/totožnosti mluvčího se jako forenzně relevantní ukazují zvláštní nevědomé řečové návyky mluvčích, které během

maskování neskryjí. Jedním z takových případů je například hezitace (Jessen, 2007a: 50-51). Mluvčí se totiž zpravidla soustředí na techniku maskování hlasu, ale nejsou schopni kontrolovat další aspekty své řeči (jako dialekt či hezitační zvuky). Hezitační zvuky se mohou navzájem lišit v závislosti na ne/přítomnosti nazálního konsonantu, přítomnosti glotalizace, nazalizace apod. Stejně tak se mohou lišit i svým trváním a četností výskytu (Jessen, 2008: 690-691).

V kriminálních případech se často stává, že pachatel má tendenci snižovat míru maskování svého hlasu s počtem svých řečových projevů (obvykle telefonátů), někdy dokonce může maskování vymizet úplně. Maskování hlasu se však neliší jen událost od události, ale i během jedné promluvy. Představuje totiž odchylku od normální řečové produkce, není pro mluvčího běžnou záležitostí (Künzel, 2000: 151-154). Proto může být nápomocné i hledání nekonzistentních prvků v řeči, zejména pak pasáží, v nichž se objevují „pauzy v maskování“ (Hollien, 2002: 50). Při snaze o modifikaci hlasu je totiž poměrně běžné, že mluvčí nejsou schopni udržet maskování konzistentně. Například pokud se mluvčí pokouší napodobit jiný přízvuk, jen velmi zřídka bývá autentický, neboť se během svého projevu mnohdy ke svému původnímu přízvuku vrací. Některé specifické rysy mluvčího navíc nemusí být modifikací vůbec zasaženy (Clark & Foulkes, 2007: 198-199).

5 METODA

5.1. Charakteristika výzkumného vzorku

Pro účely tohoto výzkumu bylo v průběhu roku 2011 a na počátku roku 2012 nahráno 45 mladých lidí (nejčastěji z řad spolužáků, přátel a známých autorky této práce²¹) ve věku od 20 do 31 let, kteří byli vybíráni tak, aby mohli být zařazeni do jedné ze tří předem stanovených skupin. První nahrávanou skupinu tvořili absolventi a studenti 4. a vyššího ročníku magisterského oboru Fonetika na Filozofické fakultě Univerzity Karlovy. Druhá skupina sestávala ze studentů a absolventů nelingvistických oborů vysokých škol nacházejících se v Praze a Plzni. Do třetí kategorie spadali lidé s nižším než středoškolským vzděláním s maturitou žijící v okolí Prahy a Plzně, nejčastěji se jednalo o absolventy tříletých učňovských oborů. Pro potřeby výzkumu bylo do každé z těchto tří kategorií zařazeno 15 lidí, z poslední byli dva mluvčí z důvodu dosažení středoškolského vzdělání dodatečně vyloučeni. Většina participantů nebyla za účast na tomto výzkumu odměněna²².

5.2 Nahrávání a výzkumný materiál

Po přečtení a podepsání informovaného souhlasu (viz příloha č. 1), jímž vyjádřili ochotu podílet se na tomto výzkumu, byli všichni participanté požádáni o přečtení dvou textů, které jim byly postupně předloženy. První z nich byl konstruován jako novinová zpráva vypovídající o vydírání investigativního žurnalisty (viz příloha č. 2). Předtím, než se přistoupilo k samotnému nahrávání, byli účastníci výzkumu požádáni, aby se s textem (nejen kvůli jeho složitosti) důkladně seznámili. K četbě prvního textu nebyly uloženy žádné zvláštní instrukce, mluvčí jej měli jen nahlas přečíst. Jediným požadavkem bylo, aby se mluvčí v případě přechnutí během čtení vrátili na začátek dané věty a přečetli ji celou znovu. Druhý text byl sestaven jako přepis telefonátu, v němž neznámý pachatel vyhrožuje novináři (viz příloha č. 3). Aby byla zajištěna srovnatelnost obou dvou projevů, byly texty záměrně sestaveny tak, aby se v nich opakoval co možná nejvyšší počet slov na podobných pozicích ve větě. Druhý text byl participantům poskytnut až po nahrání prvního, do této chvíle nikomu z nich nebyl sdělen účel výzkumu. Společně s obdržením druhého textu byli participanté instruováni, aby zkusili zamaskovat svůj hlas tak, aby je nikdo nepoznal. Na vymyšlení

²¹ Zbylí participanté byli nalezeni metodou sněhové koule – jednalo se o přátele či známé již nahaných mluvčích.

²² Výjimku představují pouze dva participanté z poslední skupiny mluvčích. Z důvodu nedostatku nahrávek v této skupině byla těmto dvěma mluvčím nabídnuta drobná finanční odměna.

vhodné taktiky byla mluvčím ponechána libovolně dlouhá doba, obvykle jim však stačilo pouze několik minut. Výběr techniky maskování byl ponechán čistě na participantech, a to s jedinou podmínkou – maskovaný projev musel být srozumitelný.

Nahrávky byly pořízeny v Praze a v Plzni. V Praze byly zvukové záznamy pořízeny v nahrávací kabině na půdě Fonetického ústavu FF UK. V Plzni bylo vytvořeno provizorní nahrávací studio v bytě jednoho z participantů tohoto výzkumu, jenž měl k dispozici licencovanou verzi programu Cool Edit a mixážní pult. Kvalitní mikrofon byl pro účely tohoto nahrávání autorce práce zapůjčen Fonetickým ústavem. Nahrávací kabina byla vytvořena ze šatní skříně, pro její maximální odzvučení byly její dveřní stěny a nekrytá místa přikryty prostěradly. Na jedné straně dveří pak byla umístěna ještě karimatka, na niž se pomocí špendlíku připevnil text, který měl mluvčí číst. Jedinou nevýhodu těchto podmínek představovala vyšší teplota v provizorní nahrávací kabině.

5.3. Způsob zpracování dat

5.3.1 Poslechová analýza

Za účelem poslechové analýzy byly opakovaně přehrávány jednotlivé zvukové záznamy získané od všech participantů (celkem 86 promluv, dvě od každého mluvčího). Vždy byl porovnáván nemaskovaný a maskovaný projev mluvčího, na papír byly tužkou zaznamenávány charakteristiky hlasu a lingvistické aspekty řeči mluvčího a jejich odchylky v maskovaném projevu. Vzhledem k velkému počtu participantů a délce textu nebyly projevy transkribovány.

5.3.2 Akustická analýza

Akustická analýza byla prováděna pomocí programu Praat (Boersma & Weenink, 2012). Zaměřena byla na základní frekvenci hlasu mluvčího, vokální formanty a fonační kvality měřené na vokálech a nazálních konsonantech – jitter, shimmer a harmonicitu.

5.3.2.1 Základní frekvence

Základní frekvence hlasu představuje jeho nejhlubší a určující složku. U dospělých mluvčích se pohybuje v rozsahu 80-300 Hz, u mužských hlasů se obvykle nachází v rozmezí 100-150 Hz, u ženských 200-300 Hz (Palková, 1997: 57). Jak již bylo zmíněno v teoretické části práce, v rámci promluv jednoho člověka se rozmezí průměrné F0 zpravidla neliší o více

než 5 Hz. Hodnota základní frekvence tak ztělesňuje jeden ze základních parametrů využívaných při porovnávání více mluvčích, její měření mohou být velmi užitečná zejména v těch případech, kdy je průměrná F0 podezřelého i pachatele vyšší nebo nižší než průměr v dané populaci. Ve forenzním výzkumu nejsou zanedbatelné ani hodnoty směrodatné odchylky F0, které poukazují na její variabilitu v toku řeči mluvčího (Baldwin & French, 1990: 46-48). Právě díky těmto charakteristikám však při snahách o maskování hlasu často dochází k úmyslné manipulaci průměrné základní frekvence, a to jak k jejímu zvýšení, tak i snížení. Někdy může i vymizet, pokud fonaci nahradí šepot.

Přestože představuje jednu ze zásadních složek podílejících se na výsledné podobě lidského hlasu, základní frekvence není konstantní ani v běžné řeči, bývá ovlivňována řadou faktorů, jako například paralingvistickými jevy či emočními stavy. Aby její naměřené hodnoty byly co možná nejspolehlivější, navrhují Lindh a Eriksson (2007) vzorec pro výpočet alternativní základní hladiny $F0^{23}$ - ta totiž z několika naměřených hodnot (průměrná F0, medián, základní hladina) nejlépe charakterizovala mluvčí v experimentech založených na různých emočních stavech, rozdílné kvalitě kanálu a variabilitě v mluvním úsilí²⁴. Alternativní základní hladina vychází z původní základní hladiny (base line), která byla navržena v rámci modulační teorie řeči. Ta počítá s jakousi neutrální motorickou pozicí, jež (v případě řeči) většinou není ovlivňována mluvním stylem, obsahem či mluvním úsilím (blíže viz Lindh & Eriksson, 2007: 2025-2026).

K měření základní frekvence hlasu participantů byly použity nesestříhané zvukové záznamy, z maskovaných projevů však byly odstraněny úseky nesouvisející se zkoumaným textem - jako smích či komentáře průběhu nahrávání, díky nimž by mohlo dojít ke zkreslení získaných hodnot. Všechny hodnoty byly získány prostřednictvím algoritmu detekujícího akustickou periodicitu, který je založen na metodě autokorelace. Hodnoty základní frekvence - hodnoty průměrné F0, jejího mediánu, směrodatné odchylky a alternativní základní hladiny navrhované Lindhem a Erikssonem (2007) byly získány s pomocí skriptu napsaného Lindhem (2005-2006). Pro vytvoření histogramů byl napsán jednoduchý skript, který měřil a vypisoval hodnoty základní frekvence v každých 5 ms obou projevů. Výsledné hodnoty pak byly

²³ Alternativní základní hladina představuje hodnotu v 7,64 % všech naměřených hodnot F0. Například z celkových 10000 získaných hodnot představuje 7,64 % číslo nacházející se na 764. místě, když jsou naměřené hodnoty seřazeny od nejnižší k nevyšší. To se pak rovná alternativní základní hladině.

²⁴ Tento experiment byl založen na rozdílné vzdálenosti mluvčího a posluchače.

převedeny do excelového souboru, v němž byly vymazány hodnoty „undefined“²⁵. Takto upravený soubor pak sloužil jako základ k vytvoření histogramů v programu Statistica.

5.3.2.2 Vokální formanty, jitter, shimmer, hnr

Jednou z často zkoumaných hodnot ve forenzním výzkumu jsou mimo základní frekvenci hlasu také hodnoty vokálních formantů, které se mění v závislosti na kvalitě vokálů (Butcher, 2002: 8). Bylo zjištěno, že míra schopnosti rozlišovat mluvčího na základě hodnot formantů se mění v závislosti na zkoumaném vokálu. U jednotlivých samohlásek se totiž vyskytuje rozdílná míra variability uvnitř řeči mluvčího i různé rozpětí možných formantových frekvencí hlásky (de Jong, McDougall, & Nolan, 2007; Nolan, McDougall, de Jong, & Hudson, 2006). V současnosti se ve forenzním výzkumu provádí i tzv. dlouhodobá analýza formantů, z níž vycházela i tato diplomová práce. „Dlouhodobá analýza formantů (LTF) shrnuje rezonanční charakteristiky vokálního traktu jedince a odráží individuální anatomii a artikulační zvyky“ (McDougall, Nolan, Harrison, & Kirchhübel, 2012). Tato dlouhodobá analýza formantů se provádí na základě průměrných hodnot formantových frekvencí různých vokálů²⁶.

Akustická analýza provedená v této práci se nezaměřila pouze na hodnoty formantů u vybraných vokálů, ale i na hodnoty harmonicity, jitteru a shimmeru v těchto vokálech a několika nazálech. Všechny tyto kategorie se vztahují ke zdroji hlasu. Harmonicity²⁷ značí poměr tónových a šumových složek, a tak určuje míru periodicity v signálu. Vyjadřována bývá v decibelech, dosahovat může i záporných hodnot. Jitter signalizuje krátkodobou nestabilitu základní frekvence, shimmer variací amplitudy zvukové vlny. Těchto tří kategorií bývá nejčastěji využíváno ve výzkumu patologických rysů fonace (např. Shama, Krishna, & Cholayya, 2007; Wolfe & Fitch, 1995; Yumoto, Gould, & Baer 1982;), ale používají se i v automatických systémech pro rozpoznání mluvčího (Farrús, Hernando, & Ejarque, 2007).

Aby byla usnadněna manipulace se zvukovými záznamy, byly ke všem nahrávkám nejprve vytvořeny TextGridy, které byly olabelovány²⁸ podle nádechových úseků mluvčího. Zvuk i TextGridy pak byly na základě takto vytvořených hranic od sebe odděleny skriptem,

²⁵ Nečíselné hodnoty, tj. hodnoty, které program Praat vyhodnotil jako nedefinované.

²⁶ Zpravidla se měří první tři formanty všech samohlásek vyskytujících se v daném jazyce, přičemž od každé z nich bývá zvolen stejný počet výskytů. Z těchto hodnot se pak provede průměr – výsledné hodnoty jsou pak tři – průměrné F1, F2 a F3.

²⁷ Harmonics-to-noise-ratio, jinak HNR.

²⁸ Označeny textem.

a tak rozstříhány na menší části. V nich byly dalším skriptem vytvořeny zbylé dvě vrstvy (úroveň slov a úroveň hlásek), v nichž byly hranice slov a hlásek označeny prostřednictvím automatické hláskové segmentace (Pollák, Volín, & Skarnitzl, 2007). Ke každému projevu tak vzniklo množství úseků, ke všem z nich byl k dispozici TextGrid se třemi vrstvami - s automaticky olabelovanými segmenty v první vrstvě (a slovy v druhé). Protože však skript automaticky označující hranice hlásek vykazuje chyby, bylo třeba tyto hranice před prováděním další analýzy manuálně opravit. Oba nahrávané texty (maskovaný i nemaskovaný) však sestávají z více než 1000 hlásek (tj. více než 86 000 v celém zvukovém materiálu), a proto v rámci této diplomové práce nebylo možné kontrolovat a opravit všechny hranice segmentů. Z toho důvodu bylo k analýze vybraných parametrů vybráno osm slov z každého projevu - zatím, žena, rodinou, museli, setkat, osobně, řídí, nepřeje, u jejichž segmentů byly tyto hranice manuálně opraveny na základě doporučení Machače a Skarnitzla (2009).

Průměrné hodnoty vokálních formantů, jitteru, shimmeru a harmonicity byly získány z výše zmíněných osmi slov. Za účelem měření těchto parametrů²⁹ byly napsány a použity dva skripty, jeden pro vokály, jeden pro nazální konsonanty. U nazál byly měřeny pouze hodnoty jitteru, shimmeru a harmonicity.

5.4 Poslechový test

Součástí praktické části této práce bylo i provedení poslechového testu na základě získaného řečového materiálu. Ten byl sestaven za pomoci jednoduchého skriptu vytvořeného v programu Praat (Boersma & Weenink, 2012). Celkem obsahoval 21 položek, jimž předcházely dvě položky zácvičné. Všechny od sebe byly odděleny desenzitačním zvukem, pauzou a hlasovým projevem upozorňujícím na číslo právě testované položky. Každá položka se skládala z jedné maskované věty a tří vět nemaskovaných, úkolem posluchače bylo v původním nemaskovaném projevu rozpoznat mluvčího projevu maskovaného. Všichni cíloví mluvčí (kromě NKKF, která byla zahrnuta pro kontrolu úsudku³⁰, protože svůj hlas nedokázala změnit) byli do testu zahrnuti dvakrát, přičemž skladba mluvčích v druhé položce zůstávala ve všech případech stejná.

²⁹ Vokální formanty byly získány prostřednictvím Burgova algoritmu, který byl nastaven na frekvenční rozsah 0-5 kHz a na detekci maximálně pěti formantů.

³⁰ V případě, že by nebyla rozpoznána u většího množství posluchačů, nebyly by výsledky testu v práci podrobněji prezentovány.

Zkoumané věty sestávaly z 15 až 34³¹ slabik. Murray a Cort (1971 podle Hirson & Duckworth, 1993: 196) totiž naznačují, že 15 slabik ve větě představuje minimální požadavek pro identifikaci mluvčího posluchačem. Na druhou stranu bylo kvůli klesající schopnosti koncentrace člověka třeba, aby věty (a tím celý test) nebyly ani příliš dlouhé. Celkový poslechový test tak trval necelých 21 minut. Vždy byl přehrán jeden maskovaný hlas, po něm následovaly tři nemaskované. Každá položka byla přehrána dvakrát, aby měl posluchač dostatek času na své rozhodnutí. Mezi jednotlivými zvuky v položce byla vždy umístěna pauza o délce 5 sekund. V příloze č. 4 je pak možné vidět konkrétní úseky, které byly z nahraných textů použity. Tyto věty byly vybrány z textů tak, aby se v maskovaném a nemaskovaném projevu opakovalo co možná největší množství slov.

Mluvčí byli do testu vybíráni na základě tří předem stanovených kritérií: Bylo třeba zvolit zástupce ze všech tří skupin (tj. fonetici, nelingvisti, mluvčí bez maturity) a zařadit do testu i několik z těch, jejichž maskování bylo podle názoru autorky jedním z nejúspěšnějších. Poslední kritérium výběru představoval požadavek, aby bylo v testu zahrnuto více typů maskování hlasu. Nakonec bylo vybráno jedenáct mluvčích, pro kontrolu úsudku do nich byla zahrnuta i mluvčí (NKKF), která (jak již bylo zmíněno výše) svůj hlas změnit nedokázala.

Samotný test byl proveden na 16 studentech bakalářského studia oboru Fonetika, z nichž dva byli kvůli odlišnému mateřskému jazyku³² dodatečně vyřazeni. Nakonec tak byly vyhodnoceny výsledky pouze 14 posluchačů, 10 žen a 4 mužů. Test byl přehrán na počítači v malé výukové místnosti Fonetického ústavu FF UK, a to všem posluchačům najednou, bez použití sluchátek. Jednotlivé odpovědi byly zanášeny na záznamový arch, jež všichni posluchači obdrželi společně s instrukcemi k testu. Tyto instrukce je možné vidět v příloze č. 5.

³¹ Široké rozpětí slabik ve větách mělo několik důvodů, a to hlavně snahu nekrátit sémantické celky (resp. souvětí) a maximalizovat počet stejných slov v obou přehrávaných větách.

³² Jejich mateřským jazykem nebyla čeština, ale slovenština.

6 VÝSLEDKY

6.1 Poslechová analýza

Úkolem mluvčích bylo změnit svůj hlas tak, aby je nikdo nepoznal. K jeho splnění měli participanti neomezené množství možností, proto se použité strategie mezi sebou výrazně lišily. Tato kapitola poskytuje přehled jednotlivých skupin mluvčích a typů maskování, jež si jednotliví mluvčí zvolili, a poukazuje na některé forenzně relevantní rysy nalezené v maskovaných projevech.

6.1.1 Přehled použitých strategií maskování hlasu

Tabulka č. 4 poskytuje základní přehled maskování hlasu mluvčích v rámci první zkoumané skupiny, studentů a absolventů oboru Fonetika na FF UK. Každý řádek v tabulce vždy patří jednomu mluvčímu, uvedenému v prvním sloupci³³. Velké a tučně zdůrazněné písmeno X značí, že se jedná o hlavní (výraznou) změnu, malé písmeno x pak znamená, že změna není tolik nápadná nebo zřejmě nebyla bezprostředním záměrem mluvčího³⁴. Z tohoto přehledu je patrné, že jednou z nejoblíbenějších strategií použitých ve snaze o maskování hlasu v této skupině byla změna základní frekvence, nejčastěji její snížení. S jejím snížením pak někdy souvisí i použití třepené fonace. Tu si jako hlavní strategii vybral pouze mluvčí FJJM, ovšem v takovém rozsahu, že docílil nesrozumitelnosti projevu. Srozumitelnost projevu však byla jedinou podmínkou, která byla participantům uložena. Mluvčí FJJM ji tak nesplnil. Nejčastějšími změnami doprovázejícími hlavní strategii maskování pak byly změny tempa řeči (častěji docházelo k jeho zpomalení) a vyšší intonační variabilita v řeči.

Ostatní použité typy maskování se vyskytovaly v již o poznání menší míře než výše jmenované. Dvě mluvčí si zvolily chrapot, dva projevy byly denazalizované (ucpáním nosních dírek), dvakrát se objevila snaha o ledabylejší výslovnost. Jeden maskovaný hlas byl dyšný, jednou se dyšnost objevila jako důsledek hlavní strategie maskování – chrapotu. Další strategie, jako například vložení objektu do úst, položení objektu před ústa, nazalizace nebo krácení vokálů, byly použity jednou.

³³ První písmeno v jeho označení vždy představuje skupinu, do níž mluvčí spadá, poslední písmeno ukazuje, zda se jedná o ženu (F) či muže (M).

³⁴ Některé způsoby maskování hlasu mají za následek více nebo méně nezáměrné změny jiných parametrů v maskovaném projevu (Künzel, 2000: 173).

Mluvčí	Snížení F0	Zvýšení F0	Třepená fonace	Chrapot	Dyšný hlas	Nazalizace	Denazalizace	Objekt v ústech	Objekt před ústy	Napodobování jiného dialektu	Nižší intonační variabilita	Větší intonační variabilita	Snížení tempa řeči	Zvýšení tempa řeči	Napodobování řečové vady	Výraznější změna rytmu	Krácení vokálů	Dloužení vokálů	Zaokrouhlování rtů	Ledabytlejší výslovnost
FAAF		X										X		X						
FLLF	X		x										X							
FDDF		X							X			X		X						
FEEF	X			X									X							X
FEIF	X							X					X							
FIIF					X								X						X	
FJJF	X		x													X				
FJIF	X											X								
FLIF	X		x									X	X							
FMMF	(X) ³⁵		(X)	X	x								X					X		
FMIF							X				X		X		X	X				
FPPF	x					X						X								
FCCM	X											X	X							X
FJJM	x		X											X						X
FKKM							X				X					X	X		X	

Tabulka 4 Formy maskování hlasu použité studenty fonetiky (X značí hlavní a nápadnou změnu, x změnu méně výraznou)

³⁵ Tato mluvčí začíná maskování s nižší a třepenou fonací, která se postupně mění ve chrapot

Devět fonetiků použilo jednu hlavní taktiku za účelem zamaskování svého hlasu. Pět jich využilo dvou výraznějších strategií, zkombinována byla vysoká základní frekvence a umístění ruky před ústa (FDDF), snížení F0 a předmět v ústech (FEIF), ucpání nosních dírek a napodobování vady řeči (FMIF), snížení F0 a ledabylá výslovnost (FCCM) a denazalizace spolu s krácením vokálů (FKKM). Mluvčí FEEF zkombinovala dokonce tři výraznější změny – snížení F0, chrapot a ledabylou výslovnost – a její maskování bylo ze všech mluvčích nejúspěšnější.

Následující tabulka³⁶ předkládá souhrn četnosti výskytů jednotlivých typů maskování hlasu provedených fonetiky. Do tohoto přehledu byly zahrnuty jak hlavní a nápadnější strategie, tak i taktiky méně nápadné či ne zcela úmyslné. Z přehledu lze vidět, že k největšímu množství změn došlo na prozodické úrovni a ve způsobu fonace. Jazykové

Změna	Maskování	Četnost	Celkem
Fonace	Snížení F0	9	19
	Zvýšení F0	2	
	Třepená fonace	4	
	Chrapot	2	
	Dyšný hlas	2	
Rezonance	Nazalizace	1	4
	Denazalizace	2	
	Objekt před ústy	1	
Prozodie	Nižší intonační variabilita	2	22
	Vyšší intonační variabilita	6	
	Snížení tempa řeči	8	
	Zvýšení tempa řeči	3	
	Změny rytmu	3	
Artikulace a jazyk	Objekt v ústech	1	8
	Napodobování jiného dialektu	0	
	Napodobování řečové vady	1	
	Krácení vokálů	1	
	Dloužení vokálů	1	
	Zaokrouhlování rtů	1	
	Ledabylejší výslovnost	3	

Tabulka 5 Četnost jednotlivých typů maskování hlasu - fonetici

³⁶ Podoba tabulky je inspirována tabulkou v Künzelově studii (2000: 151), do výsledné podoby byla upravena autorkou. Jak je vidět již v teoretické části práce (viz kap. 4), jednoznačný návod, jak jednotlivé typy maskování zařazovat, neexistuje.

změny a změny v artikulaci společně se změnami v rezonancích³⁷ se naopak vyskytovaly nejméně.

Zvolenou strategii maskování hlasu ne každý mluvčí dokáže udržet po celou dobu projevu konzistentní (srov. např. Künzel, 2000: 151-154; Hollien, 2002: 50). Například mluvčí FDDF sice poměrně výrazně zvýšila základní frekvenci svého hlasu, ale nezvládla ji zachovat. F0 jejího hlasu v projevu kolísala a její průměrná hodnota postupně klesala – na konci promluvy byla znatelně nižší než na jejím začátku. Mluvčí FPPF se během svého projevu vrací téměř ke svému původnímu hlasu. U jiných fonetiků se problémy s udržením projevu odrazily ve způsobu dýchání (FJJF) nebo v mnohem rychlejším tempu (FJJM). Mluvčí FMIF svůj projev naopak během promluvy doplnila. Po necelé minutě začala napodobovat sigmatismus, ani pak v něm však nebyla důsledná.

Přestože se některé z hlavních strategií opakují, nachází se mezi úspěšností maskování jednotlivých mluvčích zásadní rozdíly. Někteří z nich jsou velmi snadno rozeznatelní už během prvního poslechu (např. FJJF, FLIF, FPPF, FCCM nebo FDDF), někteří svůj hlas maskují lépe (FAAF, FEIF, FK KM, FCCM). Mezi jednoznačně nejúspěšnější (na základě poslechu) lze zařadit mluvčí FEEF, FMIF a FJJM³⁸.

Tabulka č. 6 zobrazuje souhrn strategií maskování hlasu použitých druhou skupinou participantů tvořenou studenty a absolventy nelingvistických oborů vysokých škol. Stejně jako v předchozí skupině patří mezi nejoblíbenější metody manipulace základní frekvence hlasu (i když u mnoha participantů nepředstavuje zásadní či nápadnou změnu hlasu), větší intonační variabilita (na rozdíl od fonetiků ji tři mluvčí využili jako hlavní nebo jednu z hlavních strategií maskování) a změny v tempu řeči (častěji jeho zpomalení). Třepeň fonace doprovázela v této skupině pouze jeden z projevů. Častěji než v předchozí skupině se objevovala volba chrapotu či dyšného hlasu (mluvčí VKIF přechází až do šepotu). Ostatní typy maskování se vyskytovaly jen ve velmi nízkém počtu, jedna mluvčí si zakryla ústa mikinou (VMIF), jedna se uchýlila k ucpání nosních dírek (VPPF), jeden mluvčí celý projev rytmicky slabikoval, přičemž přízvuk pravidelně umisťoval na poslední slabiku rytmického celku (VJJM), jiný nápadně snížil rozpětí své základní frekvence (VTTM).

³⁷ Do změn v rezonancích však nebyl zahrnutý posun formantů vokálů, jenž by četnost výskytů v této oblasti mírně zvýšil (objevuje se např. u mluvčí FDDF).

³⁸ Mluvčí FJJM však nesplnil kritérium srozumitelnosti – v praxi by byl tento způsob maskování hlasu nereálný.

	Snížení F0	Zvýšení F0	Třepená fonace	Chrapot	Dyšný hlas	Nazalizace	Denazalizace	Objekt v ústech	Objekt před ústy	Napodobování jiného dialektu	Nižší intonační variabilita	Větší intonační variabilita	Snížení tempa řeči	Zvýšení tempa řeči	Napodobování řečové vady	Výraznější změna rytmu	Krácení vokálů	Dloužení vokálů	Zaokrouhlování rtů	Ledabyější výslovnost
VBBF		X										x		x						
VKKF		x										X								x
VHIF	x											X	x							
VKIF	x				X								x							
VMMF	X			x	x													x		
VMIF								X												
VPPF							X						x					x		
VFFM		X																		
VHIM		x		X								X			x					
VJJM	x															X	x			
VJIM			x	X	X								x							
VLLM	X												x						x	
VMMM		x		X								x		x						
VTTM											X									
VRRM		X										x								

Tabulka 6 Formy maskování hlasu použité nelingvisty

Třináct nelingvistů využilo jedné hlavní (nebo nápadnější) metody ke změně svého hlasu. Zbylí dva zkombinovali dvě výrazné techniky, mluvčí VIIM spojil chrapot s velmi vysokou intonační variabilitou (zcela se sžil s rolí vyděrače), mluvčí VJIM svůj hlas změnil na dyšný a chraplavý (místy až šeptaný).

Tabulka č. 7 udává souhrn jednotlivých typů maskování hlasu produkovaných nelingvisty. Podobně jako u předchozí skupiny mluvčích, i zde převažují změny ve způsobu fonace a změny na suprasegmentální úrovni. Druhé dvě skupiny změn jsou stejně jako u předchozí skupiny participantů zastoupeny nejméně.

Změna	Maskování	Četnost	Celkem
Fonace	Snížení F0	5	20
	Zvýšení F0	6	
	Třepená fonace	1	
	Chrapot	4	
	Dyšný hlas	3	
Rezonance	Nazalizace	0	2
	Denazalizace	1	
	Objekt před ústy	1	
Prozodie	Nižší intonační variabilita	1	15
	Vyšší intonační variabilita	6	
	Snížení tempa řeči	5	
	Zvýšení tempa řeči	2	
	Změny rytmu	1	
Artiklace a jazyk	Objekt v ústech	0	6
	Napodobování jiného dialektu	0	
	Napodobování řečové vady	1	
	Krácení vokálů	1	
	Dloužení vokálů	2	
	Zaokrouhlování rtů	1	
	Ledabylejší výslovnost	1	

Tabulka 7 Četnost jednotlivých typů maskování hlasu - nelingvisti

Podobně jako u fonetiků se i u některých nelingvistů projevila neschopnost udržet maskování konzistentní po celou dobu projevu. Mluvčí VLLM se postupně přibližuje svému původnímu hlasu, maskování však kolísá v průběhu celé promluvy. Kolísání během promluvy lze najít také u mluvčích VRRM či VMMF. Stejně tak je možné jej najít i u VFFM, který základní frekvenci svého hlasu nastavil příliš vysoko a nebyl schopen ji udržet, což se kromě jejího snižování projevilo i slyšitelným třasem jeho hlasu. Problémy s dýcháním vlivem

náročnosti použité strategie jsou znatelné u dvou mluvčích – VKIF a VPPF. U VPPF jsou tyto problémy způsobené zejména kombinací denazalizace a fonace tlačené přes hlasivky.

Nejúspěšněji v této skupině maskuje svůj hlas právě mluvčí VPPF, která kromě ucpaného nosu a sevřených hlasivek více otevírá ústa a prodlužuje některé vokály. K ní by mohla být přiřazena mluvčí FBBF, která po celou dobu projevu konzistentně udržuje zvýšenou základní frekvenci svého hlasu. O něco lépe než zbytek skupiny maskují svůj hlas také VMMM, VKIF, VJJM a VIIM. Naopak mluvčí VKKF, VIIF, VMIF, VTTM a VRRM patří k těm nejsnadněji rozeznatelným.

Techniky maskování provedené mluvčími z poslední skupiny participantů, tj. lidmi bez středoškolského vzdělání, uvádí tabulka č. 8. Stejně jako v obou předchozích skupinách k nejoblíbenějším patřila manipulace základní frekvence hlasu (10 mluvčích), ve srovnání s nimi však její zvýšení či snížení většinou nebylo příliš nápadné. Třepená fonace jako důsledek poklesu F0 se výrazněji projevila pouze jednou. Tři maskované promluvy byly charakteristické vyšší intonační variabilitou, ve dvou se vyskytla snaha zakrýt svoji identitu pomocí chrapotu. Ostatní strategie, jako denazalizace, vložení předmětů do úst, napodobování řečové vady či dlužení vokálů, byly použity jednou. Jedině v této skupině se jeden mluvčí uchýlil k napodobování jiného dialektu. Oproti fonetikům a nelingvistům mluvčí bez maturity nápadněji neměnili tempo řeči, temporální změny bylo možné slyšet u dvou z nich (jednou se jednalo o zpomalení, jednou o zrychlení řečového tempa).

Sedm mluvčích bez maturity k maskování svého hlasu využilo jednu hlavní strategii, jedna mluvčí sice snížila svoji základní frekvenci, ale vnímaný rozdíl je tak nepatrný, že tato změna nebyla značena velkým písmenem. Malým písmenem byly značeny i techniky užití mluvčím NTTM. Jeho hlavní strategií bylo totiž větší napětí hlasivek, které bylo doplněné o snížení F0 a nazalizaci některých hlásek. Tři lidé z této poslední skupiny zkombinovali dvě výraznější techniky, mluvčí NDDM dlužil vokály a od nemaskovaného projevu se odlišoval i intonačně, NPPM důsledně napodoboval rotacismus a rotacismus bohemicus a do projevu vložil emoce, což se odrazilo v jeho intonaci (i základní frekvenci jeho hlasu). Mluvčí NVVM zvýšil svoji základní frekvenci a rozšířil svůj intonační rozsah. Tomuto mluvčímu se také jako jedinému z celého vzorku zcela podařilo změnit kvalitu svého hlasu. U dvou participantů z této skupiny nejsou v tabulce zaznamenány žádné z uvedených změn. Maskovaný a nemaskovaný projev mluvčího NJIM byl téměř totožný, mluvčí NKKF svůj hlas nebyla schopna změnit vůbec.

	Snížení F0	Zvýšení F0	Třepená fonace	Chrapot	Dyšný hlas	Nazalizace	Denazalizace	Objekt v ústech	Objekt před ústy	Napodobování jiného dialektu	Nižší intonační variabilita	Větší intonační variabilita	Snížení tempa řeči	Zvýšení tempa řeči	Napodobování řečové vady	Výraznější změna rytmu	Krácení vokálů	Dloužení vokálů	Zaokrouhlování rtů	Ledabylejší výslovnost
NKKF																				
NMMF	x																			
NSSF	X		x						x											
NAAM		x								X			x							
NDDM		x										X				x		X		
NJJM		x		X																
NJIM																				
NKKM							X													
NPPM		x										X		x	X					
NSSM		x						X												
NTTM	x																			
NTIM	x			X																
NVVM		X										X								

Tabulka 8 Formy maskování hlasu použité mluvčími bez maturity

Tabulka č. 9 poskytuje přehled četnosti výskytů jednotlivých typů maskování řeči provedených mluvčími bez maturity. Na první pohled je patrné, že mluvčí bez maturity využili mnohem menšího počtu technik než obě skupiny předchozí. Nejvíce byly zastoupeny fonační změny, které následovaly změny prozodické (jichž je ovšem již mnohem méně než tomu bylo v předešlých dvou skupinách). Zbylé dvě skupiny změn se objevovaly jen minimálně.

Změna	Maskování	Četnost	Celkem
Fonace	Snížení F0	4	13
	Zvýšení F0	6	
	Třepená fonace	1	
	Chrapot	2	
	Dyšný hlas	0	
Rezonance	Nazalizace	1	3
	Denazalizace	1	
	Objekt před ústy	1	
Prozodie	Nižší intonační variabilita	0	6
	Vyšší intonační variabilita	3	
	Snížení tempa řeči	1	
	Zvýšení tempa řeči	1	
	Změny rytmu	1	
Artiklace a jazyk	Objekt v ústech	1	4
	Napodobování jiného dialektu	1	
	Napodobování řečové vady	1	
	Krácení vokálů	0	
	Dloužení vokálů	1	
	Zaokrouhlování rtů	0	
	Ledabylejší výslovnost	0	

Tabulka 9 Četnost jednotlivých typů maskování hlasu - mluvčí bez maturity

I v této skupině měli mluvčí problémy s udržení konzistentního maskování během celé promluvy. Například u mluvčího NKKM míra denazalizace projevu značně kolísá, mluvčí NTTM na jednom místě svoji techniku také neudržel a téměř se vrátil ke svému původnímu hlasu.

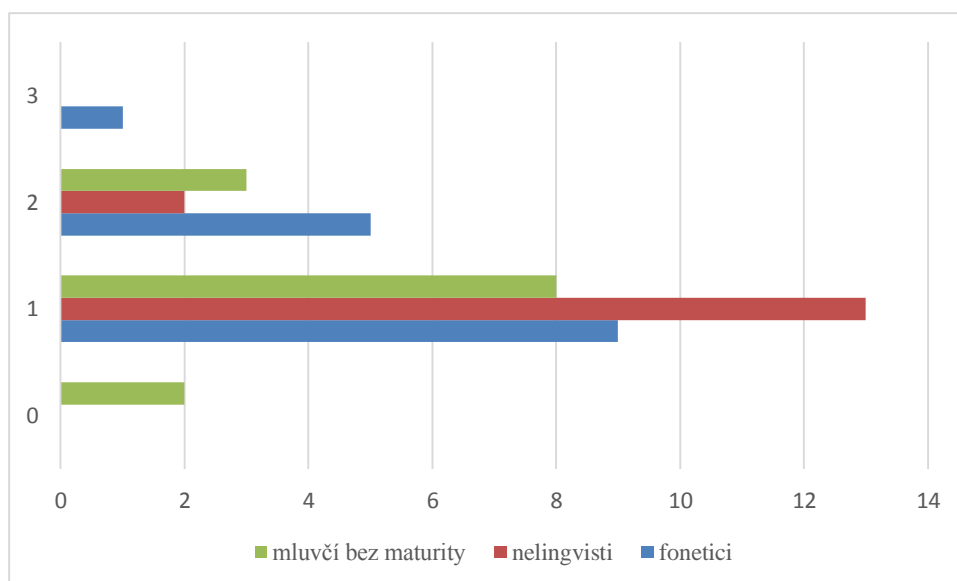
Nejúspěšněji svůj hlas v této skupině maskoval jednoznačně mluvčí NVVM, jenž zcela změnil barvu svého hlasu (původně hluboký a sytý hlas se změnil ve vyšší a „plošší“). Ke schopnějším by pak bylo možné ještě zařadit mluvčí NPPM a NTTM. U většiny mluvčích v této skupině by se daly jejich snahy o maskování hlasu označit za neúspěšné, mluvčí NKKF

svůj hlas jako jediná nezměnila vůbec, u dalších dvou (NJIM, NMMF) byla změna jen nepatrná.

Celkově je z uvedených přehledů možné pozorovat, že nejoblíbenější technikou maskování byly fonační změny, což odpovídá i výsledkům předchozích studií (srov. např. Masthoff, 1996 nebo Masthoff, 2000). Z fonačních změn převažovalo zvýšení a snížení základní frekvence hlasu, ale poměrně často se objevoval také chrapot a dyšná fonace, u fonetiků ve větší míře docházelo i k fonaci třepené. Šepot, který bývá v literatuře často označován jako jedna z oblíbených strategií (viz např. Künzel, 2000: 149), k maskování svého hlasu nikdo nepoužil³⁹. Další frekventovanou skupinou změn byly změny prozodické. Jejich četnost však významně ovlivnil výběr textů, pro maskovaný projev bylo využito autentického textu, čímž bylo participantům umožněno snadnější ztotožnění se s rolí vyděrače. S tím pak souvisí i častější změny v intonaci a v prvních dvou skupinách také doprovodné změny temporální. Také způsob maskování ovlivnil některé z prozodických faktorů výsledné řeči, například denazalizace ve dvou případech vedla k větší monotónnosti projevu (FKKF, FMIF), ve třech pak snížila tempo řeči (FKKF, FMIF, VPPF). V rámci temporálních změn bylo celkově možné častěji pozorovat snížení tempa řeči, které je většinou způsobeno větší potřebou koncentrace a artikulačních snah (Künzel, 2000: 173). Z dalších typů maskování hlasu, které bývají v literatuře uváděny, byla čtyřikrát použita denazalizace, třikrát si mluvčí položil objekt před ústa, dvakrát si vložil předmět do úst. Dva mluvčí záměrně napodobovali řečové vady (u dalšího došlo k řečovým vadám pravděpodobně díky stylu projevu, artikulace konsonantů je zasažena také v případech vkládání předmětů do úst), jeden jiné nářečí. Jen málo mluvčích se uchýlilo ke zkracování či prodlužování vokálů nebo výrazné změně rytmu řeči. Díky povaze druhého textu se třikrát objevila i nápadnější tendence směrem k ledabylejší (a hovorovější) výslovnosti oproti projevu nemaskovanému. O napodobování cizineckého přízvuku nebo imitaci jiného člověka se nepokoušel nikdo.

K maskování hlasu byla nejčastěji využita jedna hlavní technika, několikrát se objevila kombinace dvou hlavních strategií, tři zkombinovala pouze jedna mluvčí. I tento výsledek je ve shodě s dříve provedenými výzkumy (srov. např. Braun, 2006; Masthoff, 1996; Masthoff, 2000). Rozložení počtu hlavních technik maskování hlasu napříč všemi zkoumanými skupinami ilustruje graf na obrázku č. 1.

³⁹ Šepotem je zde myšlena řeč za nepřítomnosti základní frekvence hlasu.



Obrázek 1 Počet hlavních technik použitých k maskování hlasu napříč všemi skupinami mluvčích

Podobně jako v předchozích studiích (srov. např. Masthoff, 2000) bylo zjištěno, že mluvčí často nejsou schopni konzistentně udržet maskování svého hlasu a často se v průběhu projevu navrací směrem ke svému původnímu řečovému projevu. Tento jev byl zpozorován ve všech třech zkoumaných skupinách.

Úspěšnost výsledného maskování hlasu se u jednotlivých mluvčích lišila, většina z nich však byla již při prvním poslechu snadno identifikovatelná a maskování hlasu tak nebylo dostatečně účinné (srov. Masthoff, 2000; Zhang & Tan, 2008). Během poslechu byl zpozorován jeden z aspektů, které měly vliv na schopnost maskování hlasu, a to zdatnost ve čtení. Ti, již se při hlasitém čtení potýkali s většími problémy⁴⁰, zpravidla také nedokázali svůj projev dostatečně zamaskovat či maskování hlasu delší dobu udržet. Vzdělanostní faktor se však neukázal být tak významným, jak bylo na počátku výzkumu předpokládáno. Mezi třemi zkoumanými skupinami sice lze najít rozdíly (podle očekávání maskovali fonetici svůj hlas lépe, zatímco v poslední skupině bylo několik mluvčích, kteří svůj hlas nezměnili téměř vůbec), mezi ty, kterým se svoji identitu podařilo zdařile zamaskovat, však byli zařazeni mluvčí ze všech tří zkoumaných skupin (z fonetiků FEEF, FMIF, FJJF, z nelingvistů VPPF a VBBF, z mluvčích bez maturity NVVM).

⁴⁰ Vzhledem k obtížnosti textu se menší problémy objevovaly u většiny mluvčích, zde jsou však na mysli obtíže, které s náročností textu nesouvisí.

6.1.2 Některé forenzně relevantní rysy nalezené v projevech

Kromě schopnosti maskovat svůj hlas hrají při identifikaci mluvčího stejně jako u nemaskované řeči důležitou roli i charakteristické znaky jeho hlasu. Kromě hlasové kvality se na tomto faktu podílí i některé další faktory, jako například dialekt mluvčího či jeho řečové vady, které se projeví i v některých projevech nahrávaných v rámci této práce.

Dialekt mluvčího představuje důležitý aspekt ve forenzní praxi (Jessen, 2008: 690). Není neobvyklé, že i přes snahu o maskování hlasu nářečí v řeči zůstává. Přestože participantů tohoto výzkumu četli předem daný text (který tak zredukoval možnost výskytu dialektu v lexiku a syntaxi), u některých z nich se výrazné nářeční rysy udržely na fonetické úrovni. U mnoha mluvčích bylo takovým výrazným nářečním rysem široké plzeňské /a:/ (značná část participantů pocházela právě z plzeňského kraje). Objevilo se u všech tří nahrávaných skupin - z fonetiků svůj plzeňský původ nezapřela mluvčí FEIF, u nelingvistů bylo nejvíce znatelné v projevu mluvčího VTTM. Nejčastěji se tento charakteristický nářeční rys objevoval v poslední skupině (většina mluvčích byla z Plzně), například u NDDM či NSSF. Toto plzeňské /a:/ se dokonce velmi výrazně projevilo i v projevu mluvčího NAAM, jehož hlavní strategií maskování bylo napodobování dialektu z oblasti Moravy. V maskovaných promluvách bylo možné zaslechnout i další nářeční prvky, jako třeba asimilaci před jedinečnými hláskami (ve skupině studentů a absolventů nelingvistických oborů; např. mluvčí VKIF) či otevřenější výslovnost zavřených vokálů.

Další forenzně relevantní (a ještě nápadnější a hůře odbouratelnější) charakteristikou jsou řečové vady, které se také vyskytovaly u mluvčích ze všech zkoumaných skupin. Z těch participantů, jejichž řečové vady byly výraznější, nebyl nikdo zařazen mezi nejúspěšnější - co se maskování hlasu týče. Z řečových vad se objevoval rotacismus, rotacismus bohemicus a slabší sigmatismus, někdy byly zasaženy i afrikáty nebo prealveolární explozívy. U mluvčí FDDF se pravděpodobně objevuje porucha funkce véla.

Jako důležitý identifikátor může být ve forenzní fonetice využit i způsob dýchání. Velmi specifický (až patologický) způsob dýchání je možné najít v obou projevech mluvčí NKKF. Plynulost promluvy je narušena neschopností vystačit s dechem - často jí dochází dech uprostřed slova, udělá pauzu, nadechne se a v projevu pokračuje dál.

6.2 Akustická analýza

6.2.1 Základní frekvence

V Tabulce č. 10⁴¹ jsou zachyceny hodnoty F0 první skupiny participantů, studentů vyšších ročníků a absolventů oboru Fonetika na FF UK. Jak lze vidět, ke zvýšení základní frekvence v maskovaných projevech se uchýlili čtyři z nich, tři ženy a jeden muž, ke snížení osm, z toho jeden muž a sedm žen. Téměř stejná F0 u obou nahrávek se objevila jen u dvou žen z celého vzorku (u FIIF rozdíl v alternativní základní hladině⁴² 1,36 Hz, u FPPF rozdíl 6,33 Hz; průměrná základní frekvence i její medián u obou mluvčích nepřekročily rozdíl 4 Hz mezi oběma typy promluv). Minimální změnu F0 lze pozorovat i u mluvčího FCCM, u něž se základní hladina obou promluv lišila o 6,81 Hz. Z předložené tabulky je možné dále vyčíst, že míra snížení či zvýšení F0 se ve snaze o maskování hlasu mezi jednotlivými mluvčími nápadně liší, rozdíly se pohybují od cca 9 až do 106 Hz (resp. 0,88 až 15,86 púltónu). Největší odchylky v základní hladině jsou pozorovatelné u čtyř mluvčích - žen (FDDF, FEEF, FLLF, FMMF), tři z nich svoji F0 nápadně snížily, jedna zvýšila. Na základě výsledků průměrné F0 a jejího mediánu by se k této skupině řadila i mluvčí FAAF, základní hladina však ukázala rozdíl pouhých 23,3 Hz. Podobnou situaci lze pozorovat například i u FEEF, kde se průměrná základní frekvence a medián v obou projevech liší pouze o cca 30 Hz, zatímco základní hladina o 106,8 Hz (což je rovno 15,86 púltónům, tedy více než jedné oktávě). Odchylkami v těchto hodnotách se tato práce nebude podrobněji zabývat, za relevantní údaj je zde považována základní hladina F0 (kontura F0 nebyla opravována, a tak je základní hladina robustnější). Pokud jde o hodnoty udávané v púltónech, výraznější rozdíl mezi čteným a maskovaným projevem se vyskytl i u mluvčího FJJM – základní hladina se snížila o 7,99 púltónu, tedy více než půl oktávy.

⁴¹ V tabulkách s výsledky první řádek vždy zobrazuje hodnoty naměřené u normálního čteného projevu daného mluvčího, druhý řádek pak hodnoty u jeho maskovaného hlasového projevu. Průměrná základní frekvence, její medián, ale i směrodatná odchylka (značena SD) však mohou být ovlivněny chybami – např. oktávovými skoky, které v promluvách nebyly opravovány. I z tohoto důvodu se zaměřím zejména na pátý sloupec, alternativní základní hladinu navrhanou Lindhem a Erikssonem (2007), která by měla být v tomto ohledu přesnější. Poslední dva sloupce v tabulkách pak zobrazují rozdíl mezi základní hladinou obou promluv, a to v hertzech a púltónech.

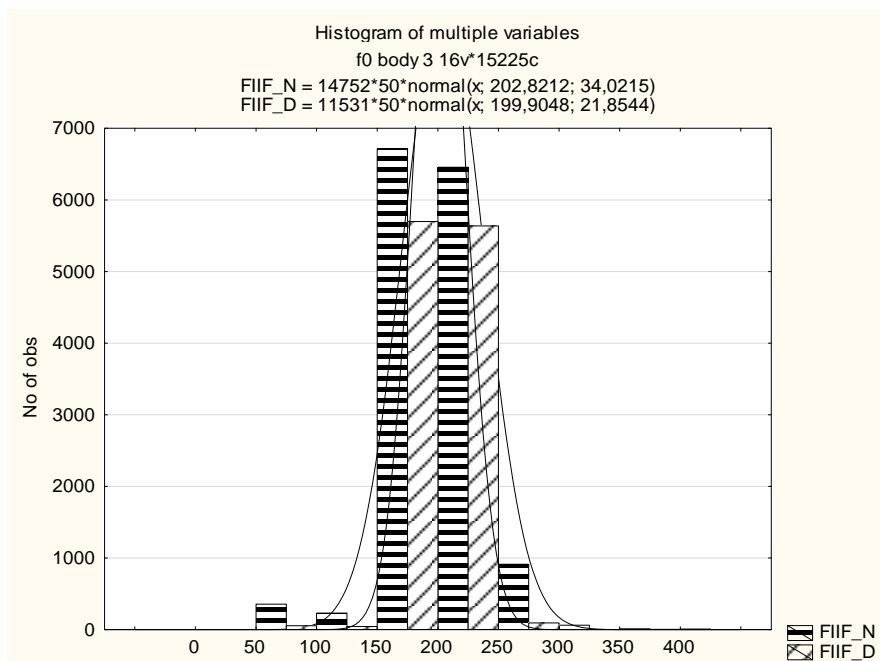
⁴² Dále jen základní hladina.

Mluvčí	Průměr	Medián	SD	Alt_base	Rozdíl Alt_B	Rozdíl v půltónech
FAAF_N	219,43	219,29	27,40	188,31		
FAAF_D	282,10	279,08	49,61	211,61	-23,3	-2,02
FCCM_N	114,76	109,54	15,34	99,37		
FCCM_D	102,27	98,69	17,37	92,56	6,81	1,23
FDDF_N	196,45	193,03	27,79	161,04		
FDDF_D	315,31	317,74	74,10	220,23	-59,19	-5,42
FEEF_N	214,69	210,37	32,19	178,05		
FEEF_D	172,98	188,94	57,06	71,25	106,80	15,86
FEIF_N	224,51	225,51	36,64	187,38		
FEIF_D	213,06	212,16	32,39	178,10	9,28	0,88
FIIF_N	203,09	200,52	34,22	171,81		
FIIF_D	199,92	199,91	22,19	173,17	-1,36	-0,14
FJIF_N	222,78	222,73	37,21	173,66		
FJIF_D	184,44	184,18	27,17	149,33	24,33	2,61
FJJF_N	203,48	202,82	24,94	174,25		
FJJF_D	176,09	177,12	24,74	155,01	19,24	2,03
FJJM_N	133,96	135,07	18,8	105,48		
FJJM_D	106,51	107,15	41,73	66,48	39,00	7,99
FKKM_N	123,12	120,46	20,64	96,83		
FKKM_D	131,13	131,09	18,61	106,41	-9,58	-1,63
FLIF_N	218,00	209,49	44,84	166,46		
FLIF_D	171,21	170,98	27,31	143,82	22,64	2,53
FLLF_N	215,03	214,29	36,10	171,37		
FLLF_D	158,59	163,48	32,01	86,43	84,94	11,85
FMIF_N	213,95	214,75	29,74	181,21		
FMIF_D	232,75	234,49	20,48	202,06	-20,85	-1,89
FMMF_N	226,71	227,10	33,28	185,45		
FMMF_D	178,82	180,97	69,79	83,99	101,46	13,71
FPPF_N	230,13	228,45	38,45	187,84		
FPPF_D	228,20	231,69	35,22	181,51	6,33	0,59

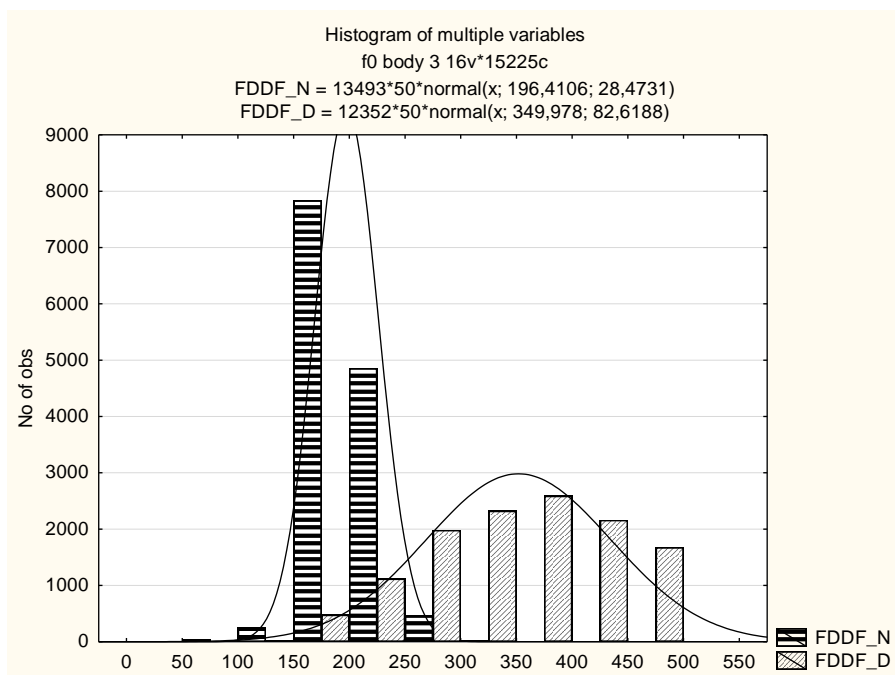
Tabulka 10 Hodnoty F0 získané z projevů 1. skupiny mluvčích - fonetiků

Následující histogramy (na obr. č. 2 a 3), založené na hodnotách základní frekvence získaných z celých záznamů v krocích po 0,005 s, slouží k lepší představě rozdílů F0 mezi oběma projevy u vybraných participantů⁴³. Na prvním z nich je možné vidět srovnání projevů mluvčí FIIF, které se mezi sebou v rozložení základní frekvence lišily jen minimálně, druhý představuje F0 mluvčí FDDF, v němž již lze vidět nápadné odchylky v hodnotách získaných z obou nahrávek.

⁴³ Všechny histogramy založené na bodech F0 viz příloha č. 6.



Obrázek 2 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí FIIF



Obrázek 3 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí FDDF

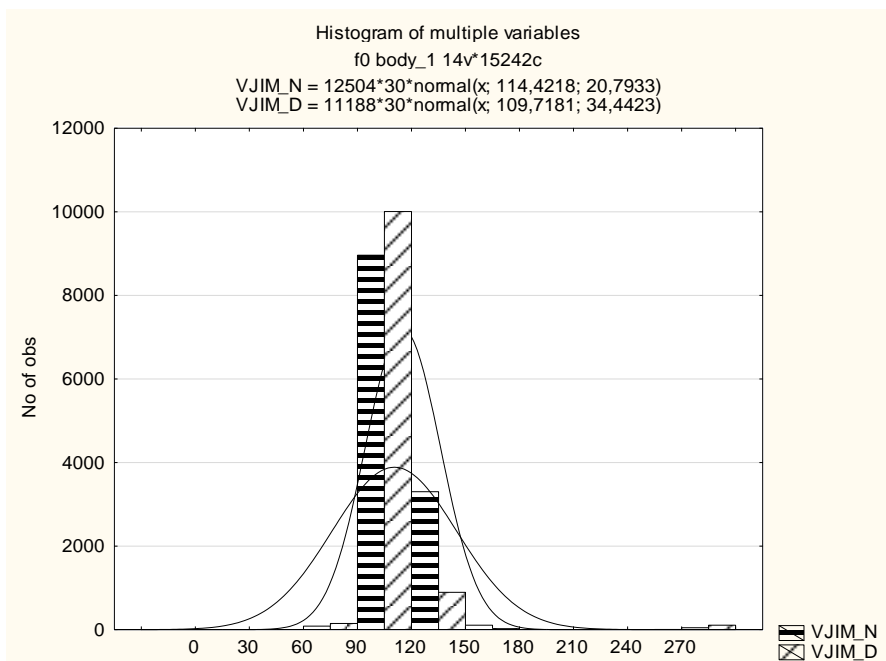
Tabulka č. 11 zobrazuje výsledné hodnoty F0 druhé skupiny participantů, 15 studentů a absolventů nelingvistických oborů vysokých škol. V maskovaných projevech svoji základní frekvenci zvýšilo šest mluvčích, čtyři muži a dvě ženy, snížilo šest, čtyři ženy a dva muži. Přibližně stejná F0 byla naměřena u tří zbývajících participantů, dvou mužů (rozdíl v základní hladině nepřekročil hranici 3 Hz) a jedné ženy (základní hladina maskovaného projevu byla o 6,17 Hz vyšší). Stejně jako u předchozí skupiny se rozdíly ve změnách F0 mezi mluvčími

nápadně liší, pohybují se v rozmezí od cca 12 až do 77 Hz (resp. 1,48 až 10,29 púltónu). Největší odchylka mezi oběma projevy se objevila u VFFM, který zvýšil svoji průměrnou základní frekvenci o 77 Hz (10,29 púltónu), mluvčí VBBF dosáhla rozdílu 60 Hz (5,40 púltónu) mezi oběma typy promluv. Z hlediska rozdílu naměřeného v púltónech by k této skupině mohli být přiřazeni ještě VMMM a VIIM (5,53 a 5,20 púltónu). Z výsledků lze vidět, že nikdo z této skupiny nezměnil svoji F0 ve stejné míře jako někteří ze skupiny první.

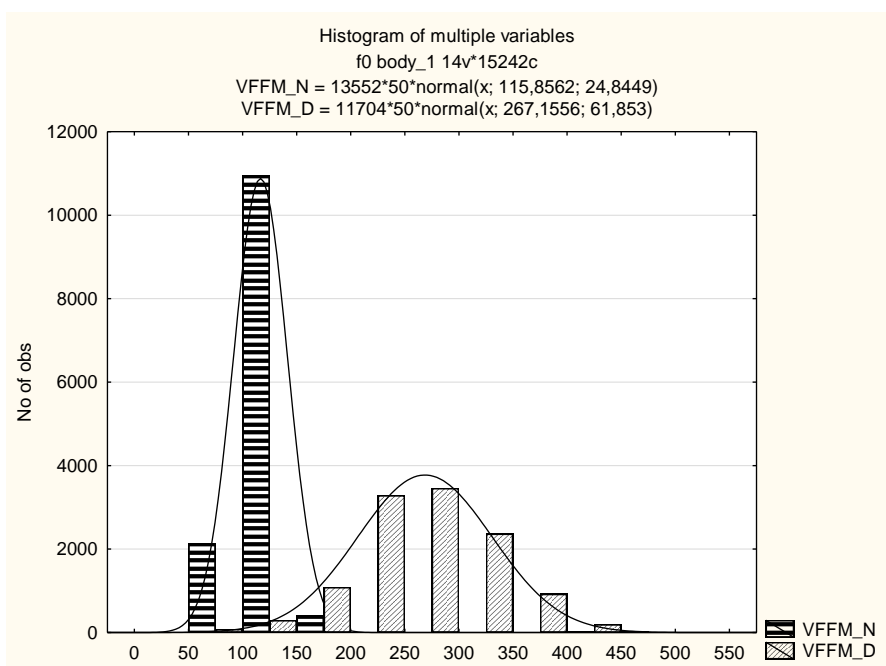
Mluvčí	Průměr	Medián	SD	Alt_base	Rozdíl Alt_B	Rozdíl v púltónech
VBBF_N	198,87	196,84	30,17	161,10		
VBBF_D	319,45	320,65	76,28	220,08	-58,98	-5,40
VFFM_N	114,75	112,5	18,14	95,07		
VFFM_D	264,33	263,29	65,85	172,28	-77,21	-10,29
VIIF_N	225,69	225,71	36,17	178,94		
VIIF_D	250,44	251,23	65,70	148,39	30,55	3,24
VIIM_N	136,68	133,83	25,72	114,12		
VIIM_D	223,11	224,60	51,38	154,09	-39,97	-5,20
VJIM_N	113,88	112,01	19,77	95,97		
VJIM_D	107,82	104,90	20,15	96,20	-0,23	-0,04
VJJM_N	126,56	126,47	17,41	105,16		
VJJM_D	134,13	133,00	12,45	121,53	-16,37	-2,50
VKIF_N	238,06	237,43	33,55	196,01		
VKIF_D	239,44	249,06	42,05	164,14	31,87	3,07
VKKF_N	192,28	192,19	31,84	148,70		
VKKF_D	242,04	246,18	50,65	168,02	-19,32	-2,11
VLLM_N	115,63	113,96	16,17	99,42		
VLLM_D	100,15	99,04	13,39	86,79	12,63	2,35
VMIF_N	201,79	201,46	27,22	174,76		
VMIF_D	220,59	223,09	27,98	180,93	-6,17	-0,60
VMMF_N	216,71	216,85	37,70	172,39		
VMMF_D	184,56	182,33	21,81	158,22	14,17	1,48
VMMM_N	123,09	122,90	19,70	98,66		
VMMM_D	129,20	134,04	32,73	71,70	26,96	5,53
VPPF_N	191,02	184,03	41,26	147,74		
VPPF_D	157,94	156,92	22,82	135,65	12,09	1,48
VRRM_N	118,61	116,26	20,20	101,63		
VRRM_D	147,98	150,60	21,59	118,07	-16,44	-2,60
VTTM_N	130,26	125,66	23,20	108,34		
VTTM_D	115,51	115,33	14,58	106,01	2,33	0,38

Tabulka 11 Hodnoty F0 získané z projevů 2. skupiny mluvčích - nelingvistů

Následující obrázky (č. 4, 5) opět ilustrují mluvčí s nízkým (VJIM) a vysokým (VFFM) rozpětím bodů F0 mezi oběma nahrávkami.



Obrázek 4 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí VJIM



Obrázek 5 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí VFFM

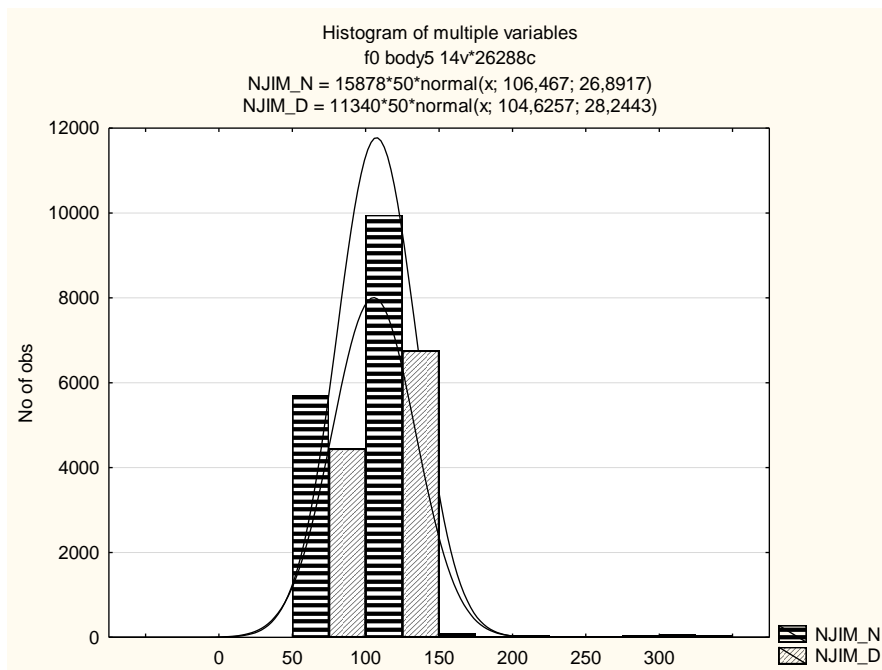
Tabulka č. 12 odkazuje na poslední nahrávanou skupinu participantů, a to lidí bez maturity. Z celkových třinácti svoji F0 zvýšili čtyři mluvčí, všichni muži. Snížení základní frekvence jako součást maskování svého hlasu zvolili také čtyři lidi, dva muži a dvě ženy. Na podobné či téměř stejné F0 v maskovaném projevu zůstali tři participant, rozdíl mezi hodnotami u obou promluv nebyl vyšší než 5 Hz. Zbylí dva pak svoji základní frekvenci v druhém projevu změnili jen minimálně (6,33 a 7,75 Hz). Odchylny F0 v maskovaném

projevu oproti normálnímu čtení se pohybují do 40 Hz (resp. od 1,78 do 7,58 půltónu), což je méně než v obou předchozích skupinách.

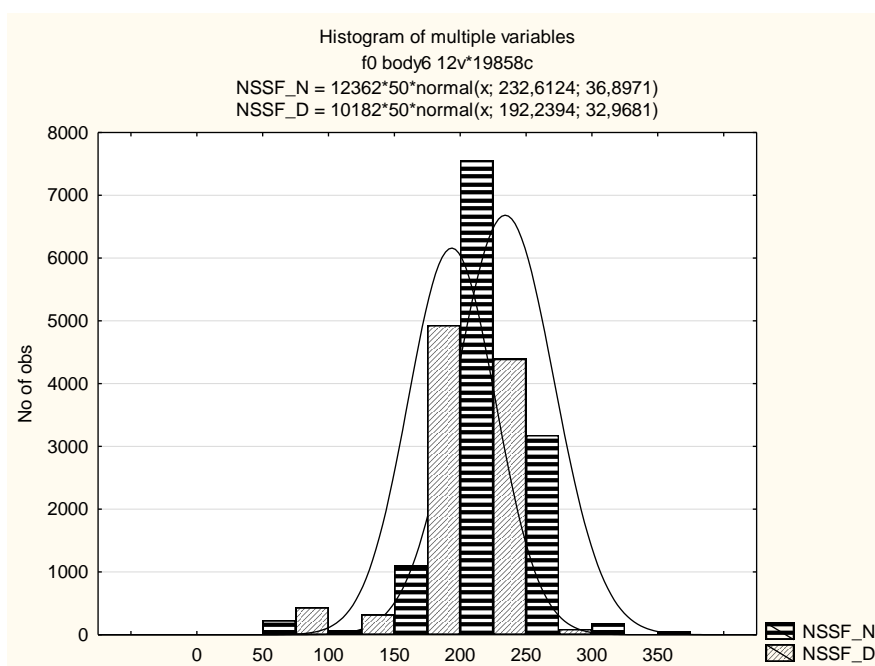
Mluvčí	Průměr	Medián	SD	Alt_base	Rozdíl AB	Rozdíl v půltónech
NAAM_N	113,83	112,27	19,92	88,41		
NAAM_D	118,98	118,98	17,14	98,03	-9,62	-1,79
NDDM_N	161,31	162,31	34,60	117,02		
NDDM_D	180,97	184,30	27,36	140,84	-23,82	-3,21
NJIM_N	106,00	103,27	24,63	87,89		
NJIM_D	105,22	102,23	25,25	89,38	-1,49	-0,29
NJJM_N	130,07	131,40	19,69	106,72		
NJJM_D	147,11	152,70	30,16	85,92	20,8	3,75
NKKF_N	166,94	166,59	26,52	139,81		
NKKF_D	167,01	165,04	28,54	138,14	1,67	0,21
NKKM_N	113,59	107,62	34,48	97,45		
NKKM_D	128,02	126,91	24,50	101,49	-4,04	-0,70
NMMF_N	198,51	197,31	29,33	167,42		
NMMF_D	196,35	197,14	36,75	151,10	16,32	1,78
NPPM_N	123,71	122,32	25,18	101,27		
NPPM_D	162,92	165,90	26,98	122,46	-21,19	-3,29
NSSF_N	232,60	233,07	36,83	193,42		
NSSF_D	192,09	197,31	35,21	153,88	39,54	3,96
NSSM_N	134,27	130,33	39,25	101,58		
NSSM_D	139,36	139,63	26,81	107,91	-6,33	-1,05
NTIM_N	124,89	121,14	30,96	102,70		
NTIM_D	128,88	124,47	63,42	66,27	36,43	7,58
NTTM_N	131,91	130,24	22,64	110,63		
NTTM_D	129,14	125,07	33,14	102,88	7,75	1,26
NVVM_N	102,21	96,79	29,56	78,24		
NVVM_D	155,26	155,87	32,87	108,66	-30,42	-5,69

Tabulka 12 Hodnoty F0 získané z projevů 3. skupiny mluvčích - bez maturity

Obrázky 6 a 7 opět předkládají příklady menších (NJIM) a větších (NSSF) rozdílů mezi rozložením bodů naměřené základní frekvence v obou typech projevů ve třetí skupině.



Obrázek 6 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí NJIM



Obrázek 7 Srovnání rozložení bodů F0 v normálním a maskovaném projevu - mluvčí NSSF

Za zmínku stojí i poslední hodnota uvedená ve výsledcích - směrodatná odchylka (přestože stejně jako výsledky průměrné F0 a jejího mediánu může být ovlivněna oktávovými skoky). Ve většině případů stoupá směrodatná odchylka tam, kde dochází k nárůstu průměrné základní frekvence; naopak při jejím poklesu se její hodnota snižuje. Z uvedených tabulek je však zřejmé, že se na její výsledné hodnotě podílejí i další parametry, z nichž nejdůležitější

představují intonační variabilita a monotónnost projevu. Někteří mluvčí se v rámci strategie maskování zaměřili zejména na větší intonační rozpětí, jež se pak projevilo ve výrazném zvýšení směrodatné odchylky (například mluvčí VKKF v druhé skupině). Naopak denazalizace měla za následek vyšší monotónnost celého projevu, což se odrazilo i ve výsledné hodnotě směrodatné odchylky (ve třech případech ze čtyř i přes mírně zvýšenou F0 směrodatná odchylka klesla).

Preference snížení či zvýšení F0 v maskovaných projevech se ukázaly jako relativně vyrovnané. V první skupině svoji základní frekvenci zvýšili 4 mluvčí, dalších 8 snížilo, ve druhé skupině jich F0 zvýšilo i snížilo 6 a ve třetí byl poměr zvýšení a snížení F0 5 ku 5. Celkově se tak k poklesu základní frekvence svého hlasu jako jednoho z prostředků maskování či jeho doprovodného jevu uchýlilo 19 mluvčích, 15 k jejímu zvýšení. Z celkového počtu 21 mužů svoji F0 zvýšilo 9 mužů, 5 naopak. Z 22 žen 5 z nich F0 zvýšilo, 13 snížilo. Tyto výsledky sice nepotvrzují zjištění prezentovaná Masthoffem (1996), že muži častěji tíhnou ke zvýšení F0, zatímco ženy ke snížení, zároveň ji však zcela nevyklučují. Künzelovu (2000) domněnku, že lidé s vyšší průměrnou F0 mají při maskování tendenci k jejímu zvýšení a naopak lidé s nižší F0 k jejímu snížení, získaná data nepodporují vůbec.

Pozorovatelný rozdíl mezi třemi skupinami mluvčích se objevil zejména v rozsahu, v jakém někteří z nich svoji základní frekvenci měnili. Rozdíl F0 nad 30 Hz v maskovaném projevu oproti normálnímu se ukázal u pěti fonetiků (čtyř žen a jednoho muže), pěti nelingvistů (tři ženy, dva muži) a tří lidí bez maturity (jedna žena, dva muži), celkový rozsah změny byl ale nejvyšší u fonetiků, následovaný nelingvisty. Podobné výsledky se ukázaly i při hodnocení rozdílů v rámci měření půltónů, o více než pět půltónů snížilo či zvýšilo svoji základní frekvenci pět fonetiků (dvě mluvčí zvýšily svůj hlas o více než jednu oktávu), čtyři nelingvisti a dva mluvčí bez maturity. Vzhledem k nízkému počtu participantů však není jasné, do jaké míry jsou tyto výsledky pouze náhodné, nicméně tato otázka může být předmětem dalšího výzkumu v oblasti maskování hlasu.

6.2.2 Vokální formanty

Tabulka č. 13 zobrazuje hodnoty formantů⁴⁴ získané z projevů studentů a absolventů oboru Fonetika. Obrázky, jež následují (č. 8, 9, 10) pak pro lepší orientaci zobrazují interval spolehlivosti spolu s průměrem získaných hodnot z jednotlivých projevů, zvláště pro hodnoty F1, F2 a F3. U mluvčí FDDF je možné vidět statisticky významný rozdíl⁴⁵ v hodnotě prvního vokálního formantu, který byl ovlivněn způsobem maskování hlasu, jež si zvolila. Zvýšení základní frekvence totiž zpravidla souvisí se zdvižením hrtanu, které zkrátí vokální trakt mluvčího, což následně vede i ke zvýšení průměrné hodnoty prvního formantu (srov. např. Thurgood, 2000: 297). Zvýšení prvního formantu oproti nemaskovanému projevu se objevuje i u mluvčího FJJM. I tato změna by mohla souviset s typem zvolené techniky maskování hlasu. Třepeň fonace totiž může vést ke zdvižení hrtanu vlivem napjatosti hlasivek, čímž dochází právě ke zvýšení prvního formantu (tamtéž). Pokles hodnoty prvního formantu lze pozorovat u mluvčí FIIF. Ta k zamaskování svého hlasu jako hlavní strategii použila dyšný hlas, který často koreluje se snížením hrtanu – a tudíž i prodloužením vokálního traktu, který má vliv na snížení hodnoty prvního formantu (tamtéž).

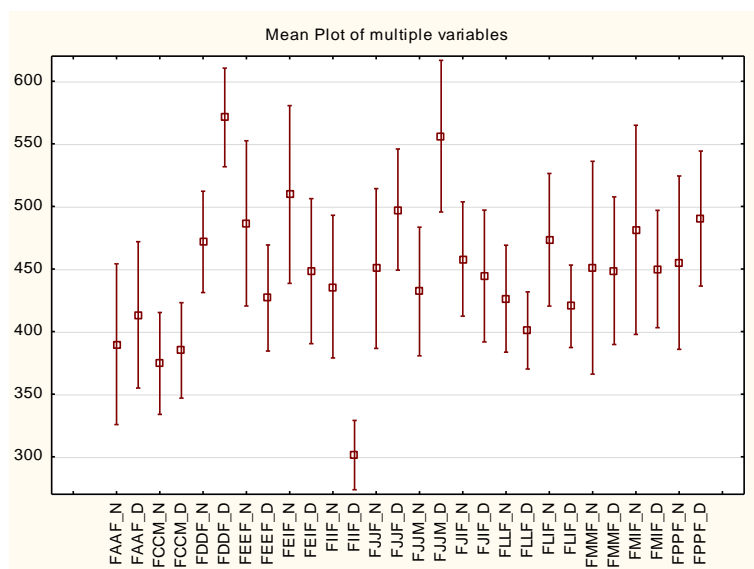
mluvčí	F1	t-test	F2	t-test	F3	t-test
FAAF_N	389,91		1577,26		2627,89	
FAAF_D	413,34	0,58	1341,39	0,09	2639,91	0,91
FCCM_N	374,51		1514,45		2594,02	
FCCM_D	384,98	0,70	1496,35	0,85	2499,20	0,31
FDDE_N	471,66		1719,06		2689,40	
FDDE_D	571,07	<0,001	1526,72	0,07	2457,57	0,02
FEFF_N	486,45		1620,97		2627,16	
FEFF_D	426,82	0,12	1703,64	0,50	2626,95	1,00
FEIF_N	509,50		1682,16		2705,92	
FEIF_D	448,28	0,17	1526,65	0,25	2446,18	0,02
FIIF_N	435,90		1733,67		2871,99	
FIIF_D	301,34	<0,0001	1789,14	0,66	2837,30	0,68
FJIF_N	455,81		1488,95		2357,83	
FJIF_D	444,37	0,73	1653,15	0,11	2618,28	0,035
FJJF_N	450,39		1669,69		2646,72	
FJJF_D	497,47	0,23	1718,04	0,71	2630,44	0,83
FJJM_N	432,01	<0,01	1526,93	0,34	2579,99	<0,01

⁴⁴ Ne všechny získané hodnoty však lze vysvětlit, výsledky by vyžadovaly manuální opravy, což by ovšem bylo časově náročné. Na tomto místě tak budou komentovány pouze ty výsledky, jež lze na první pohled interpretovat.

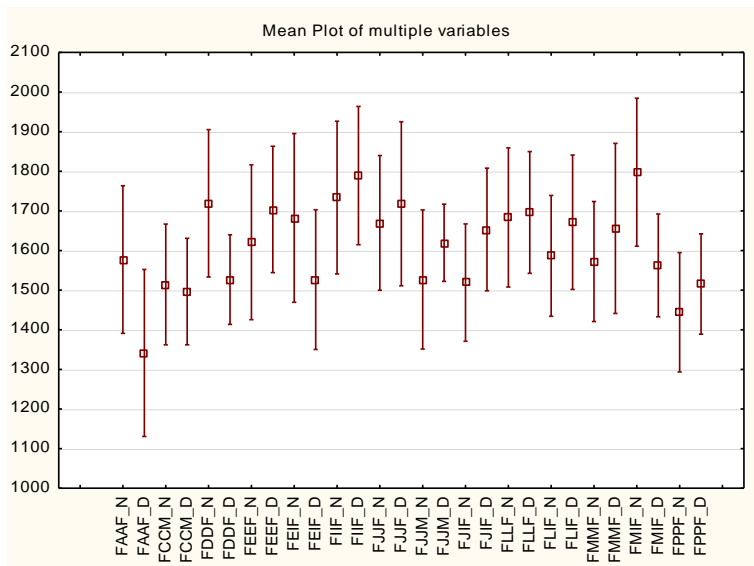
⁴⁵ Statisticky významné rozdíly jsou označeny červenou barvou.

FJJM_D	556,10		1619,58		2402,78	
FLIF_N	473,34		1586,45		2596,50	
FLIF_D	420,20	0,08	1671,36	0,44	2664,77	0,45
FLLF_N	426,29		1683,50		2780,95	
FLLF_D	400,92	0,32	1696,22	0,91	2680,39	0,28
FMIF_N	481,29		1797,69		2755,14	
FMIF_D	450,00	0,50	1562,41	0,04	2316,90	<0,0001
FMMF_N	451,00		1572,32		2628,14	
FMMF_D	448,64	0,96	1655,86	0,51	2673,32	0,54
FPPF_N	455,07		1443,99		2365,29	
FPPF_D	490,30	0,41	1515,53	0,45	2167,53	0,14

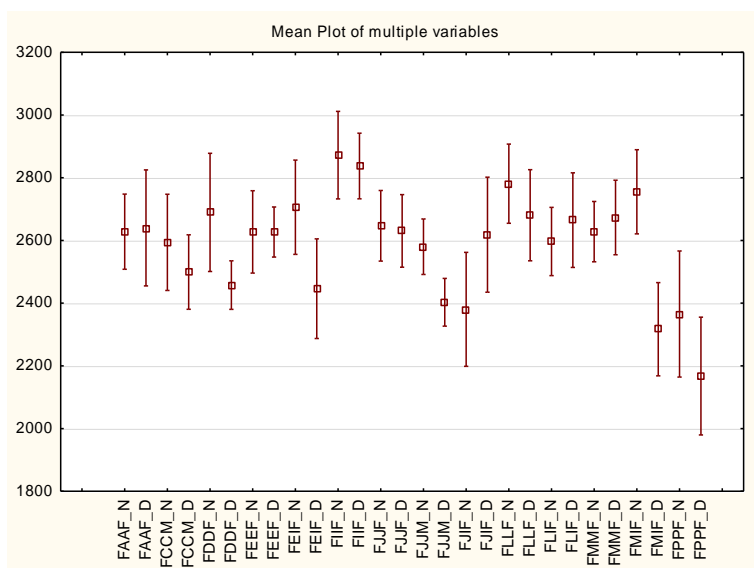
Tabulka 13 Fonetici - výsledky měření vokálních formantů



Obrázek 8 Fonetici - F1 Interval spolehlivosti a průměr



Obrázek 9 Fonetici - F2 Interval spolehlivosti a průměr



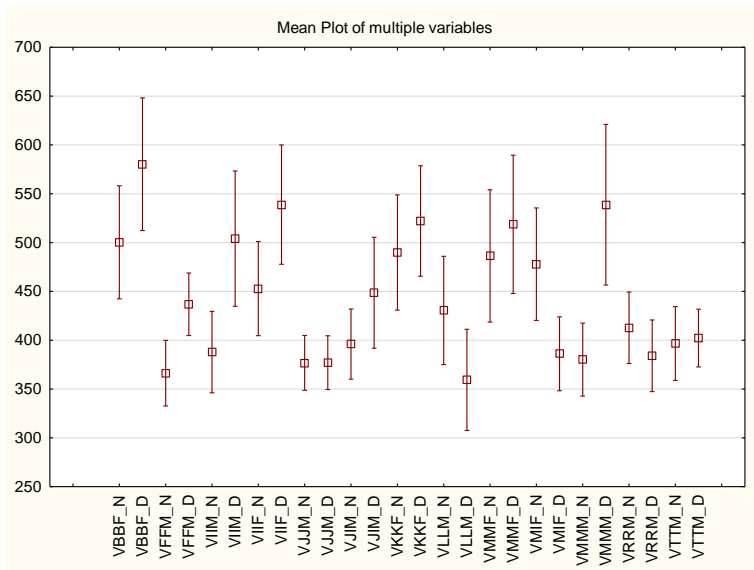
Obrázek 10 Fonetici - F3 Interval spolehlivosti a průměr

Výsledky měření hodnot formantů druhé skupiny mluvčích, studentů a absolventů nelingvistických oborů vysokých škol, jsou uvedeny v tabulce č. 14 a obrázcích č. 11, 12 a 13. Zvýšení základní frekvence hlasu pravděpodobně ovlivnilo zvýšení prvního formantu vokálů u mluvčích VFFM a VIIM. Zvýšení prvního formantu vokálů v maskovaném projevu mluvčí VIIF by mohlo být způsobeno tlačnou fonací (tj. větším napětím hlasivek), jež vede ke zkrácení vokálního traktu (srov. Thurgood, 2000: 297). Zajímavé jsou signifikantní rozdíly mezi nemaskovaným a maskovaným projevem mluvčí VMIF. Tato mluvčí svůj hlas prakticky nijak nemodifikovala, k maskování hlasu si vybrala položení předmětu (mikiny) před ústa. Jak je však možné ve výsledcích pozorovat, statisticky významné rozdíly měření potvrdilo

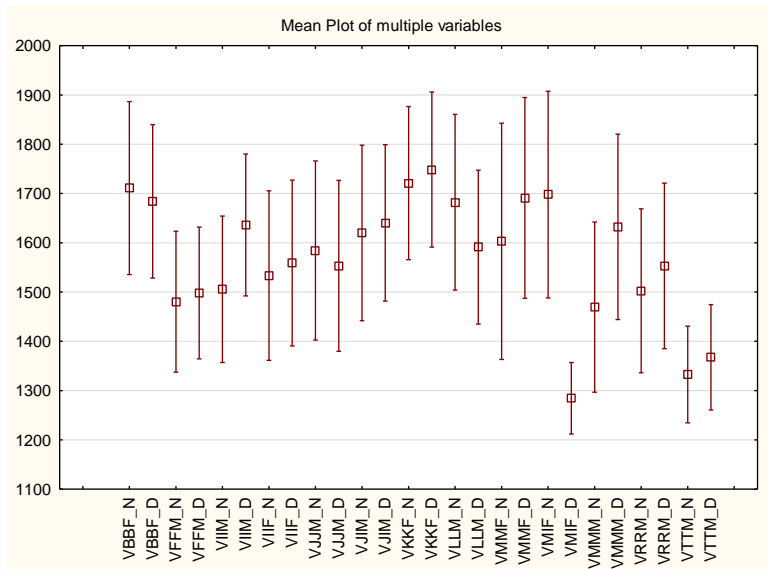
u průměrné hodnoty všech zkoumaných formantů. Tyto výsledky je možné dát do souvislosti s problémy softwarových nástrojů rozpoznat mluvčího při tomto typu maskování hlasu (které uvádí např. Perrot a kol., 2009; Perrot a kol., 2007b; Zhang & Tan, 2008).

mluvčí	F1	t-testy	F2	t-testy	F3	t-testy
VBBF_N	500,30		1711,13		2776,92	
VBBF_D	580,35	0,07	1684,21	0,81	2667,85	0,26
VFFM_N	366,27		1480,59		2611,23	
VFFM_D	436,83	<0,01	1498,11	0,85	2433,02	0,01
VIIF_N	452,92		1533,43		2433,98	
VIIF_D	538,86	0,03	1558,99	0,82	2603,98	0,06
VIIM_N	387,81		1505,73		2364,24	
VIIM_D	504,08	0,005	1642,27	0,16	2500,49	0,02
VJIM_N	396,04		1620,00		2533,25	
VJIM_D	448,60	0,11	1640,49	0,86	2704,72	<0,01
VJJM_N	376,79		1584,34		2611,77	
VJJM_D	377,02	0,99	1553,28	0,80	2579,88	0,74
VKKF_N	489,84		1721,27		2756,08	
VKKF_D	522,16	0,32	1748,72	0,96	2750,40	0,89
VLLM_N	430,51		1682,36		2776,47	
VLLM_D	359,37	0,06	1591,33	0,43	2767,47	0,92
VMIF_N	477,87		1697,88		2707,55	
VMIF_D	386,15	0,01	1284,58	<0,001	2164,71	<0,0001
VMMF_N	486,35		1603,10		2692,35	
VMMF_D	518,71	0,49	1690,98	0,56	2523,23	0,09
VMMM_N	380,19		1469,47		2501,69	
VMMM_D	538,78	<0,001	1632,46	0,19	2535,42	0,61
VRRM_N	412,79		1502,50		2565,94	
VRRM_D	384,04	0,25	1553,32	0,66	2555,28	0,85
VTTM_N	396,69		1332,85		2195,82	
VTTM_D	402,26	0,81	1367,60	0,62	2493,19	0,02

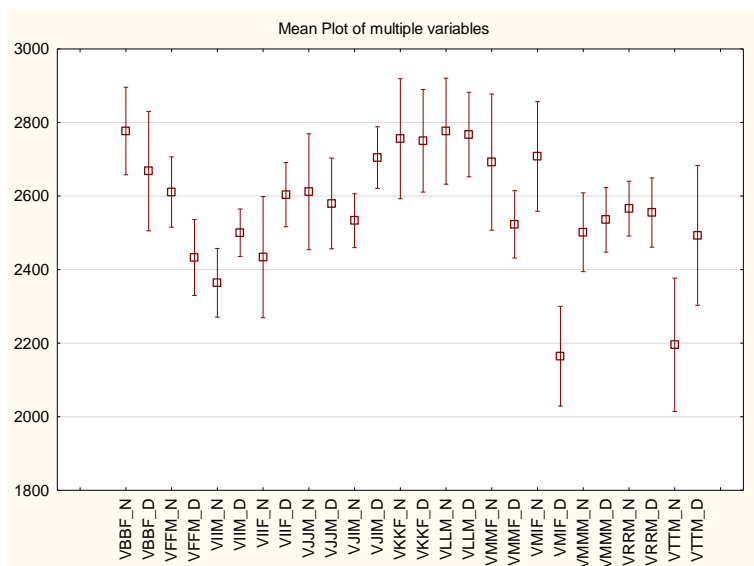
Tabulka 14 Nelingvisti - výsledky měření vokálních formantů



Obrazek 11 Nelingvisti - F1 Interval spolehlivosti a průměr



Obrazek 12 Nelingvisti - F2 Interval spolehlivosti a průměr



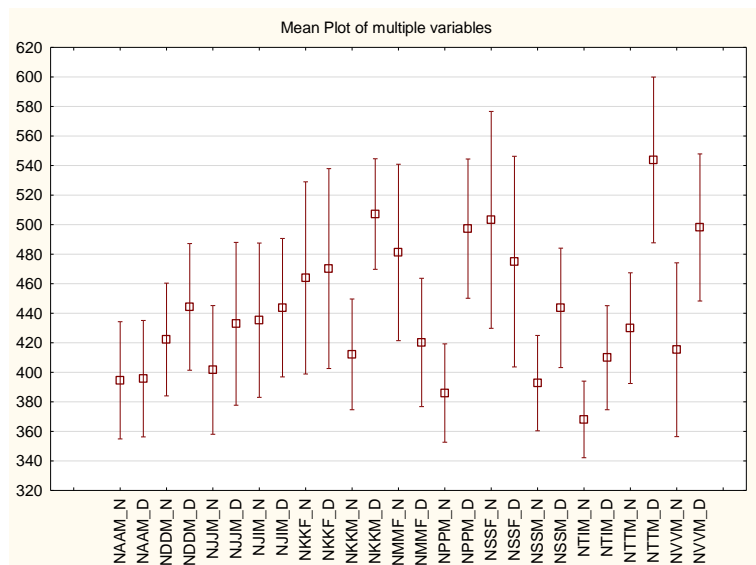
Obrázek 13 Nelingvisti - F3 Interval spolehlivosti a průměr

Tabulka č. 15 a obrázky 14, 15 a 16 ukazují výsledky naměřených hodnot vokálních formantů v projevech třetí skupiny mluvčích – mluvčích bez maturity. Signifikantní rozdíly mezi nemaskovaným a maskovaným projevem v průměrné hodnotě prvního formantu ovlivněné zejména zvýšením základní frekvence lze vidět u mluvčích NPPM a NVVM. Zvýšená hodnota prvního formantu u mluvčího NTTM pravděpodobně také souvisí se způsobem fonace, protože hlavní použitou technikou jeho maskování byla tlačaná fonace.

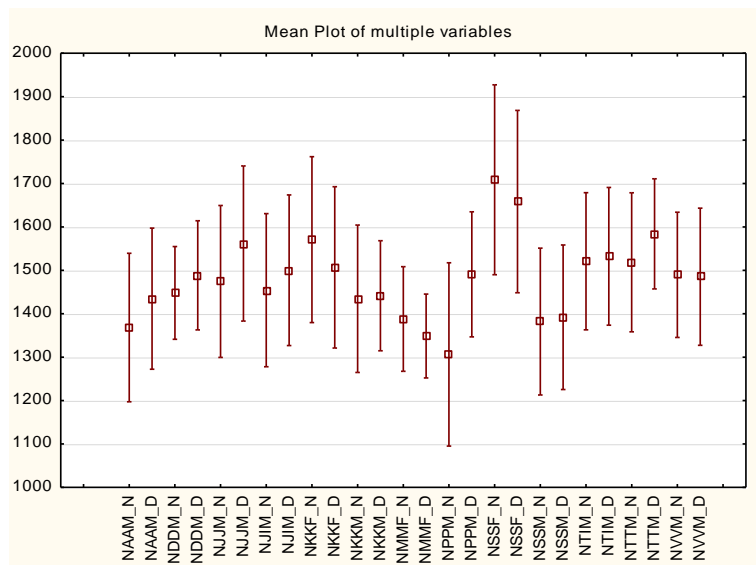
mluvčí	F1	t-testy	F2	t-testy	F3	t-testy
NAAM_N	394,58		1368,12		2515,92	
NAAM_D	397,64	0,91	1434,72	0,56	2399,55	0,14
NDDM_N	422,21		1447,93		2272,53	
NDDM_D	444,30	0,43	1488,41	0,61	2478,68	0,06
NJIM_N	435,30		1454,23		2661,67	
NJIM_D	443,78	0,80	1500,21	0,70	2608,37	0,75
NJJM_N	401,64		1474,33		2540,78	
NJJM_D	432,85	0,47	1561,72	0,48	2703,51	0,01
NKKF_N	463,93		1570,66		2526,47	
NKKF_D	470,26	0,89	1506,78	0,61	2517,14	0,89
NKKM_N	412,19		1434,49		2953,10	
NKKM_D	507,18	<0,001	1441,38	0,95	2166,96	<0,0001
NMMF_N	481,16		1387,97		2824,01	
NMMF_D	420,24	0,09	1348,64	0,60	2629,69	0,22
NPPM_N	385,97		1306,26		2531,83	
NPPM_D	497,29	<0,001	1490,71	0,14	2548,25	0,87
NSSF_N	503,24		1708,60		2525,41	
NSSF_D	474,96	0,57	1658,29	0,73	2502,77	0,78

NSSM_N	392,70		1381,99		2369,03	
NSSM_D	443,65	0,046	1391,93	0,93	2356,96	0,89
NTIM_N	368,10		1520,83		2510,25	
NTIM_D	409,88	0,05	1532,26	0,92	2466,27	0,44
NTTM_N	429,93		1518,50		2367,93	
NTTM_D	543,79	<0,01	1583,96	0,51	2259,77	0,21
NVVM_N	415,28		1489,57		2480,24	
NVVM_D	498,09	0,03	1485,27	0,97	2423,23	0,43

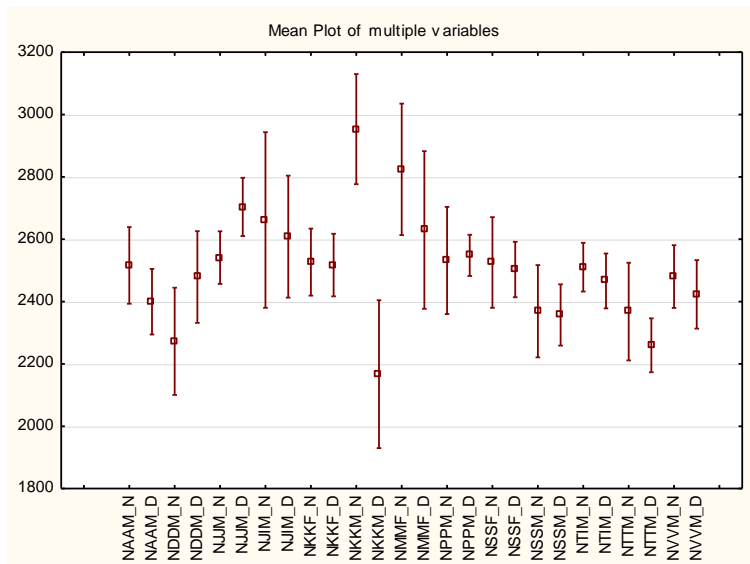
Tabulka 15 Mluví bez maturity - výsledky měření vokalicckých formantů



Obrázek 14 Mluví bez maturity - F1 Interval spolehlivosti a průměr



Obrázek 15 Mluví bez maturity - F2 Interval spolehlivosti a průměr



Obrázek 16 Mluví bez maturity - F3 Interval spolehlivosti a průměr

U několika mluvčích ukázalo měření vokalických formantů statisticky významné rozdíly mezi hodnotami získanými z jejich nemaskovaného a maskovaného projevu. Z uvedených výsledků se tak zdá, že společně s maskováním hlasu dochází i ke změně rezonančních vlastností vokálů. Naměřené hodnoty vokalických formantů však nebyly manuálně opravovány⁴⁶, a proto se ve výsledcích mohou objevovat chyby. Z důvodu chybně naměřených hodnot byli také z výsledků vyřazeni tři mluvčí (FKKM, VKIF, VPPF), u nichž vlivem maskování hlasu došlo v Praatu k chybné detekci formantů, a tudíž k velkému celkovému zkreslení všech získaných hodnot.

Z grafů na uvedených obrázcích pak je možné pozorovat nejen rozdíly mezi maskovaným a nemaskovaným projevem jednoho mluvčího, ale také mezi projevy všech mluvčích ve skupině. Tyto grafy tak ukazují, že hodnoty vokalických formantů jsou do jisté míry pro mluvčího specifické a že dlouhodobá analýza formantů má v rámci forenzní fonetiky své místo.

⁴⁶ Resp. byly z výsledků odstraněny pouze takové hodnoty, které byly na první pohled zjevné – např. příliš vysoký F1 u vysokých vokálů nebo příliš nízká hodnota F2 u /i/ (pod 1200 Hz).

6.2.3 HNR, jitter, shimmer

Tabulka č. 16 zobrazuje výsledné hodnoty jitteru, shimmeru a HNR v první skupině mluvčích, jež sestávala ze studentů a absolventů oboru Fonetika na FF UK. U každého mluvčího jsou vypsány nejprve hodnoty získané z vybraných vokálů, další část tabulky pak ukazuje výsledky měření nazálních konsonantů. Za průměrnými hodnotami z obou typů projevů se nacházejí výsledky t-testů, které porovnávaly hodnoty získané z normálního a maskovaného projevu mluvčího. Statisticky významné rozdíly jsou v nich označeny červenou barvou.

Signifikantní rozdíl u všech zkoumaných kategorií lze pozorovat pouze u mluvčí FMMF. Ta byla jednou z těch, kteří si jako hlavní strategii maskování svého hlasu zvolili chrapot. Vzhledem k tomu, že se jitter, shimmer i HNR používají právě při zkoumání patologických rysů fonace, mezi něž chrapot patří, není tento výsledek překvapivý. Jitter naměřený ve vokálech stoupl z hodnoty 0,008 na 0,017, v nazálách z 0,005 na 0,012, tedy v obou případech o dvojnásobek, podobné výsledky je možné vidět i v případě shimmeru. Podíl tónových složek u vokálů i nazálních konsonantů výrazně klesl, hodnota HNR se snížila o 10 a 15 dB. Chrapot jako jednu z hlavních technik maskování si vybrala také mluvčí FEEF. Výsledky t-testů založených na porovnání hlásek z jejího normálního a maskovaného projevu ukazují signifikantní změny v pěti zkoumaných případech, pokles harmonicity u vokálů statisticky významný nebyl. Signifikantní rozdíly mezi hodnotami v normálním a maskovaném projevu se objevily i u mluvčího FJJM, jenž maskoval svůj hlas extrémní formou třepené fonace, a to v shimmeru a HNR jak u vokálů, tak i u nazál. Hodnoty jitteru se sice zvýšily, ale pravděpodobnost náhodné chyby v t-testech byla již vyšší než 0,05 (v tomto případě se však dá předpokládat, že při vyšším počtu výskytů by byl rozdíl statisticky významný). Třepená fonace se pravděpodobně odrazila i ve výsledcích získaných z projevů mluvčí FJJF, signifikantní rozdíly potvrdily testy u jitteru a harmonicity jak u nazál, tak u vokálů. Další významné rozdíly lze pozorovat u nazál mluvčí FMIF. Vlivem denazalizace došlo v maskovaném projevu ke snížení HNR a zvýšení jitteru a shimmeru. Tento výsledek není překvapivý, podoba nazalizovaných konsonantů se totiž během denazalizace mění, více či méně přechází ve znělé explozivy, jejichž HNR je nižší. U vokálů se naopak jitter u této mluvčí významně snížil, což také souvisí s použitou technikou maskování. U mluvčího FKKM lze v rámci nazálních konsonantů pozorovat pouze snížení jitteru, další signifikantní změnou bylo zvýšení harmonicity vokálů. Tyto výsledky jsou poněkud překvapivé, statisticky významný rozdíl by vzhledem k vybrané technice maskování byl očekáván zejména ve

mluvčí	Vokály						Nazály					
	jitter	t-test	shimmer	t-test	hnr	t-test	jitter	t-test	shimmer	t-test	hnr	t-test
FAAF_N	0,011	0,11	0,088	0,07	17,69	0,22	0,008	0,73	0,050	0,93	23,31	0,89
FAAF_D	0,006		0,062		19,98		0,007		0,048		23,69	
FCCM_N	0,010	0,20	0,120	0,44	13,40	0,34	0,012	0,05	0,134	0,28	18,05	0,31
FCCM_D	0,020		0,10		11,59		0,004		0,068		20,28	
FDDE_N	0,007	0,27	0,090	0,80	14,53	0,13	0,005	0,71	0,072	0,16	18,77	0,31
FDDE_D	0,006		0,085		17,12		0,005		0,048		20,44	
FEEF_N	0,009	0,02	0,080	<0,01	13,46	0,42	0,005	0,03	0,060	<0,01	22,88	<0,01
FEEF_D	0,022		0,149		12,08		0,015		0,217		11,01	
FEIF_N	0,010	0,46	0,110	0,18	15,22	0,37	0,006	0,61	0,060	0,64	23,45	0,67
FEIF_D	0,016		0,082		16,85		0,010		0,047		25,29	
FIIF_N	0,007	0,42	0,102	0,045	16,82	<0,001	0,005	0,19	0,080	0,31	24,15	0,30
FIIF_D	0,009		0,072		21,06		0,004		0,052		26,15	
FJIF_N	0,009	0,49	0,078	0,46	16,29	0,97	0,007	0,40	0,073	0,85	20,37	0,88
FJIF_D	0,007		0,090		16,24		0,006		0,080		19,93	
FJJF_N	0,006	0,03	0,102	0,79	15,48	0,02	0,005	0,049	0,058	0,17	21,44	0,01
FJJF_D	0,014		0,106		11,51		0,015		0,125		11,44	
FJJM_N	0,017	0,55	0,114	0,02	10,48	<0,0001	0,013	0,05	0,073	0,03	15,61	0,01
FJJM_D	0,021		0,177		3,13		0,030		0,157		5,16	
FKKM_N	0,024	0,22	0,136	0,07	11,36	<0,001	0,011	0,02	0,138	0,97	14,09	0,70
FKKM_D	0,016		0,094		15,82		0,026		0,140		14,9	
FLIF_N	0,012	0,91	0,145	0,06	13,09	0,11	0,009	0,51	0,115	0,60	18,52	0,56
FLIF_D	0,012		0,108		10,41		0,013		0,090		17,11	
FLLF_N	0,008	0,43	0,102	0,97	14,99	0,08	0,005	0,23	0,068	0,39	19,43	0,19
FLLF_D	0,009		0,103		12,49		0,024		0,105		12,74	
FMIF_N	0,006	0,03	0,088	0,06	18,17	0,66	0,004	0,01	0,050	0,049	23,08	0,01
FMIF_D	0,004		0,066		18,78		0,013		0,250		12,98	
FMMF_N	0,008	0,01	0,079	<0,001	16,17	<0,0001	0,005	0,03	0,067	<0,001	22,83	<0,0001
FMMF_D	0,017		0,164		6,39		0,012		0,280		7,4	
FPPF_N	0,005	0,81	0,072	0,19	16,27	0,96	0,007	0,64	0,060	0,55	22,59	0,56
FPPF_D	0,005		0,095		16,17		0,006		0,043		24,49	

Tabulka 16 Fonetici - hodnoty HNR, jitteru a shimmeru získané z vybraných vokálů a nazál

snížení hodnoty HNR u nazál (na základě měření však HNR vychází ještě o něco vyšší než v projevu nemaskovaném). Naopak HNR vokálů by měla být spíše nižší, protože mluvčí vokály výrazně krátí. Po prozkoumání signálu bylo zjištěno, že jsou v něm vysoké frekvence (tj. šumové) slabší, proto byla hodnota HNR naměřena jako vyšší, přestože sluchově je dojem opačný. Jistě by bylo zajímavé zjistit, do jaké míry souvisí harmonicita se spektrálním sklonem. Podobný výsledek HNR u vokálů je pak možné vidět i u mluvčí FIIF.

Tabulka č. 17 podává přehled hodnot jitteru, shimmeru a harmonicity u druhé skupiny mluvčích, studentů a absolventů nelingvistických oborů vysokých škol. Z mluvčích, kteří si jako jednu z hlavních technik k maskování svého hlasu zvolili chrapot (VJIM, VIIM, VMMM), se statisticky významné rozdíly v některých hodnotách ze čteného a maskovaného projevu objevily pouze u mluvčího VMMM. U vokálů vyšel signifikantní nárůst shimmeru a snížená HNR, u nazálních konsonantů pouze HNR (přičemž pravděpodobnost náhodné chyby u jitteru byla pouze 0,07 - i zde se dá předpokládat, že při větším množství výskytů by i tento rozdíl mohl být významný). Statisticky významný rozdíl v shimmeru a HNR ve vokálech i nazálech lze spatřit u mluvčí VKIF, jejíž maskovaný hlas byl dyšný, téměř šeptaný. Z tabulky je možné vidět, že se výsledný poměr harmonických a šumových složek u vybraných vokálů snížil z 18 na 6 dB, což značí rapidní pokles harmonických složek. U některých vokálů byly dokonce naměřeny hodnoty záporné – podíl šumových složek v nich byl vyšší než tónových. Významné rozdíly v naměřených hodnotách shimmeru a HNR u vokálů je možné vidět u mluvčího VJIM. Na rozdíl od předchozí mluvčí u něj byly nalezeny tendence opačné – snížení hodnot jitteru a zvýšení harmonicity, což může být nepřímo ovlivněno pravidelným rytmem jeho maskované řeči. Signifikantní rozdíly v HNR se u vokálů i nazál projeví u mluvčí VMIF, u nazál i v shimmeru. Její maskování spočívalo v položení mikiny před ústa, poslechově byl rozdíl mezi normálním a maskovaným projevem minimální. U této mluvčí se statisticky významné rozdíly potvrdily i v hodnotách formantů u vokálů (viz předchozí kap.). Tyto výsledky ukazují, že akustická analýza bez analýzy poslechové může být často nedostatečná. To dokazují i deklarované problémy při identifikaci mluvčího automatickými systémy právě při tomto typu maskování hlasu (např. Perrot a kol., 2009; Perrot a kol., 2007b; Zhang & Tan, 2008). Poslední mluvčí, u níž se objevují statisticky významné rozdíly v naměřených hodnotách, je VPPF, která svůj projev denazalizovala. Maskování ovlivnilo harmonicitu u vokálů i nazál, zejména u nazál její hodnota zdatelně poklesla. Jak již bylo řečeno výše, zejména u denazalizovaných konsonantů se hodnoty HNR výrazně mění.

mluvčí	Vokály						Nazály					
	jitter	t-testy	shimmer	t-testy	hnr	t-testy	jitter	t-test	shimmer	t-test	hnr	t-test
VBBF_N	0,009	0,42	0,096	0,77	16,14	0,19	0,005	0,79	0,117	0,09	21,16	0,91
VBBF_D	0,006		0,092		14,54		0,006		0,047		21,52	
VFFM_N	0,013	0,28	0,136	0,52	13,09	0,25	0,006	0,19	0,065	0,41	21,00	0,48
VFFM_D	0,010		0,118		15,15		0,013		0,098		18,26	
VHIF_N	0,009	0,24	0,114	0,94	12,97	0,93	0,008	0,70	0,095	0,40	18,85	0,46
VHIF_D	0,015		0,112		12,83		0,009		0,073		16,66	
VHIM_N	0,011	0,90	0,143	0,49	11,78	0,51	0,006	0,86	0,107	0,98	17,12	0,83
VHIM_D	0,011		0,127		10,54		0,006		0,108		16,15	
VJIM_N	0,015	0,49	0,121	0,51	10,15	0,52	0,008	0,26	0,053	0,33	17,71	0,35
VJIM_D	0,019		0,105		8,96		0,012		0,070		14,78	
VJJM_N	0,015	0,17	0,126	0,01	12,85	<0,01	0,012	0,18	0,085	0,85	17,44	0,10
VJJM_D	0,008		0,077		17,37		0,008		0,078		22,14	
VKIF_N	0,006	0,06	0,063	<0,001	18,44	<0,0001	0,004	0,15	0,042	<0,01	23,58	<0,01
VKIF_D	0,012		0,124		6,62		0,008		0,120		14,11	
VKKF_N	0,013	0,11	0,130	0,14	12,12	0,32	0,008	0,67	0,072	0,70	18,48	0,41
VKKF_D	0,008		0,106		13,42		0,007		0,063		21,35	
VLLM_N	0,013	0,6	0,128	0,16	11,11	0,61	0,007	0,37	0,095	0,06	16,43	0,70
VLLM_D	0,011		0,084		10,18		0,009		0,028		17,35	
VMIF_N	0,006	0,64	0,119	0,12	15,47	<0,01	0,007	0,43	0,105	0,01	17,18	0,04
VMIF_D	0,006		0,082		20,85		0,006		0,057		22,79	
VMMF_N	0,012	0,13	0,097	0,10	13,98	0,17	0,005	0,08	0,052	0,56	21,54	0,78
VMMF_D	0,008		0,075		16,18		0,003		0,040		22,21	
VMMM_N	0,014	1,00	0,091	0,02	12,86	<0,001	0,012	0,39	0,052	0,07	16,61	0,02
VMMM_D	0,014		0,157		6,90		0,017		0,133		9,42	
VPPF_N	0,010	0,28	0,101	0,41	13,88	0,04	0,005	0,02	0,045	0,05	23,39	<0,01
VPPF_D	0,020		0,089		10,31		0,010		0,090		9,05	
VRRM_N	0,021	0,62	0,079	0,11	9,94	0,14	0,017	0,58	0,044	0,10	16,77	0,57
VRRM_D	0,016		0,104		12,35		0,011		0,135		18,99	
VTTM_N	0,015	0,25	0,117	0,73	12,45	0,36	0,011	0,29	0,098	0,54	16,35	0,65
VTTM_D	0,011		0,108		10,63		0,007		0,074		14,68	

Tabulka 17 Nelingvisti - hodnoty HNR, jitteru a shimmeru získané z vybraných vokálů a nazál

Tabulka č. 18 předkládá výsledky třetí skupiny, mluvčích bez maturity. Nejvíce statisticky významných rozdílů se objevilo u mluvčího NTIM, který jako hlavní strategii k zakrytí své identity použil chrapot, místy až šeptaný. Není tedy překvapivé, že byl pokles hodnot HNR v jeho maskovaném projevu u vokálů i nazál signifikantní. Jako významný se ukázal také rozdíl v shimmeru u vokálů. Chrapot jako hlavní techniku svého maskování si zvolil i mluvčí FJJM, u něj se však žádné statisticky významné rozdíly nepotvrdily. U vokálů se signifikantní rozdíl v shimmeru a HNR objevil u mluvčího NDDM. Tónové složky ve vokálech se zvýšily, hodnota shimmeru klesla. Tyto hodnoty byly pravděpodobně ovlivněny jednou z jeho hlavních taktik maskování hlasu – dloužením vokálů. Významné snížení harmonicity u nazálních konsonantů se potvrdilo u mluvčího NKKM, který svůj maskovaný projev denazalizoval – jak již bylo naznačeno výše, denazalizace HNR nazál snižuje. Další statisticky významné (a většinou i poslechem potvrditelné) rozdíly související s výběrem maskování hlasu je možné pozorovat i u mluvčích NAAM (zvýšení shimmeru u vokálů), NSSM (snížení jitteru u vokálů), NTTM (snížení hodnoty HNR u vokálů) a NVVM (zvýšení shimmeru u vokálů a snížení jitteru u nazál).

mluvčí	Vokály						Nazály					
	jitter	t-testy	shimmer	t-testy	hnr	t-testy	jitter	t-test	shimmer	t-test	hnr	t-test
NAAM_N	0,014	0,64	0,067	0,02	12,01	0,35	0,014	0,68	0,093	0,72	14,16	0,99
NAAM_D	0,013		0,117		13,52		0,017		0,080		14,20	
NDDM_N	0,015	0,05	0,137	0,01	12,02	0,01	0,006	0,31	0,037	0,17	17,20	0,87
NDDM_D	0,008		0,093		16,22		0,004		0,082		17,71	
NJIM_N	0,013	0,49	0,095	0,11	11,70	0,37	0,012	0,83	0,098	0,6	14,96	0,92
NJIM_D	0,015		0,136		10,16		0,013		0,160		14,58	
NJJM_N	0,013	0,64	0,085	0,20	14,58	0,74	0,010	0,89	0,076	0,67	16,54	0,42
NJJM_D	0,015		0,104		15,09		0,011		0,087		18,69	
NKKF_N	0,011	0,95	0,077	0,87	15,78	0,44	0,007	0,32	0,055	0,95	24,63	0,42
NKKF_D	0,011		0,079		14,48		0,011		0,057		21,55	
NKKM_N	0,012	0,09	0,143	0,31	9,83	0,48	0,010	0,30	0,045	0,22	16,23	<0,01
NKKM_D	0,031		0,108		8,48		0,025		0,140		6,10	
NMMF_N	0,008	0,75	0,082	0,70	20,46	0,37	0,004	0,38	0,053	0,16	24,11	0,56
NMMF_D	0,007		0,089		19,15		0,006		0,080		22,71	
NPPM_N	0,017	0,05	0,104	0,94	12,67	0,37	0,015	0,76	0,078	0,04	17,03	0,32
NPPM_D	0,008		0,100		14,17		0,013		0,223		12,72	
NSSF_N	0,007	0,06	0,089	0,63	13,92	0,24	0,006	0,19	0,065	0,18	20,88	0,03
NSSF_D	0,019		0,102		11,49		0,022		0,116		10,30	
NSSM_N	0,020	0,03	0,131	0,86	13,13	0,79	0,011	0,67	0,082	0,82	16,35	0,82
NSSM_D	0,012		0,127		13,57		0,009		0,088		17,00	
NTIM_N	0,019	0,9	0,108	<0,0001	13,80	<0,0001	0,008	0,21	0,093	0,16	15,53	0,03
NTIM_D	0,019		0,251		4,80		0,020		0,210		7,42	
NTTM_N	0,012	0,26	0,122	0,21	14,88	0,01	0,014	0,64	0,062	0,35	18,28	0,14
NTTM_D	0,009		0,090		10,00		0,011		0,112		12,72	
NVVM_N	0,019	0,48	0,073	0,04	12,65	0,66	0,023	0,03	0,071	0,62	14,02	0,97
NVVM_D	0,015		0,132		11,99		0,003		0,040		13,89	

Tabulka 18 Mluví bez maturity - hodnoty HNR, jitteru a shimmeru získané z vybraných vokálů a nazál

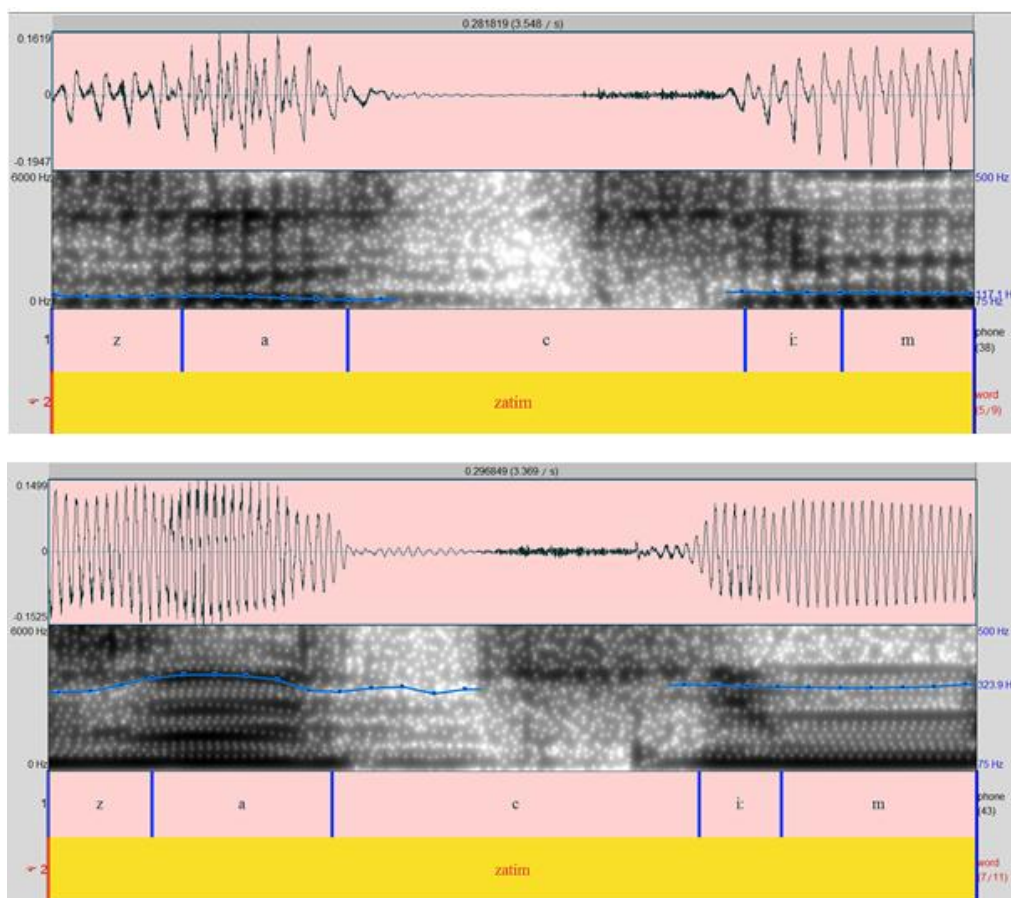
Výsledky všech tří zkoumaných skupin i přes poměrně nízký počet zkoumaných hlásek u některých mluvčích ukazují statisticky významné rozdíly mezi hodnotami jitteru, shimmeru a HNR mezi oběma typy projevů. Tyto signifikantní změny se týkají jak vokálů, tak i nazálních konsonantů a zpravidla se nacházejí v souvislosti s použitou technikou maskování hlasu. Asi nejzajímavější výsledek pak představují naměřené hodnoty u mluvčí VMIF, jejíž technika, položení mikiny před ústa, významně ovlivnila některé ze zkoumaných hodnot.

6.3 Ilustrace konkrétních případů maskování hlasu

Cílem této kapitoly je poukázat na některé z provedených typů maskování hlasu, jejichž realizace jsou dobře zobrazitelné v oscilogramech a spektrogramech získaných pomocí programu Praat (Boersma & Weenink, 2012).

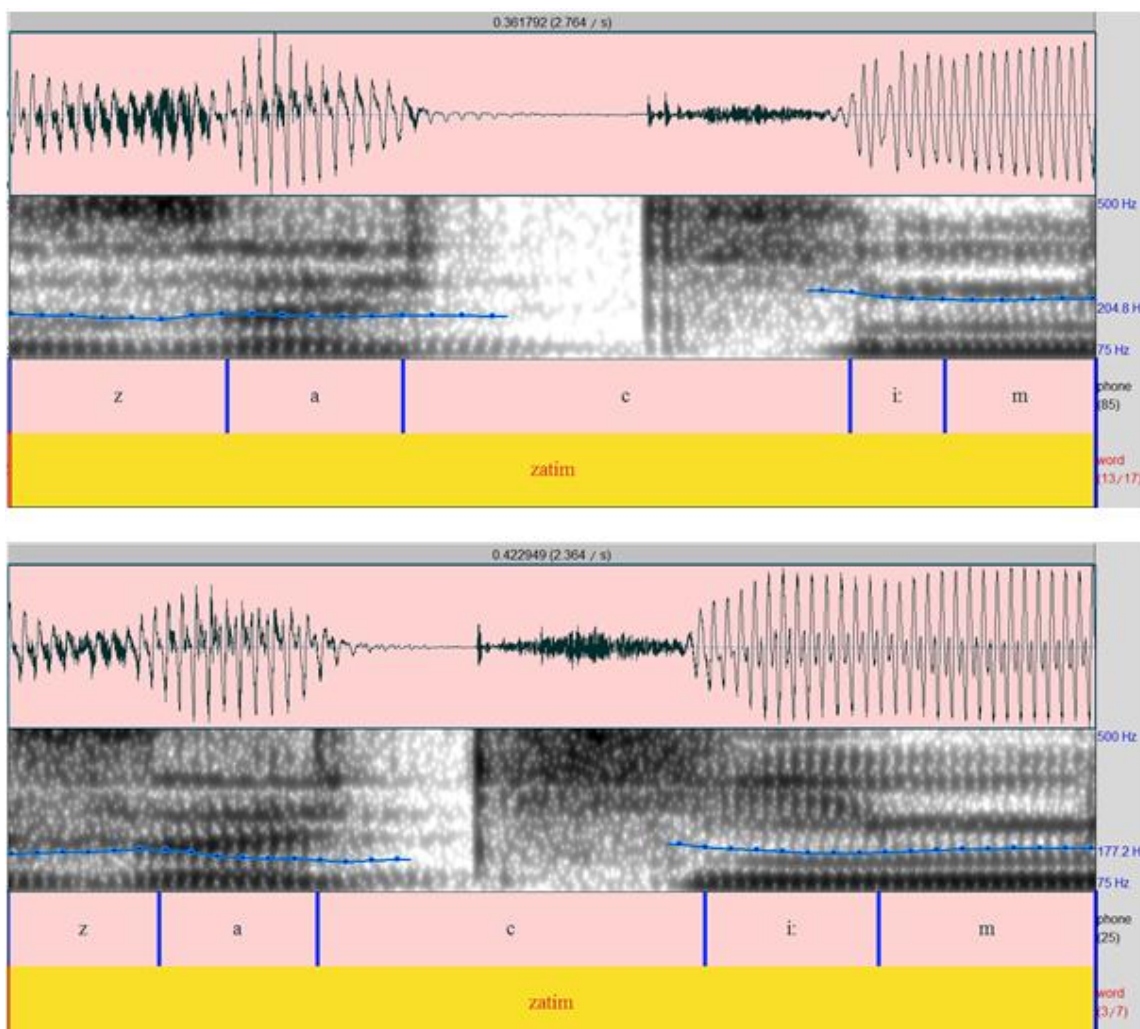
6.3.1 Fonační změny

Fonační změny byly nejčastěji prováděným typem maskování hlasu. Obrázek č. 17 zobrazuje slovo „zatím“ vyslovené mluvčím VFFM, jehož hlavní strategií maskování hlasu představovalo zvýšení základní frekvence. V horní části je možné vidět realizaci tohoto slova v běžném čteném projevu, v dolní části jeho maskovanou verzi, průběh F0 je znázorněn modrou čarou umístěnou ve spektrogramech. Při porovnání obou verzí lze na první pohled pozorovat výrazný rozdíl, křivka F0 je v druhém případě umístěna znatelně výše. Tento rozdíl je na první pohled viditelný i z tvaru zvukové vlny – periody znělých hlásek jsou v horní části obrázku od sebe vzdálenější než v jeho dolní části (což se odráží i ve výsledné podobě spektrogramu).



Obrázek 17 VFFM Průběh F0 ve slově „zatím“ ve čteném a maskovaném projevu

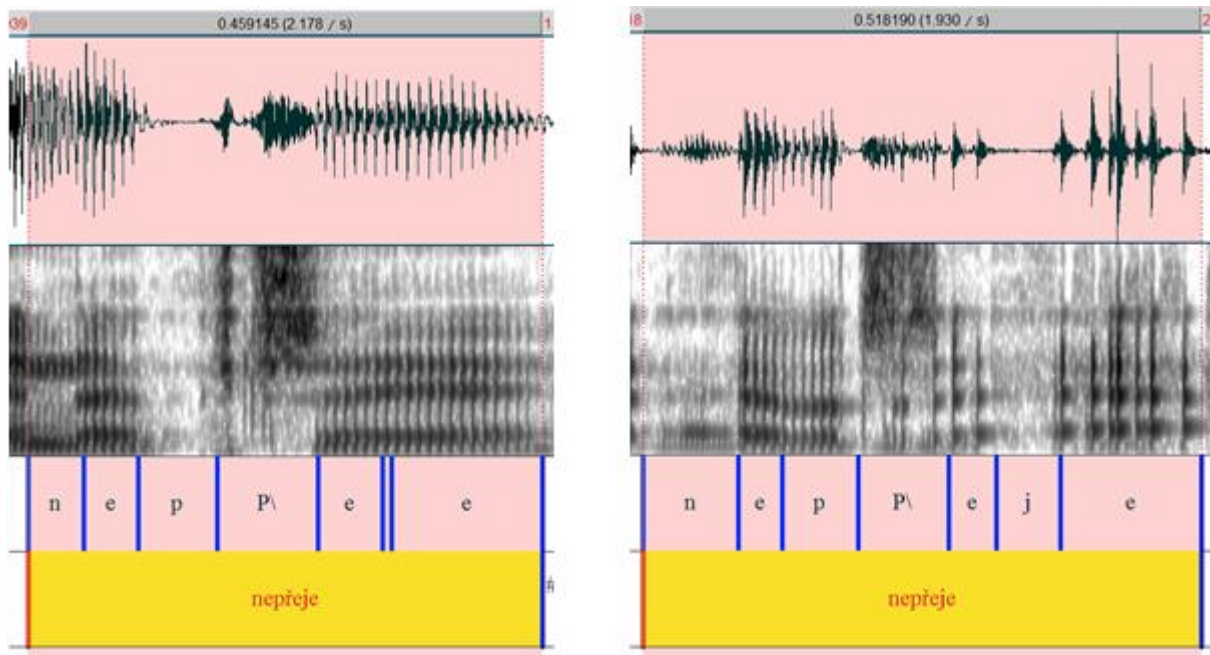
Obr. č. 18⁴⁷ odkazuje k opačné taktice změně F0, a to k jejímu snížení. Mluvčí FLLF však svoji základní frekvenci již nesnížila v takové míře, jako ji předchozí mluvčí zvýšil, proto rozdíl mezi oběma částmi obrázku již není tak nápadný.



Obrázek 18 FLLF Průběh F0 ve slově „zatím“ ve čteném a maskovaném projevu

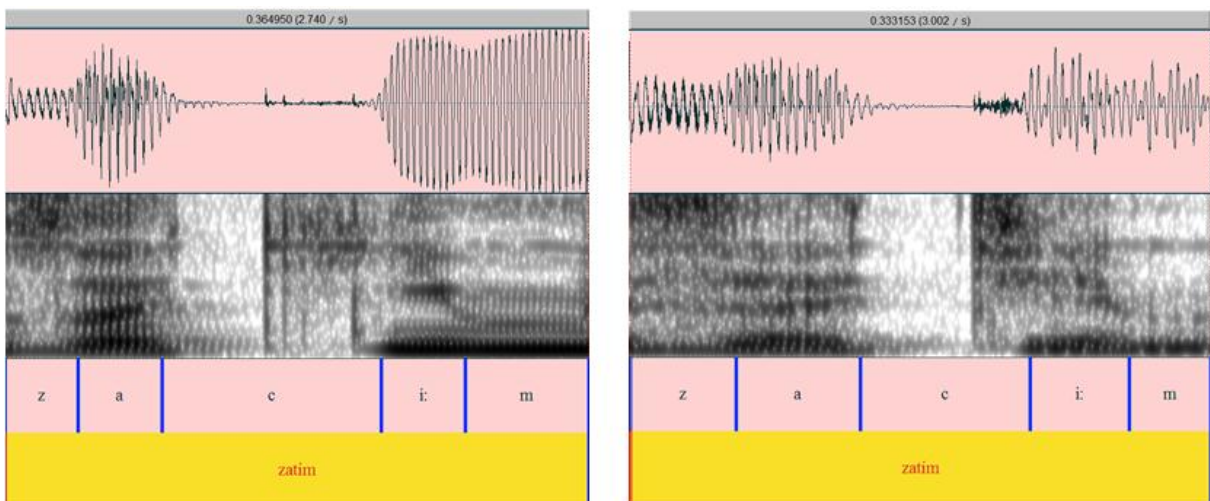
Se snížením základní frekvence souvisí i obr. č. 19, na němž se nachází srovnání realizací slova „nepřeje“ mluvčím FJJM. Po levé straně lze vidět jeho zobrazení v běžném čteném projevu, po pravé v maskovaném. V maskovaném projevu je možné ve znělých hláskách pozorovat charakteristické jevy spojené s třepenou fonací - méně pravidelnou periodu s většími rozestupy, místy jakoby přerušovanou zvukovou vlnu (častý pokles na hodnoty těsně kolem nuly) či výrazně nižší amplitudu kmitání.

⁴⁷ Všechny spektrogramy byly nastaveny na rozmezí 0-6 kHz, výjimku tvoří pouze spektrogram na obrázku č. 19, v němž bylo rozmezí nastaveno na 0-5 kHz.



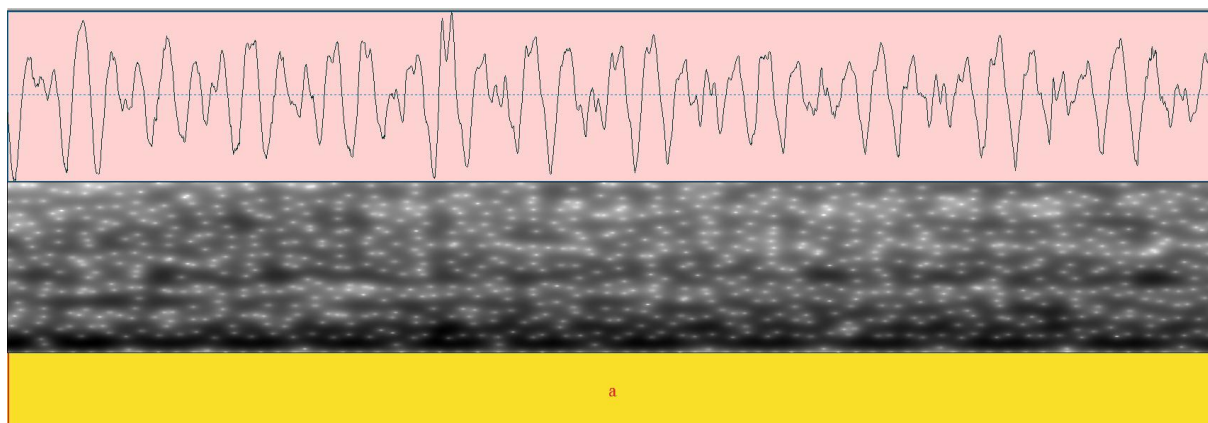
Obrázek 19 FJMM Normální realizace a třepená fonace u slova „nepřeje“

Jednou z častěji prováděných fonačních změn byl i chrapot. Na obrázku č. 20 lze vidět realizaci slova „zatím“ v projevu mluvčí FEEF. To, že se jedná o chrapot, lze vyčíst jednak z většího množství šumu ve spektrogramu (méně výrazné formanty, jakoby vyčárkovaný spektrogram), jednak z narušení periodicity kmitání hlasivek (to je nejlépe vidět na konci slova).



Obrázek 20 FEEF – slovo „zatím“ ve čteném a maskovaném projevu

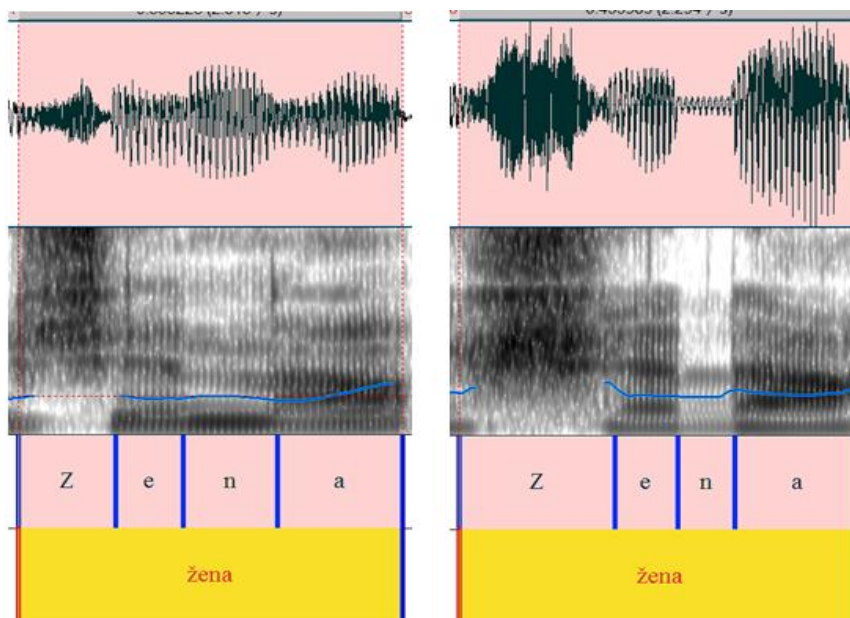
Obrázek č. 21 zobrazuje vokál /a/ ve slově „žena“ v maskovaném projevu stejné mluvčí. Ze zvukové vlny je možné vidět, že zde dochází k tzv. diplofonii. Diplofonie je stav, kdy každá hlasivka kmitá jinak, čímž dochází k aperiodicitě řečového signálu.



Obrázek 21 F0-F3 - /a/ ve slově "žena"

6.3.2 Denazalizace

Další poměrně oblíbenou a ve spektrogramu dobře zobrazitelnou strategií maskování hlasu je denazalizace. Ta se dobře dokládá u nazál, v jejichž standardní výslovnosti vzduch částečně proudí skrze nosní dutinu. Rozvětvením cesty výdechového proudu tak dochází k vytváření antiformantů, které lze obvykle bez obtíží nalézt ve spektrogramu (bílá místa). Při ucpání nosních dírek se cesta skrze nosní dutinu uzavře, a tak se nazály zobrazují podobně jako explozívy, jak ilustruje obrázek č. 22 na slově „žena“ v projevu mluvčí VPPF⁴⁸. Denazalizace se ve spektrogramu projevuje i v zobrazení formantů, při srovnání obou realizací je patrná jejich větší vyrýsovanost v projevu maskovaném.

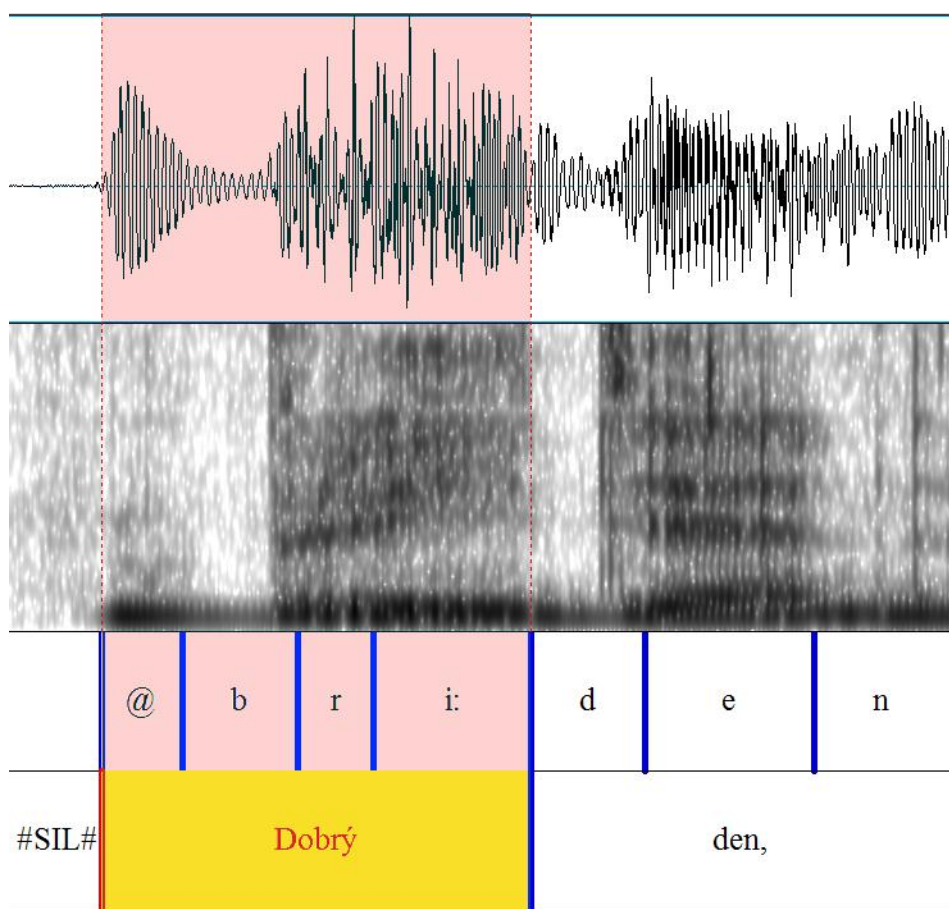


Obrázek 22 VPPF slovo "žena" ve čteném a maskovaném projevu

⁴⁸ U této mluvčí došlo k desonorizaci /ž/ ve slově žena v obou projevech, u dalších dvou se však desonorizace objevila jen u maskovaného projevu.

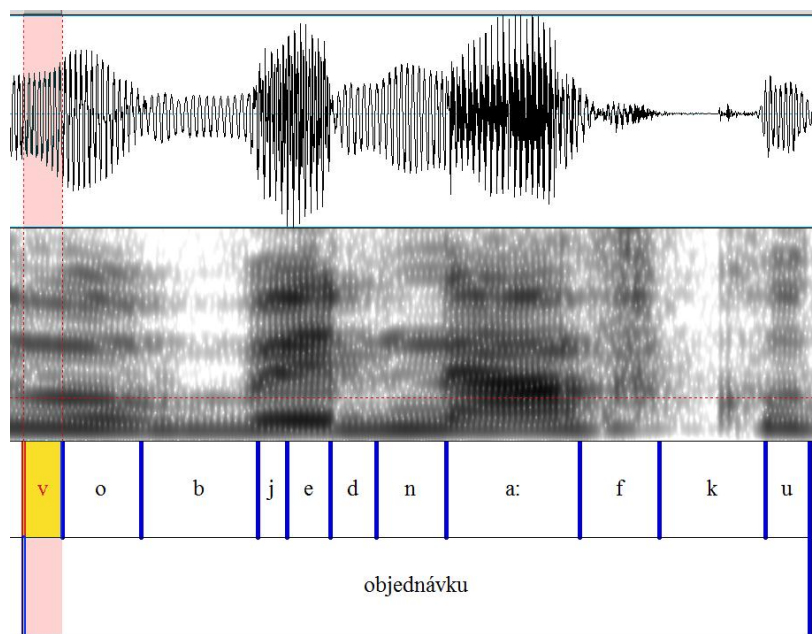
6.3.3 Změny v artikulaci hlásek

Někteří mluvčí se v rámci strategií o maskování změnu uchýlovali i ke změně artikulace některých hlásek, ať už konsonantů či vokálů. Mnoho z těchto změn je patrných i ve spektrogramech. Ledabylou výslovnost lze pozorovat zejména tam, kde dochází k odstranění některých hlásek, jak ilustruje obrázek č. 23. V něm je možné vidět, že mluvčí FEEF nahradila slabiku „do“ ve slově „dobrý“ pouze krátkým šva tak, jak je obvyklé v běžné ústní komunikaci.



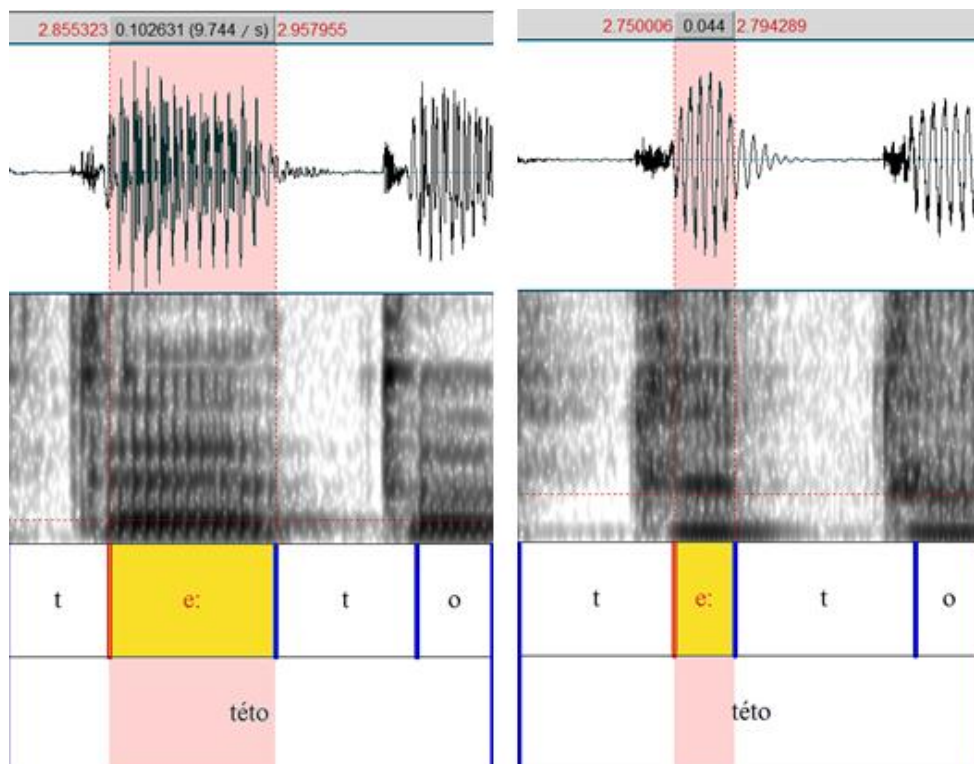
Obrázek 23 FEEF - "dobrý den" v maskovaném projevu

Obrázek 24 zobrazuje změnu opačnou, a to přidání protetického v- před slovo začínající na o-. Protetické v- je charakteristickým rysem obecné češtiny, která je rozšířená zejména v oblasti Čech a západní Moravy.

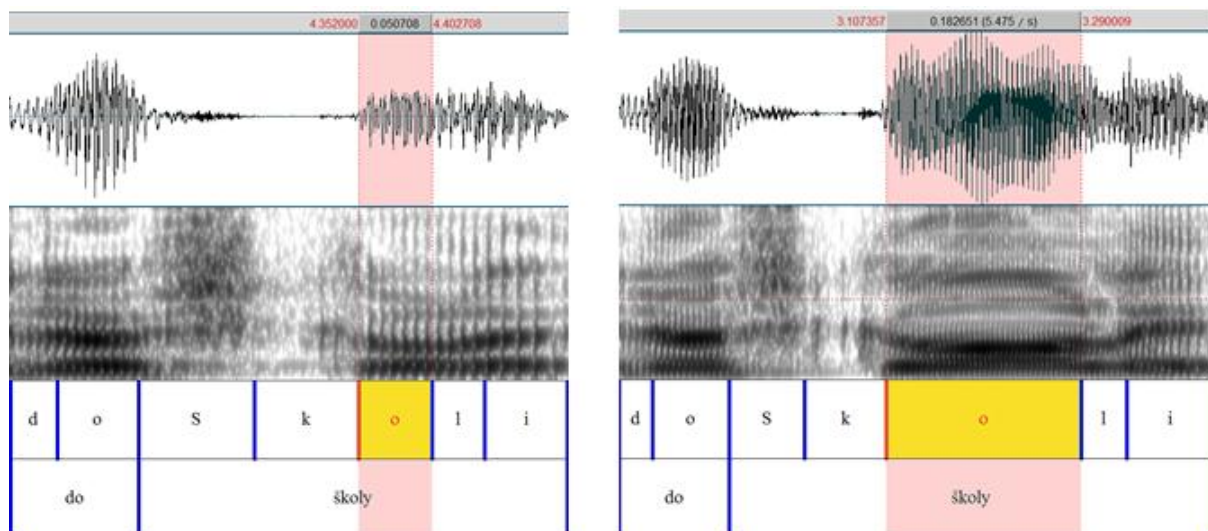


Obrázek 24 VKKF - protetické v- ve slově "objednátku"

Někteří mluvčí záměrně měnili i temporální charakteristiku vokálů, zkrácení vokálu v maskovaném projevu je zobrazeno na obrázku č. 25, obrázek č. 26 naopak překládá jeho prodloužení (po levé straně je vždy spektrogram nemaskovaného projevu, napravo maskovaného).

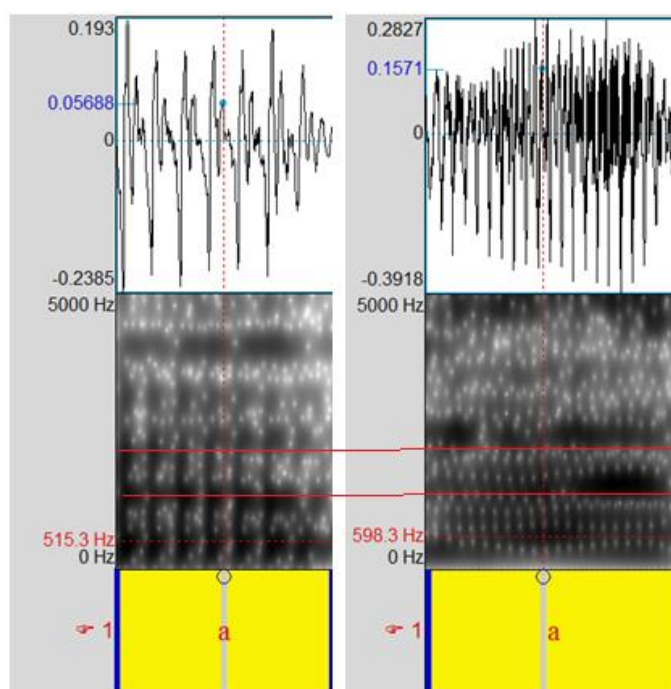


Obrázek 25 FKMM - krácení vokálu /e:/ ve slově "tэто"



Obrázek 26 NDDM - dlužení vokálu /o/ ve slově "školy"

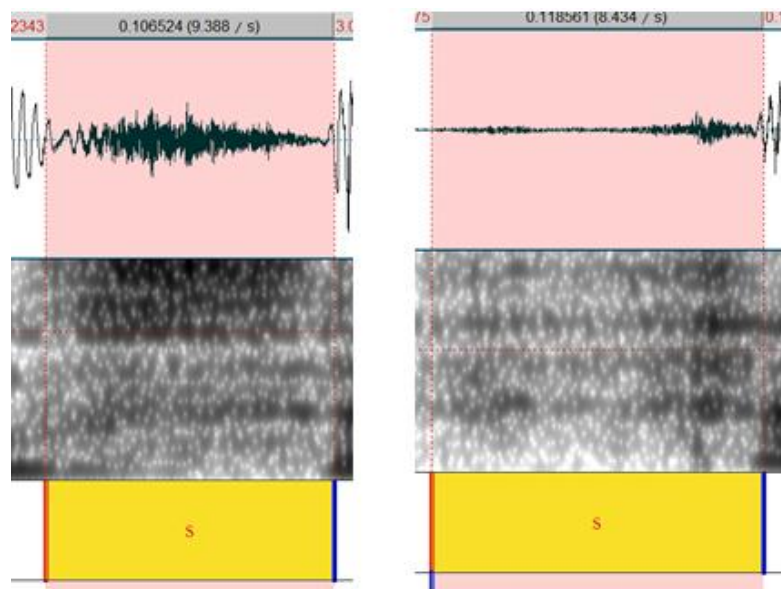
Vlivem použité strategie maskování někdy dochází k posunu formantu vokálů směrem nahoru či dolů. Příklad posunu formantů směrem nahoru ve vokálu „a“ získaném ze slova „setkat“ je zobrazen na obrázku č. 27.



Obrázek 27 VIIM - posun formantů ve vokálu /a/ ve slově setkat

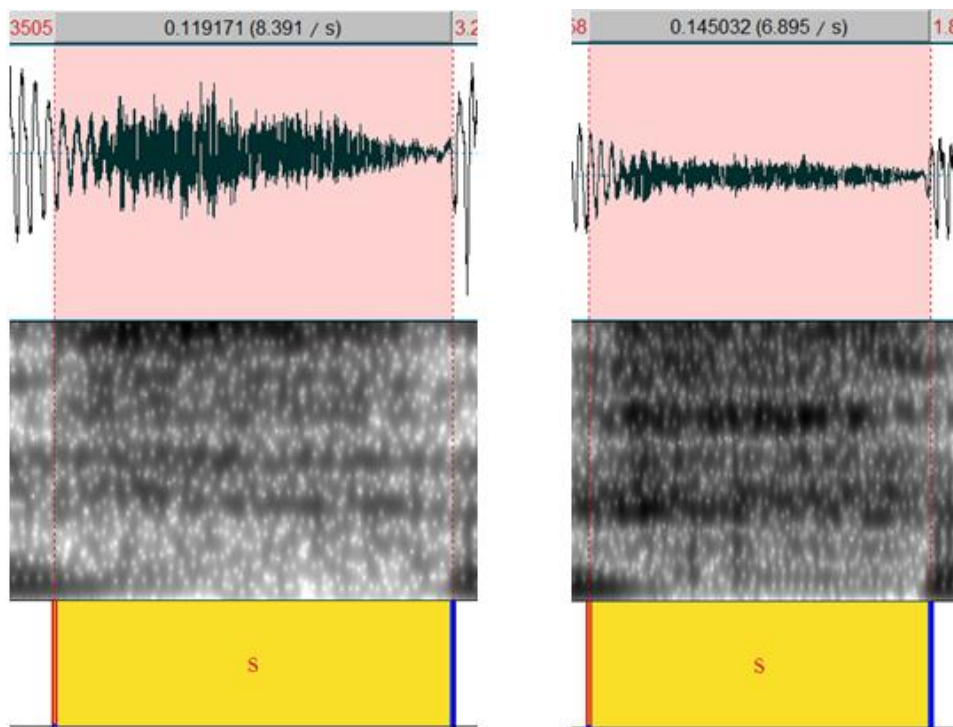
Dva participanti si jako jednu ze strategií maskování svého hlasu zvolili napodobování řečových vad. Mluvčí FMIF se rozhodla pro sigmatismus, realizaci její normální a maskované hlásky /s/ vyslovené ve slově setkat je možné spatřit na obr. č. 28. V normální realizaci je

možné vidět výrazný šumový formant ve vyšších frekvencích, zatímco dentální /s/ v maskovaném projevu jej má rozprostřený v mnohem širším frekvenčním pásmu.



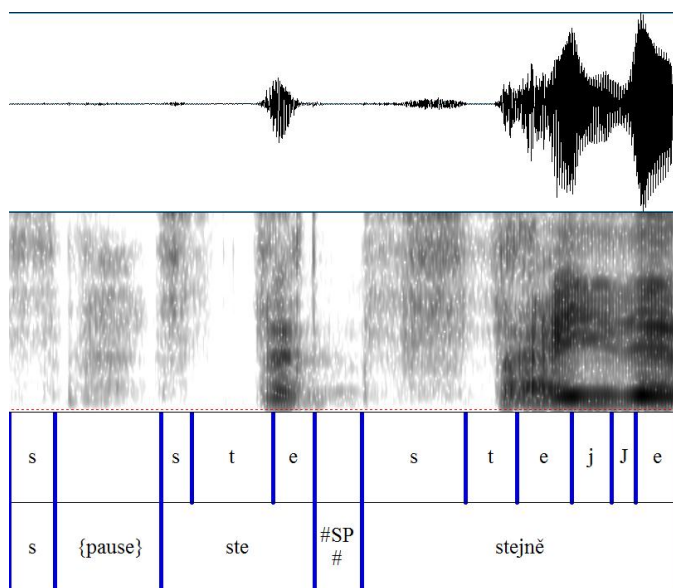
Obrázek 28 FMIF - /s/ ve slově setkat ve čteném a maskovaném projevu

Podobně jako napodobování řečových vad se projevuje i umístění cizího předmětu do úst (samozřejmě záleží i na rozměru a přesném umístění překážky). Změnu výslovnosti hlásky /s/ (posun jazyka směrem dozadu) ve slově „setkat“ následkem překážky nacházející se uvnitř úst demonstruje obr. č. 29.



Obrázek 29 FEIF - /s/ ve slově setkat v maskovaném a čteném projevu

Maskování svého hlasu o zajímavou strategii doplnil mluvčí NPPM, když se záměrně zakoktal na slově „stejně“⁴⁹ na začátku věty. Realizaci tohoto slova je možné nalézt na obr. č. 30.



Obrázek 30 NPPM - zadrnutí se na slově "stejně"

⁴⁹ Celá věta zněla: Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

6.4 Poslechový test

Součástí této práce bylo i provedení poslechového testu, jehož hlavním cílem bylo zjistit, zda a v jakém měřítku jsou posluchači na základě jedné věty schopni identifikovat neznámého mluvčího maskujícího svůj hlas. Test měl také poukázat na úspěšnost jednotlivých strategií maskování hlasu a případné další aspekty ovlivňující rozhodování posluchače, jako například délku pronesené věty či umístění původního (správného) mluvčího v položce.

Výsledky poslechového testu zobrazuje tabulka č. 19. První sloupec odkazuje na pořadí položky v testu, druhý sloupec uvádí mluvčího, jehož rozpoznání bylo zkoumáno. Zbylé tři sloupce ukazují, kolik posluchačů označilo kterého z mluvčích⁵⁰, správné identifikace se vždy nachází v šedivě označeném poli. Závěčné položky se do výsledků nezapočítávaly, úmyslně byly jednodušší. Procento správné identifikace u nich proto bylo poměrně vysoké, u první 100 %⁵¹, u druhé 69 %. Nejsnáze rozpoznatelnou mluvčí v testu se podle očekávání stala mluvčí NKKF (92,9 %), jejíž projev byl do testu začleněn jako kontrolní položka. Dále měli posluchači nejmenší problémy s rozpoznáním mluvčích VKIF (v první nahrávce ji správně rozpoznalo 100 %, ve druhé 71,4 %) a NSSM (v první nahrávce jej správně určilo 92,9 %, ve druhé 85,7 % posluchačů). Naopak nejobtížněji identifikovatelná byla mluvčí FEEF (v obou projevech byla správně označena 14,3 % posluchačů) a mluvčí NVVM (v jedné položce rozeznán pouze jedním posluchačem - 7,1 %, v druhé 28,5 %). Někdy se míra rozpoznání jednoho mluvčího v obou položkách značně lišila, například NKKM byl poprvé správně identifikován všemi posluchači, podruhé pouze 28,5 %. Podobně také mluvčí FKKM - poprvé jej rozpoznalo 92,9 %, podruhé 35,7 %.

Položka	Maskovaný	Mluvčí 1	Mluvčí 2	Mluvčí 3
0. (první závěčná)	FLIF	0	13	0
0. (druhá závěčná)	FLLF	0	4	9
1.	NPPM	3	9 (3 muži)	2 (1 muž)
18.	NPPM	2	7 (2 muži)	5 (2 muži)
2.	FEEF	0	12 (4 muži)	2
13.	FEEF	2	2 (1 muž)	10 (3 muži)

⁵⁰ Za těmito čísly je v závorce uveden počet mužů - posluchačů, kteří položku označili. Pokud závorka chybí, jedná se pouze o ženy.

⁵¹ Úroveň náhody je 33 %.

3.	FKKM	2 (1 muž)	5 (2 muži)	7 (1 muž)
5.	FKKM	13 (3 muži)	1 (muž)	0
4.	NVVM	6 (1 muž)	1 (muž)	7 (2 muži)
11.	NVVM	7 (1 muž)	3 (1 muž)	4 (2 muži)
6.	VIIM	0	6 (muž)	8 (3 muži)
16.	VIIM	3	9 (3 muži)	2 (1 muž)
7.	VKIF	0	14	0
12.	VKIF	3 (1 muž)	10 (3 muži)	1
8.	NKKM	0	0	14
9.	NKKM	6 (1 muž)	4 (1 muž)	4 (2 muži)
10.	NSSM	2 (1 muž)	0	12 (3 muži)
20.	NSSM	0	1	13 (4 muži)
14.	NKKF	0	1	13 (4 muži)
15.	FMIF	4 (1 muž)	9 (2 muži)	1 (muž)
21.	FMIF	6 (2 muži)	5 (1 muž)	3 (1 muž)
17.	VFFM	2	4 (2 muži)	8 (2 muži)
19.	VFFM	7 (1 muž)	3 (1 muž)	4 (2 muži)

Tabulka 19 Výsledky poslechového testu

Uvedené výsledky naznačují, že rozdíly ve správné identifikaci mluvčího jsou založeny i na jiných aspektech, jakými mohou být například délka projevu, různá hlásková seskupení v něm či pozice mluvčího v testové položce. Porovnávání konkrétních typů vět však nápadnější rozdíly v úspěšnosti identifikace neodhalilo. Potvrdit či odmítnout to, jestli (a do jaké míry) může být rozhodování posluchače odlišnostmi v délce trvání projevu (v rámci počtu slabik) a hláskovém seskupení ovlivněno, však nelze, poněvadž poslechový test byl proveden pouze na malém vzorku lidí. Stejně tak výsledky testu neukazují, že by pořadí mluvčího ve zkoumané položce mělo vliv na posluchačův úsudek. Na stejných pozicích často docházelo k naprosto odlišným výsledkům, a naopak na různých pozicích k podobným. Jaké konkrétní vlivy a jakou měrou se kromě samotného provedení maskování na posluchačově hodnocení ve skutečnosti podílí, by bylo třeba zjišťovat v jiném, samostatném výzkumu.

Pro lepší orientaci byla sestavena tabulka č. 20, která shrnuje úspěšnost identifikace jednotlivých mluvčích z obou dvou výskytů (kromě mluvčí NKKF, která byla do testu

zařazena jen jednou). V ní je možné pozorovat výrazný rozdíl mezi procentem úspěšné identifikace mluvčích FEEF a NVVM a zbytku mluvčích. Tito dva mluvčí byli již dříve přiřazeni (viz kap. 6.1) k nejméně úspěšným, co se jejich schopnosti maskování hlasu týče - a výsledky tohoto testu to potvrzují. Mluvčí FEEF zakryla svůj původní hlas chrapotem a značně ledabylou výslovností, mluvčí NVVM kompletně změnil barvu svého hlasu i způsob mluvy (vyšší F0, jiná intonace, tempo apod.). Dalšími mluvčími, kteří nebyli identifikováni ve více než 50 % případů, byla FMIF, která zkombinovala denazalizaci se sigmatismem a NPPM, který zkombinoval vyšší intonační variabilitu, rotacismus a tlačnou fonaci. Zajímavé by také bylo zjistit, do jaké míry se na tomto rozpoznávání podílely řečové vady některých participantů - mluvčí VKIF trpí rotacismem, u VFFM je zřejmý posun hlásek /d/ a /n/ směrem dopředu (výslovnost na zubech). Za účelem sledování vlivu řečových vad na rozpoznání mluvčího byl projev mluvčího NPPM, který jako součást maskování svého hlasu zvolil rotacismus, v položce zkombinován s přirozeným rotacismem mluvčího NTIM. Ve výsledcích (viz položka 1 a 18 v tab. č. 19) je pak možné pozorovat, že v prvním projevu byl NPPM nesprávně identifikován jako NTIM třemi posluchači, v druhém dokonce sedmi z nich, což poukazuje na možný výrazný podíl řečových vad při identifikaci jedince na základě jeho hlasu. Podobně lze nahlížet i výsledky u mluvčí VKIF, v jejímž druhém maskovaném projevu se objevilo pouze jedno /r/, které mohlo ovlivnit její horší identifikaci než v položce předchozí. To, že řečové vady a podobné odchylky od „normální“ výslovnosti přispívají k rozpoznání mluvčího, samozřejmě není překvapující (srov. např. Jessen, 2007b: 190).

Mluvčí	% rozpoznání
FEEF	14,29
NVVM	17,86
FMIF	50,00
NPPM	50,00
VFFM	53,57
VIIM	60,71
FKKM	64,29
NKKM	64,29
VKIF	85,71
NSSM	89,29
NKKF	92,86

Tabulka 20 Úspěšnost rozpoznání mluvčích

Posluchači se v úspěšnosti rozpoznání mluvčího navzájem mezi sebou výrazně lišili. Nejhorší výsledky měla žena, jež správně identifikovala mluvčí pouze v 7 položkách (33 %), jako jediná také správně neurčila kontrolní mluvčí NKKF, jejíž hlas zněl v obou typech

nahrávek stejně. Nejlepšími posluchači byla žena a muž, kteří správně rozpoznali mluvčí v 16 položkách (76,2 %). Průměrně pak posluchači zodpověděli správně 11,8 otázky z celkových 21, tj. 56,19 %.

Výsledky testu naznačují, že posluchači jsou do jisté míry schopni správně identifikovat mluvčího maskujícího svůj hlas i na základě jedné věty (jejíž minimální délka byla stanovena na 15 slabik). Zároveň však potvrzují, že výrazné rozdíly se nenacházejí jen v různé schopnosti mluvčích maskovat hlas, ale zároveň i v rozpoznávací schopnosti jednotlivých posluchačů. Míra správné identifikace se v některých případech značně lišila i u dvou položek téhož mluvčího, proto by do budoucna bylo jistě zajímavé zjistit, jakým způsobem se na identifikaci mluvčího podílí i další faktory, jako například skladba hlásek v daném projevu.

7 DISKUSE

7.1 Poslechová analýza

Analýza získaných nahrávek ukázala, že mluvčí k maskování svého hlasu využívají poměrně široké spektrum technik. Stejně jako v dříve provedených výzkumech v zahraničí (např. Masthoff, 1996 nebo Masthoff, 2000), nejoblíbenějšími strategiemi provedenými s cílem zakrýt svoji identitu byly fonační změny, zejména snížení či zvýšení základní frekvence hlasu. Otázkou zůstává, z jakého důvodu mluvčí zvýšení či snížení základní frekvence svého hlasu upřednostňují. Pravděpodobným vysvětlením by mohlo být, že podvědomě vnímají, že základní frekvence má nejvyšší podíl na celkové charakteristice jejich hlasu a že její modifikací by mohli docílit právě zakrytí své identity. Jiným možným vysvětlením může být také to, že mírné snížení či zvýšení F0 není tak náročnou technikou.

Další frekventovanou skupinou změn byly změny na prozodické úrovni. Z nich se nejčastěji objevovala vyšší intonační variabilita či snížení tempa řeči. Prozodické změny se však v této míře objevovaly patrně díky povaze obou textů. Zatímco první text byl konstruován do formy novinové zprávy, druhý byl sestaven tak, aby maximálně napodobil výhrušný telefonát. Povaha těchto textů tak způsobila, že zatímco první text byl opravdu velmi často uchopen jako zpráva (čemuž byl přizpůsoben způsob čtení), v druhém se mluvčí často sžili s rolí vyděrače, což se odrazilo právě v prozodických vlastnostech jejich promluv (zejména v jejich intonační variabilitě). Zároveň se v prozodických vlastnostech promluv projeví i některé techniky maskování, například denazalizace vedla v některých případech k nižší variabilitě v řeči či snížení řečového tempa.

Rezonanční změny a změny artikulace společně se změnami jazykovými byly ze čtyř stanovených skupin změn zastoupeny nejméně. Největší podíl v těchto dvou skupinách měla denazalizace, dloužení vokálů a ledabylejší výslovnost. Ledabylejší výslovnost však také do jisté míry souvisí s povahou nahrávaného textu. Její výskyt se zpravidla omezuje na běžnou ústní komunikaci - jejím použitím se tak mluvčí snažili textu dodat autentičnost a přiblížit se reálnému spontánnímu projevu.

Ve shodě s dříve provedenými výzkumy (srov. např. Braun, 2006; Masthoff, 1996; Masthoff, 2000) se ukázalo, že si mluvčí k maskování svého hlasu nejčastěji vybírají pouze jednu nápadnější strategii, maximálně však tři. Tři strategie zkombinovala pouze jedna mluvčí, dvě techniky použilo deset mluvčích z celého výzkumného vzorku. Takový výsledek

bylo možné předpokládat, protože použití více než jedné strategie maskování hlasu představuje větší úsilí při produkci maskované promluvy, a tudíž i větší problémy maskování hlasu udržet. Většině z těch, kteří se uchýlili ke dvěma (příp. třem) strategiím, se však maskování podařilo zachovat konzistentní po celou dobu projevu. Naopak mnohým z těch, již techniky maskování nekombinovali, se to nepodařilo. Neschopnost udržet maskovaný projev konzistentní je daná tím, že mluvčí musí být během maskování svého hlasu schopen společně se zakrytím své identity sdělit zprávu srozumitelnou posluchači (Masthoff, 1996: 166).

Jednotliví mluvčí se v úspěšnosti maskování svého hlasu lišili, avšak ve většině případů nebylo maskování hlasu účinné – mluvčí byli rozpoznatelní již na první či druhý poslech. Přesto bylo možné v rámci každé skupiny vybrat jednoho nebo několik mluvčích, jimž se podařilo dosáhnout velmi efektivního maskování a jejichž identifikace by byla přinejmenším velmi náročná.

S úspěšností maskování jsou spojené i některé další řečové aspekty mluvčího, jako například dialekt, řečové vady či dýchání, které se v promluvách vyskytovaly. Nikdo z mluvčích, v jejichž projevech bylo možné řečové vady, nářeční rysy nebo atypický způsob dýchání zachytit, nebyl zařazen do skupiny těch úspěšnějších.

Participantů byli rozděleni do skupin podle úrovně dosaženého vzdělání kvůli předpokladu, že by se mezi těmito vzdělanostními skupinami mohly objevit významné rozdíly. Přestože se tři zkoumané skupiny mírně lišily a skupina studentů a absolventů fonetiky vykazovala o něco lepší výsledky než zbylé dvě skupiny, významný rozdíl mezi nimi se nepotvrdil⁵². Navíc se ukázalo, že schopnost maskování hlasu jakož i jeho udržení souvisí se zdatností ve čtení. Ta byla nejslabší v poslední skupině mluvčích (tj. mluvčích bez maturity), ale problémy se objevovaly i ve skupině studentů a absolventů nelingvistických oborů. Proto zůstává otázkou, do jaké míry by se výsledky změnily, pokud by mluvčí měli možnost text doma důkladně natrénovat nebo pokud by k maskování hlasu mohli použít řeč spontánní.

7.2 Akustická analýza

Analýza základní frekvence hlasu v nemaskovaných a maskovaných nahrávkách ukázala velké rozdíly mezi mluvčími v rozsahu její změny. Někteří mluvčí svoji základní frekvenci nezměnili vůbec, jiní naopak o více než jednu oktávu. Tato analýza poukázala také

⁵² Provedený chí-kvadrát na počet použitých strategií ukázal $p > 0,05$.

na nápadné rozdíly mezi třemi zkoumanými skupinami, největší odchylky od původní F0 bylo možné pozorovat u skupiny fonetiků (dvě mluvčí snížily základní frekvenci svého hlasu o více než jednu oktávu), nejmenší naopak u poslední zkoumané skupiny – mluvčích bez maturity.

Preference snížení či zvýšení F0 v maskovaném projevu byly poměrně vyrovnané, přičemž u žen se objevovala spíše tendence k jejímu snížení, zatímco u mužů k jejímu zvýšení. Tato tendence však byla slabá a k jejímu potvrzení by bylo třeba nahrát větší množství participantů.

Za relevantní hodnotu základní frekvence byla považovaná základní hladina navrhovaná Lindhem a Erikssonem (2007), protože průměrná hodnota F0 a její medián mohly být ovlivněny oktávovými skoky, které nebyly opravovány. Hodnoty průměrné základní frekvence (a jejího mediánu) se s hodnotami základní hladiny často rozcházely, a to nejen ve velikosti odchylky F0 maskovaného hlasu od hlasu původního, ale někdy dokonce i v určení toho, zda mluvčí základní frekvenci svého hlasu snížil, nebo zvýšil. S tímto jevem do určité míry může souviset i to, že se s hodnotou základní hladiny asi ve dvou případech neshodují ani výsledky získané analýzou poslechovou.

Výsledky měření vokálních formantů se ukázaly jako nejproblematictější část této práce. Přestože naznačují, že společně s maskováním hlasu poměrně často dochází ke změně rezonančních vlastností vokálů, mnoho statisticky významných rozdílů nelze snadno vysvětlit. Jako nejpravděpodobnější se jeví možnost, že maskování hlasu ovlivnilo schopnost Burgova algoritmu správně detekovat některé z formantů a že by bylo potřeba důkladnější kontroly naměřených hodnot. To naznačuje i nutnost vyloučení tří mluvčích ze vzorku, u nichž došlo ke zcela chybné detekci formantů. Přesto se však zdá, že dlouhodobá analýza formantů dokáže reflektovat změny provedené při maskování hlasu. Zároveň je také možné (zejména z předložených grafů) pozorovat, že tato analýza je na základě získaných hodnot do jisté míry schopná poukazovat i na rozdíly mezi jednotlivými mluvčími.

Také měření dalších tří hodnot – jitteru, shimmeru a harmonicity – naznačuje, že má ve forenzním výzkumu své opodstatnění, a to zejména u těch mluvčích, kteří nenapodobovali některé z patologických rysů fonace. Naopak u těch mluvčích, již si k maskování svého hlasu nápodobu těchto rysů vybrali, se často vyskytovaly signifikantní rozdíly mezi oběma typy projevů. Ve výsledcích harmonicity se však u dvou mluvčích ukázal i paradox – měření ukázalo vyšší hodnotu HNR v maskovaném projevu, přestože sluchový dojem byl opačný.

Zdá se, že by harmoničita mohla souviset se spektrálním sklonem, avšak tuto domněnku by samozřejmě bylo třeba ještě prověřit.

Zajímavé je, že akustická analýza vokálních formantů, shimmeru a HNR selhala u mluvčích, která si za účelem maskování svého hlasu položila mikinu před ústa. Problémy se správnou identifikací mluvčích při tomto typu maskování hlasu bývají udávány i ve výzkumech zabývajících se automatickými systémy pro rozpoznání mluvčího. Při poslechové analýze však člověka maskujícího hlas tímto způsobem není těžké během několika prvních vět rozeznat. Takový jev jen dokládá, že oba typy analýz mají své opodstatnění a že by tyto analýzy měly být prováděny společně.

7.3 Poslechový test

Výsledky poslechového textu ukázaly, že i na základě jedné věty o minimální délce 15 slabik jsou posluchači do jisté míry správně identifikovat mluvčího maskujícího svůj hlas. Test také potvrdil rozdílnou míru úspěšnosti maskování hlasu různých mluvčích, ale také rozdílnou schopnost posluchačů správně identifikovat mluvčího (úspěšnost rozpoznání mluvčího posluchači se pohybovala od 33 % do 76 %). Bylo však také zjištěno, že procento správné identifikace se někdy značně liší i mezi dvěma položkami téhož mluvčího. Otázkou tak zůstává, jaké faktory se kromě schopností mluvčího a posluchače na rozpoznání mluvčího podílí.

7.4 Limity práce

Tato diplomová práce se zabývá tématem, které v českém kontextu nebylo dosud zkoumáno. Je ji proto třeba vnímat jako první vhléd do problematiky analýzy maskování hlasu, s čímž je samozřejmě spojeno i mnoho omezení, která tuto práci doprovázejí.

Omezení této práce vyplývají již ze způsobu výběru participantů tohoto výzkumu. Výzkumný vzorek není reprezentativní, a tudíž výsledky mají pouze informativní charakter a rozhodně není možné je generalizovat.

V rámci této práce také nebylo možné zkoumat všechny řečové parametry, jež by pro forenzní výzkum mohly být relevantní. Provedená akustická analýza má sloužit pouze jako první sonda do zkoumaného tématu, její výsledky jsou limitovány tím, že nebyly opravovány. Získané průměrné hodnoty základní frekvence, její medián a směrodatná odchylka jsou zřesleny oktávovými skoky, které vzhledem k rozsáhlému zvukovému materiálu nebylo

možné v rámci této diplomové práce opravovat. Detailněji kontrolovány a opravovány nebyly ani výsledné hodnoty vokálních formantů, odstraněny byly pouze nejnápadnější chyby, a tak je pravděpodobné, že i u těchto výsledných hodnot došlo k jistému zkreslení. Výsledné hodnoty formantů uváděných pro jednotlivé mluvčí jsou ovlivněny i tím, že se v těchto vybraných slovech vyskytuje větší množství zavřených vokálů (z celkových 20 vokálů je pětkrát obsaženo krátké nebo dlouhé /i/, sedmkrát /e/). Další omezení části akustické analýzy (výsledky hodnot vokálních formantů a hodnot HNR, jitteru a shimmeru) může představovat i fakt, že se opírá pouze o segmenty z osmi vybraných slov, což může mít vliv i na prezentované výsledky.

7.5 Návrhy k dalšímu výzkumu

Do budoucna by jistě bylo zajímavé provést detailnější analýzu všech získaných projevů. K podrobnému zkoumání forenzně relevantních rysů by bylo užitečné všechny promluvy transkribovat, a tak zachytit i jevy, které mohou zůstat i během důkladného a opakovaného poslechu nepostřehnuty. Pro potřeby přesnější akustické analýzy by bylo vhodné opravit chyby, které mohou naměřené hodnoty významným způsobem ovlivňovat. Díky opravě oktávových skoků v kontuře základní frekvence by tak v budoucnosti mohly být například porovnány výsledky průměrné základní frekvence, mediánu a alternativní základní hladiny. Takto získané výsledky by pak mohly být porovnány i s výsledky poslechové analýzy, protože, jak již bylo řečeno, i ta se svými závěry někdy s výslednou hodnotou základní hladiny rozchází. Analýza vokálních formantů by pak mohla být zpřesněna nejen opravením chyb v naměřených hodnotách, ale také zvýšením počtu zkoumaných hlásek a vyvážením výskytu jednotlivých vokálů. Navýšením počtu zkoumaných hlásek by došlo i ke zpřesnění výsledků dalších naměřených hodnot – harmonicity, jitteru a shimmeru.

V rámci akustické analýzy provedené v této práci nebyly zkoumány i některé další forenzně relevantní jevy, jako například trvání jednotlivých typů hlásek či artikulační tempo. Stejně tak nebyla provedena analýza dlouhodobého průměrného spektra, které je považováno za důležitý ukazatel kvality hlasu. Aby bylo možné důkladně porozumět jednotlivým typům maskování hlasu a jeho projevům v řeči mluvčího, další výzkum by se měl zaměřit na tyto jmenované i další aspekty.

Záměrem této práce bylo mimo jiné také položit základ databázi maskovaných hlasů, kterou by bylo vhodné do budoucna dále rozšiřovat. Rozšíření této databáze by pak pomohlo

nejenom potvrdit či vyloučit rozdíly mezi skupinami mluvčích zkoumanými v této práci, ale větší množství mluvčích by také přispělo ke snadnější rekonstrukci původního hlasu mluvčího a pochopení specifík jednotlivých typů maskování hlasu. Tato databáze by mohla být rozšířena i o různé podmínky nahrávání, například simulaci telefonního hovoru či poskytnutí několikadenní přípravy na výběr maskování hlasu.

8 ZÁVĚR

Tato práce se zabývala vlivem vzdělání na schopnost maskovat svůj hlas i maskováním hlasu obecně. Byly získány nahrávky čteného nemaskovaného a maskovaného projevu 43 mluvčích, kteří byli zařazeni do jedné ze tří předem stanovených skupin stanovených na základě úrovně a typu dosaženého vzdělání. První skupinu tvořili absolventi a studenti vyšších ročníků oboru Fonetika na Univerzitě Karlově v Praze, druhou studenti a absolventi nelingvistických oborů na vysokých školách v Praze a Plzni, do poslední skupiny byli zahrnuti mluvčí, kteří dosáhli nižšího než úplného středoškolského vzdělání.

Cílem této práce bylo na základě pořízených dat zjistit, jaké techniky maskování hlasu si mluvčí za účelem skrytí své identity nejčastěji vybírají, zda využívají i jejich kombinací a jak jsou mluvčími použité strategie maskování hlasu efektivní. Zároveň se tato práce zaměřovala na to, jestli a do jaké míry se jednotlivé skupiny participantů mezi sebou liší ve výběru těchto strategií, v jejich kombinacích i v úspěšnosti zamaskování svého hlasu. V teoretické části práce byla představena sociofonetika a forenzní fonetika s jejich metodami i možnostmi jejich využití a byl rozebrán fenomén maskování hlasu i výsledky některých dřívějších studií. V praktické části práce pak byly prezentovány výsledky poslechové a akustické analýzy, které byly doplněny o ilustraci některých příkladů z konkrétních promluv. Praktickou část uzavírají výsledky poslechového testu provedeného na 14 posluchačích.

Výsledky provedeného výzkumu poukazují na to, že nejfrekventovanějšími změnami hlasu za účelem jeho zamaskování byly změny fonační, zejména snížení nebo zvýšení základní frekvence. Mezi další oblíbené strategie maskování hlasu patřily prozodické změny, jejichž výskyt však byl do jisté míry ovlivněn povahou textů použitých při nahrávání řečového materiálu. Naopak nejméně byly zastoupeny rezonanční změny a změny artikulace společně se změnami jazykovými. Dále bylo zjištěno, že se participantů nejčastěji uchylovali pouze k jedné nápadné technice maskování hlasu, maximálně však ke třem. Úspěšnost jednotlivých strategií se v závislosti na mluvčím lišila, ve většině případů však maskování hlasu nebylo dostatečně účinné.

Mezi zkoumanými skupinami bylo sice možné pozorovat rozdíly, avšak ne v takové míře, jak bylo na začátku výzkumu předpokládáno. Kombinace více než jedné strategie maskování stejně jako mluvčí úspěšně maskující svůj hlas se objevili ve všech třech zkoumaných skupinách. Přesto se zdá, že studenti a absolventi fonetiky jsou schopni svůj hlas

maskovat o něco lépe a že naopak skupina mluvčích bez maturity má tuto schopnost oproti zbylým dvěma skupinám oslabenu. Tuto tendenci naznačují i výsledky analýzy základní frekvence - v maskovaném projevu byly odchylky od původní F0 u fonetiků největší, zatímco u mluvčích bez maturity naopak. Aby však bylo možné tuto domněnku potvrdit či vyvrátit, bylo by třeba získat nahrávky dalších mluvčích spadajících do těchto skupin.

Výsledky poslechového testu potvrzují nejenom rozdílnou schopnost mluvčích maskovat svůj hlas, ale zároveň i rozdílnou schopnost posluchačů správně mluvčího identifikovat. Také se zdá, že na správné identifikaci mluvčího se podílí i další faktory, jako například skladba hlásek v přehrávaném projevu.

Přese všechna omezení, která tuto práci doprovází, se podařilo získat poměrně bohatý řečový materiál a poskytnout tak základ databázi maskovaných hlasů, která by mohla být v budoucnosti rozvíjena a využívána k dalšímu výzkumu. Kromě toho výsledky této práce otvírají další otázky spojené s fenoménem maskování hlasu a nabízí další možnosti směřování jeho zkoumání.

SEZNAM POUŽITÝCH ZDROJŮ

- Amino, K., & Arai, T. (2009). Speaker-dependent characteristics of the nasals. *Forensic Science International*, 185, 21-28.
- Anolli, L., & Ciceri, R. (1997). The voice of deception: Vocal strategies of naive and able liars. *Journal of Nonverbal Behavior*, 21 (4), 259-284.
- Baldwin, J., & French, P. (1990). *Forensic Phonetics*. London/New York: Pinter Publishers.
- Baranowski, M. (v tisku). Sociophonetics. In R. Bayley, R. Cameron, & C. Lucas (Eds.), *The Oxford handbook of sociolinguistic*. Oxford: Oxford University Press.
- Blatchford, H., & Foulkes, P. (2006). Identification of voices in shouting. *Journal of Speech, Language and the Law*, 13 (2), 241-254.
- Boersma, P., & Weenink, D. (2012) Praat: doing phonetics by computer (Verze 5.3.11) [software]. Získáno 20.08.2012 z http://www.fon.hum.uva.nl/praat/download_win.html
- Braun, A. (2006). Stimmverstellung und Stimmenimitation in der forensischen Sprechereerkennung. In T. Kopfermann (Ed.), *Das Phänomen Stimme. Imitation und Identität* (s. 177-180). St. Ingbert: Röhrig Universitätsverlag.
- Broeders, A. P. A., & van Amelsvoort, A. G. (2001). A practical approach to forensic earwitness identification: Constructing a voice line-up. *Problems of Forensic Sciences*, 47, 237-245.
- Broeders, A. P. A., & van Amelsvoort, A. G. (1999). Lineup construction for forensic earwitness identification: A practical approach. *ICPhS*, San Francisco, 1373-1376.
- Butcher, A. (2002). *Forensic phonetics: Issues in speaker identification evidence*. Získáno 28.06.2012 z <http://www.scribd.com/doc/55778545/Forensic-Phonetics>
- Butcher, A. (1996). *Getting the voice line-up right: Analysis of multiple auditory confrontation*. 97-102. Získáno 30.06.2012 z <http://www.assta.org/sst/SST-96/cache/SST-96-Chapter5-p3.pdf>
- Clark, J., & Foulkes, P. (2007). Identification of voices in electronically disguised speech. *Journal of Speech, Language and the Law*, 14 (2), 195-221.

Drager, K. (2010). Sociophonetic variation in speech production. *Language and Linguistics Compass*, 4 (7), 473-480.

Drager, K., & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. Získáno 28.06.2012 z <http://www2.hawaii.edu/~kdrager/drager&hay2012.pdf>

de Figueiredo, R. M., & de Souza Britto, H. (1996). A report on the acoustics effects of one type of disguise. *Forensic Linguistics*, 3 (1), 168-175.

Eriksson, E., Sullivan, K., Zetterholm, E., Czigler, J. G., Skagerstrand, & van Doorn, J. (2010). Detection of imitated voices: who are reliable earwitnesses? *The International Journal of Speech, Language and the Law*, 17 (1), 25-44.

Farrús, M., Hernando, J., & Ejarque, P. (2007). Jitter and shimmer measurements for speaker recognition. Získáno 01.01.2013 z http://nlp.lsi.upc.edu/papers/far_jit_07.pdf

Farrús, M., Wagner, M., Anguita, J., & Hernando, J. (2008). How vulnerable are prosodic features to professional imitators? *Odyssey 2008: The Speaker and Language Recognition Workshop Stellenbosch*, Jižní Afrika, 21-24. leden.

Foulkes, P. (2002). Current trends in British sociophonetics. Získáno 28.06.2012 z <http://www-users.york.ac.uk/~pf11/PennWP2002.pdf>

Foulkes, P. (2005). Sociophonetics. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (s. 495-499). Amsterdam: Elsevier Press.

Foulkes, P., & Barron, A. (2000). Telephone speaker recognition amongst members of a close social network. *Forensic Linguistics* 7 (2), 180-198.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34, 409-438.

Foulkes, P., & French, J. P. (2001). *Forensic phonetics and sociolinguistic*. Získáno 27.06.2012 z <http://www-users.york.ac.uk/~pf11/foulkes&french.pdf>

Foulkes, P., Scobbie, J. M., & Watt, D. (2010). Sociophonetics. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (s. 703-754), Singapur: Wiley-Blackwell.

- Hansen, J. H. L., & Patil, S. (2007). Speech under stress: Analysis, modeling and recognition. In C. Müller (Ed.), *Speaker Classification I* (s. 108-137). Berlin/Heidelberg: Springer Verlag.
- Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics*, 34, 439-457.
- Harrington, J., Palethorpe, S., & Watson, C. I. (2000). Does the Queen speak the Queen's English? *Nature*, 408, 927-928.
- Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, 36, 89-103.
- Hirson, A., & Duckworth, M. (1993). Glottal fry and voice disguise: a case study in forensic phonetics. *Journal of Biomedical Engineering*, 15, 193-200.
- Hollien, H. (2002). *Forensic Voice Identification*. London/San Diego: Academic Press.
- Jessen, M. (2008). Forensic phonetics. *Language and Linguistics Compass*, 2 (4), 671-711.
- Jessen, M. (2007a). Forensic reference data on articulation rate in German. *Science and Justice*, 47, 50-67.
- Jessen, M. (2007b). Speaker classification in forensic phonetics and acoustic. In C. Müller (Ed.), *Speaker Classification I* (s. 180-204). Berlin/Heidelberg: Springer-Verlag.
- Jessen, M. (2010). The forensic phonetician. Forensic speaker identification by experts. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (s. 378-394). London/New York: Routledge.
- Jessen, M., & Jessen, M. (2008). Forensische Sprechererkennung und Tonträgerauswertung in Praxis und Forschung - Teil 1. *Die Kriminalpolizei*, (4), 136-139.
- Jessen, M., & Jessen, M. (2009). Forensische Sprechererkennung und Tonträgerauswertung in Praxis und Forschung - Teil 2. *Die Kriminalpolizei*, (1), 30-34.
- de Jong, G., McDougall, K., & Nolan, F. (2007). Sound change and speaker identity: An acoustic study. In C. Müller (Ed.), *Speaker Classification II* (s. 130-141). Berlin/Heidelberg: Springer-Verlag.
- Jovičić, S. T., & Šarić, Z. (2008). Acoustic analysis of consonants in whispered speech. *Journal of Voice*, 22 (3), 263-274.

- Köster, O., Jessen, M., Khairi, F., & Eckert, H. (2007). Auditory-perceptual identification of voice quality by expert and non-expert listeners. *ICPhS*, Saarbrücken, 6-10.08.2007, 1845-1848.
- Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Oxford: Wiley-Blackwell.
- Künzel, H. J. (1994). Current approaches to forensic speaker recognition. ESCA Workshop on Automatic Speaker Recognition, Identification, and Verification. Martigny, Švýcarsko. 5-7. 4. 1994, s. 135-141. Získáno 28.06.2012 z http://www.isca-speech.org/archive_open/archive_papers/asriv94/sr94_135.pdf
- Künzel, H. J. (2000). Effect of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7 (2), 149-179.
- Künzel, H. J., Gonzales-Rodriguez, J., & Ortega-García, J. (2004). Effect of voice disguise on the performance of a forensic automatic speaker recognition system. ODYSSEY04 – The Speaker and Language Recognition Workshop. Toledo, 31.5.-3.06.2004.
- Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, 34, 500-515.
- Lavner, Y., Gath, I., & Rosenhouse, J. (2000). The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels. *Speech Communication*, 30, 9-26.
- Lindh, J. (2005-2006). Collect_F0_data_incl_altbase.praat. Získáno 15.08.2012 z <https://sites.google.com/site/phoneticolindh/home/scripts-and-audio/scripts>
- Lindh, J., & Eriksson, A. (2007). Robustness of long time measures of fundamental frequency. In *INTERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association*, Antwerp, Belgie, 27-31. srpen, 2007, 2025-2028.
- Machač, P., & Skarnitzl, R. (2009). *Fonetická segmentace hlásek*. Praha: Epocha.
- Markham, D. (1999). Listeners and disguised voices: the imitation and perception of dialectal accent. *Forensic Linguistics*, 6 (2), 289-299.
- Masthoff, H. (1996). A report on voice disguise experiment. *Forensic Linguistics*, 3 (1), 160-167.

Masthoff, H. (2000). Die Auswirkung von Stimmverstellung auf ausgewählte phonetische Merkmale. In B. Nolte (Ed.), *Anwendungen der Akustik in der Wehrtechnik. Tagung am 26., 27. und 28. September 2000* (s. 58-66). Hamburg: Deutsche Gesellschaft für Wehrtechnik, Arbeitskreis Anwendungen der Akustik in der Wehrtechnik.

McDougall, K., Nolan, F., Harrison, P., & Kirchhübel, C. (2012). Characterising speakers using formant frequency information: A comparison of vowel formant measurements and long-term formant analysis. Získáno 21.10.2012 z http://www.iafpa2012.com/AbstracsPDF/MCDOUGALL_KIRSTY.pdf

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34, 202-240.

Neuhauser, S. (2008). Voice disguise using a foreign accent: phonetic nad linguistic variation. *The International Journal of Speech, Language and the Law*, 15 (2), 131-159.

Nolan, F. (2003). A recent voice parade. *Forensic Linguistics*, 10 (2), 277-291.

Nolan, F. (1999). Speaker recognition and forensic phonetics. In W. J. Hardcastle & J. Laver (Eds.), *The Handbook of Phonetic Sciences*. Blackwell Publishing. Blackwell Reference Online. Získáno 28.12.2007 z http://www.blackwellreference.com/subscriber/uid=532/tocnode?id=g9780631214786_chunk_g978063121478625

Nolan, F., McDougall, K., de Jong, G., & Hudson, T. (2006). A forensic study of 'dynamic' sources of variability in speech: The DyViS project. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Sciences & Technology*, 6-8. prosinec 2006, Auckland: Australasian Speech Science and Technology Association, 13-18.

Orchard, T. L., & Yarmey, A. D. (1995). The effects of whispers, voice-sample duration, and voice distinctiveness on criminal speaker identification. *Applied Cognitive Psychology*, 9, 249-260.

Palková, Z. (1997). *Fonetika a fonologie češtiny*. Praha: Karolinum.

Perrot, P., Aversano, G., & Chollet, G. (2007a). Voice disguise and automatic detection: Review and perspectives. In Y. Stylianon, M. Faundez-Zanny, & A. Espozito (Eds.),

Workshop on Nonlinear Speech Processing 2005, LNCS 4391 (s. 101-117). Berlin/Heidelberg: Springer Verlag.

Perrot, P., & Chollet, G. (2008). The question of disguised voice. *Acoustics*, 29.6.-4. 7. 2008, Paris, 5681-5685.

Perrot, P., Morel, M., Razik, J., & Chollet, G. (2009). Vocal forgery in forensic sciences. In M. Surel (Ed.), *e-Forensics*, LNICST 8 (s. 179-185).

Perrot, P., Preteux, C., Vasseur, S., & Chollet, G. (2007b). Detection and recognition of voice disguise. *Proceedings, IAFPA 2007*, The College of St Mark & St John, Plymouth.

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America*, 116 (4), 1905-1908.

Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistic*, 11 (4), 478-504.

Pollák P., Volín J., & Skarnitzl R. (2007). Accuracy analysis of phonetic segmentation with multiple word-pronunciation variants and segmentation tool in Praat environment. In R. Vích (Ed.), *17th Czech-German Workshop - Speech Processing* (s. 37-42), Praha.

Rak, R., Matyáš, V., & Říha, Z. a kol. (2008). *Biometrie a identita člověka ve forezních a komerčních aplikacích*. Praha: Grada Publishing.

Reich, A. R., & Duke, J. E. (1979). Effects of selected vocal disguises upon speaker identification by listening. *Journal of Acoustical Society of America* 66 (4), 1023-1028.

Reich, A. R., Moll, K. L., & Curtis, J. F. (1976). Effects of selected vocal disguises upon spectrographic speaker identification. *Journal of Acoustical Society of America*, 60 (4), 919-925.

Rodman, R. D. (1998). Speaker recognition of disguised voices: A program for research. *Consortium on Speech Technology Conference on Speaker by man and machine: direction for forensic applications*, COST 250, Turecko. Získáno 26.06.2012 z <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=191B2E2A13D6823B286FFFB78F92E1EC?doi=10.1.1.116.7877&rep=rep1&type=pdf>

Rodman, R., McAllister, D., Bitzer, D., Cepeda, L., & Abbit, P. (2002). Forensic speaker identification based on spectral moments. *Forensic Linguistics*, 9 (1), 22-43.

Rodman, R. D., & Powell, M. S. (2000). Computer recognition of speakers who disguise their voice. Získáno 28.06.2012 z <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.138.2110&rep=rep1&type=pdf>

Rose, P. (2002). *Forensic Speaker Identification*. London/New York: Taylor & Francis.

Schlichting, F., & Sullivan, K. P. H. (1997). The imitated voice - a problem for voice line-ups? *Forensic Linguistics*, 4 (1), 148-165.

Shama, K., Krishna, A., & Cholayya, N. U. (2007). Study of harmonics-to-noise-ratio and critical-band energy spectrum of speech as acoustic indicators of laryngeal and voice pathology. *EURASIP Journal of Advances in Signal Processing*, 2007.

Sjöström, M., Eriksson, E. J., Zetterholm, & E., Sullivan, K. P. H. (2008). A bidialectal experiment on voice identification. *Working papers* 53, 145-158.

Smith, E. A., Hall, K. C., & Munson, B. (2010). Bringing semantics to sociophonetics: Social variables and secondary entailments. Získáno 28.06.2012 z http://www.tc.umn.edu/~munso005/SmithHallMunson_InPress.pdf

Thomas, E. R. (2011). *Sociophonetics. An Introduction*. Hampshire: Palgrave Macmillan.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. *American Speech*, 77 (2), 115-147.

Thomas, E. R. (2000). Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics*, 28, 1-25.

Thurgood, G. (2000). Voice quality differences and the origin of diphthongs. *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistic Society: General Session and Parasession on Aspect*, 295-303.

Walker, A. (2007). The effect of phonetic detail on perceived speaker age and social class. *Proceeding of the International Congress of Phonetic Sciences, Germany*, 16, 1453-1456.

Watt, D. (2009). The identification of the individual through speech. Získáno 30.06.2012 z <http://www-users.york.ac.uk/~dw539/watt2009.pdf>

- Wolfe, V., & Fitch, J. (1995). Acoustic prediction of severity in commonly occurring voice problems. *Journal of Speech and Hearing Research*, 38 (2), 273-279.
- Yarmey, A. D., Yarmey, L., Yarmey, M. J., & Parliament, L. (2001). Commonsense beliefs and the identification of familiar voices. *Applied Cognitive Psychology*, 15, 283-299.
- Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise-ratio as an index of the degree of hoarseness. *Journal of Acoustic Society of America*, 71 (6), 1544-1550.
- Zetterholm, E. (2000). An impersonator and some of his voices. A phonetic study of human voice flexibility. Proceedings Fonetik 2000, The Swedish Phonetics Conference, 145-148.
- Zetterholm, E. (2007). Detection of speaker characteristics using voice imitations. In C. Müller (Ed.), *Speaker Classification II*, LNAI 4441 (s. 192-205). Berlin/Heidelberg: Springer-Verlag.
- Zetterholm, E. (2006). Same speaker - different voices. A study of one impersonator and some of his different imitations. In *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, 2006, 70–75.
- Zetterholm, E., Sarvar, F., & Allwood, C. M. (2009). Earwitnesses: The effect of voice differences in identification accuracy and the realism in confidence judgments. Proceedings, FONETIK.
- Zhang, C., & Tan, T. (2008). Voice disguise and automatic speaker recognition. *Forensic Science International*, 175, 118-122.

PŘÍLOHY

Příloha 1 Informovaný souhlas

Já, (uved'te jméno), souhlasím s tím, aby mnou poskytnuté řečové projevy a/nebo údaje z poslechových testů byly použity pro účely vědeckého výzkumu.

Fonetický ústav FF UK se zavazuje, že níže uvedené osobní údaje ani jméno nebudou zveřejněny a budou uchovávány odděleně od samotných výzkumných materiálů.

Pohlaví:

Věk:

Povolání:

Jste rodilým mluvčím češtiny?

Ano, prosím uveďte kraj, kde jste prožili dětství:

Ne, prosím uveďte rodný jazyk:

Jak dlouho žijete v České republice?

Prosím uveďte další skutečnosti, které mohly ovlivnit Vaši výslovnost (pobyt v zahraničí, cizí jazyky, nářečí, rodný jazyk nebo nářečí rodičů atd.):

.....
.....

Uveďte prosím případné vady řeči nebo sluchu (návštěvy logopeda v dětství, dyslexie apod.):

.....
.....

Kouříte, případně kouřil(a) jste v minulosti? Přibližně kolik?

.....

Uveďte prosím emailovou adresu (či jiný kontakt), abychom se s Vámi mohli v případě potřeby spojit. Děkujeme.

.....

Datum: Podpis:.....

Příloha 2 Text 1 – normální čtený projev

Výhrůžky kvůli televizní reportáži

Jan Novák z televize Ona byl nedávno vydírán v souvislosti s natáčením reportáže, jež by se měla v televizi objevit tento týden. Vyděrač, který Janu Novákovi zavolal na mobilní telefon, se představil jako zatím ještě jeho přítel. Muž na druhém konci drátu projevil nejprve radost nad tím, že reportér vynechal schůzku u firmy Lopos. Ta je totiž jedním z hlavních aktérů, o nichž reportáž Jana Nováka pojednává. Reportérovi bylo doslova řečeno, že by vyděrač jen velmi nerad podnikal další kroky v této věci, Novák by měl být rozumný, určitě by mu to nestálo za další problémy. Novákovi bylo vyhrožováno tím, že vyděrač přesně ví, kde se pohybuje, a tudíž není problém jej najít. Nezůstalo pouze u jeho osoby, pod pohrůžku se dostala i reportérova bližší rodina. Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy. Prý mu nepřeje, aby se s ním či jeho rodinou museli setkat osobně. Kdo celou záležitost řídí, však Novák neznámému muži neprozradil.

Reportér již byl jednou varován při první reportáži, která se na dané téma točila. Nedávný rozhovor je prý třeba vnímat jako varování poslední. Pokud reportér natočí další reportáž, výhrůžky se prý již zrealizují.

Podle Policie ČR by bylo lepší vyvstalé problémy přenechat jejím složkám. Firma, o níž reportáž pojednává, si očividně nepřeje, aby reportér ve své práci pokračoval. Identita vyděrače stále zůstává nejasná.

Příloha 3 Text 2 – maskovaný projev

Pan Novák? Dobrý den, zdravím Vás, tady zatím ještě váš přítel. Jsem rád, že jste dnes vynechal schůzku u firmy Lopos. Zdá se, že jste po úterním telefonátu věc pochopil. Jen velmi nerad bych podnikal další kroky v této věci, buďte rozumný, určitě by vám to nestálo za to. Víme, kde se pohybujete, takže není problém si vás najít. Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy. Určitě byste nerad, abysme se s vámi nebo s vaší rodinou museli setkat osobně.

Kdo celou věc řídí, mi asi jen tak nebudete chtít povědět, že? Kdo za tím je. Jako na čí objednávku tohleto celé děláte. Neříkejte, že vám to vnuknul duch svatej. Byl jste už jednou varován při první reportáži, zatím pouze ústně. Tento hovor berte jako varování poslední. Jestli natočíte další reportáž, bude asi následovat spoustu jiných věcí. Ne nadarmo vám tady tímto způsobem vyhrožuji. Ano? Doufám, že jste věc řádně pochopil, takže... Věřte mi, můj klient nedává druhé šance.

To asi opravdu nebude lepší, když nějaké problémy budete zjišťovat vy. Firma, která si mě zavolala, si prostě nepřeje, abyste ve věci pokračoval. Ano?

Doufám, že jste pochopil, že tohleto je poslední varování. Ve vašem zájmu... Ve vašem zájmu by bylo lepší, kdybyste se na to vyprd. Nashle.

Příloha 4 Věty použité v poslechovém testu

0 (první zácvičná) – Muž na druhém konci drátu projevil nejprve radost nad tím, že reportér vynechal schůzku u firmy Lopos.

Jsem rád, že jste dnes vynechal schůzku u firmy Lopos.

0 (druhá zácvičná) – Muž na druhém konci drátu projevil nejprve radost nad tím, že reportér vynechal schůzku u firmy Lopos.

Jsem rád, že jste dnes vynechal schůzku u firmy Lopos.

1 – Reportérovi bylo doslova řečeno, že by vyděrač jen velmi nerad podnikal další kroky v této věci.

Jen velmi nerad bych podnikal další kroky v této věci.

2 – Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy.

Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

3 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

4 – Muž na druhém konci drátu projevil nejprve radost nad tím, že reportér vynechal schůzku u firmy Lopos.

Jsem rád, že jste dnes vynechal schůzku u firmy Lopos.

5 – Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy.

Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

6 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

7 – Muž na druhém konci drátu projevil nejprve radost nad tím, že reportér vynechal schůzku u firmy Lopos.

Jsem rád, že jste dnes vynechal schůzku u firmy Lopos.

8 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

9 – Prý mu nepřeje, aby se s ním či jeho rodinou museli setkat osobně.

Určitě byste nerad, abysme se s vámi nebo s vaší rodinou museli setkat osobně.

10 – Reportérovi bylo doslova řečeno, že by vyděrač jen velmi nerad podnikal další kroky v této věci.

Jen velmi nerad bych podnikal další kroky v této věci.

11 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

12 – Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy.

Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

13 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

14 – Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy.

Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

15 – Prý mu nepřeje, aby se s ním či jeho rodinou museli setkat osobně.

Určitě byste nerad, abysme se s vámi nebo s vaší rodinou museli setkat osobně.

16 – Prý mu nepřeje, aby se s ním či jeho rodinou museli setkat osobně.

Určitě byste nerad, abysme se s vámi nebo s vaší rodinou museli setkat osobně.

17 – Reportérovi bylo doslova řečeno, že by vyděrač jen velmi nerad podnikal další kroky v této věci.

Jen velmi nerad bych podnikal další kroky v této věci.

18 – Reportér již byl jednou varován při první reportáži, která se na dané téma točila.

Byl jste už jednou varován při první reportáži, zatím pouze ústně.

19 – Vyděrač Novákovi oznámil, že stejně tak ví, kde pracuje jeho žena a kam chodí jeho děti do školy.

Stejně tak víme, kde pracuje vaše žena a kam chodí vaše děti do školy.

20 – Prý mu nepřeje, aby se s ním či jeho rodinou museli setkat osobně.

Určitě byste nerad, abysme se s vámi nebo s vaší rodinou museli setkat osobně.

21 – Reportérovi bylo doslova řečeno, že by vyděrač jen velmi nerad podnikal další kroky v této věci.

Jen velmi nerad bych podnikal další kroky v této věci.

Příloha 5 Instrukce k poslechovému testu

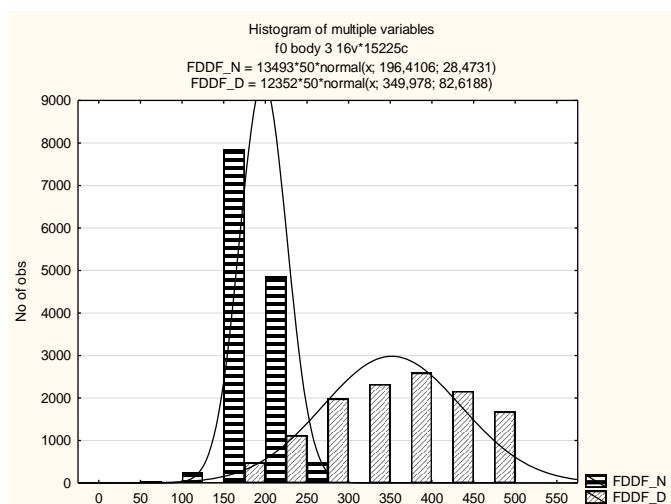
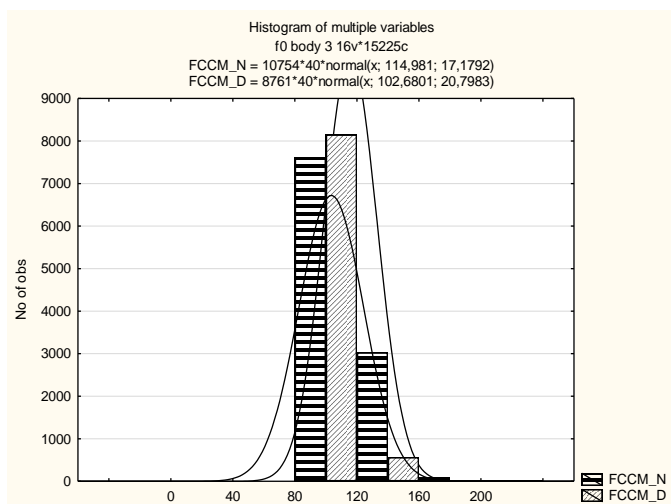
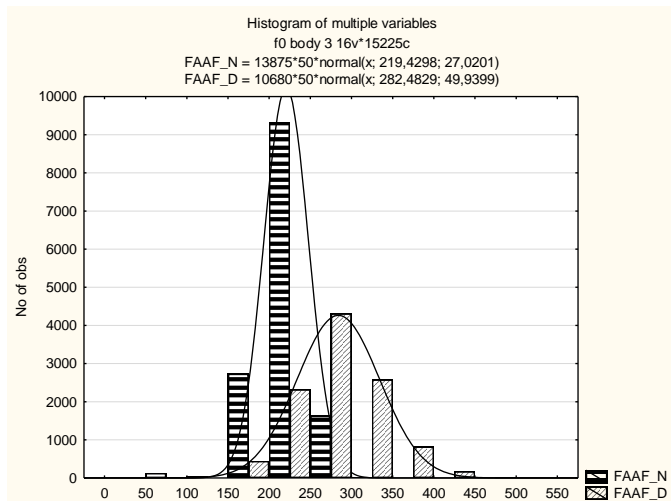
Instrukce

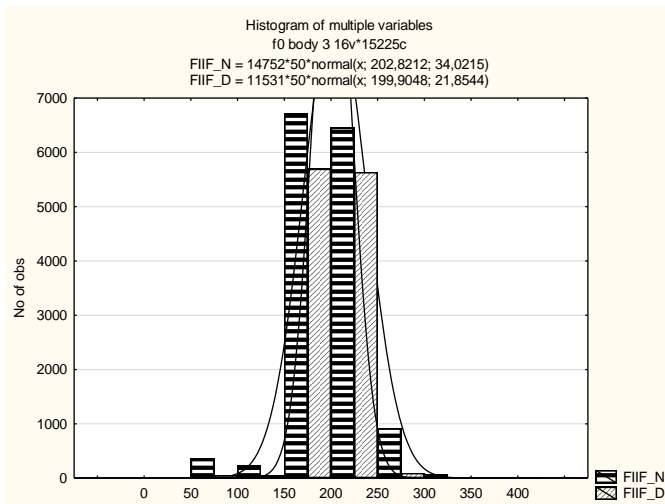
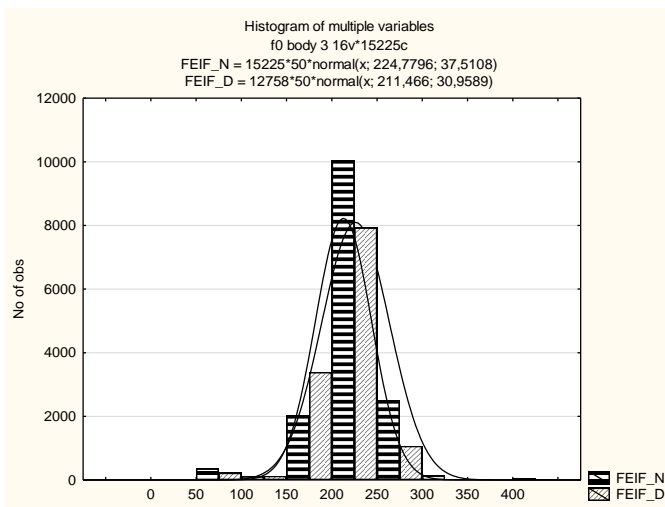
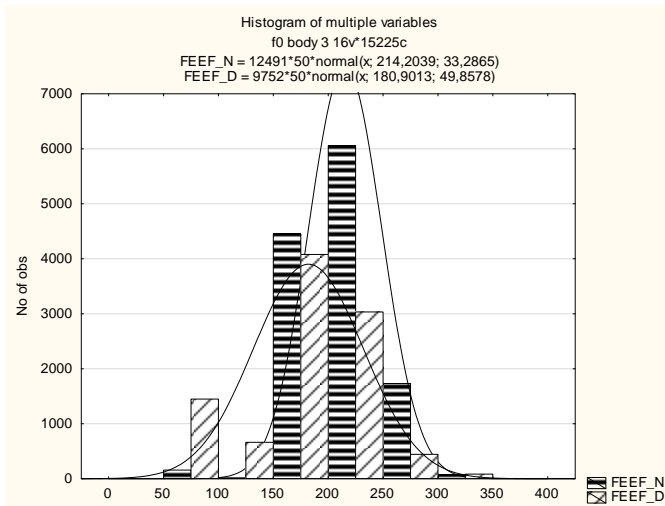
V následujícím poslechovém testu uslyšíte vždy jeden maskovaný hlas, po něm tři hlasy nezměněné. Vaším úkolem je určit, který z těchto tří nezměněných hlasů odpovídá hlasu maskovanému. Ten pak označte ve Vaší tabulce. Všechny čtyři hlasy uslyšíte vždy dvakrát po sobě, první dvě položky jsou zácvičné.

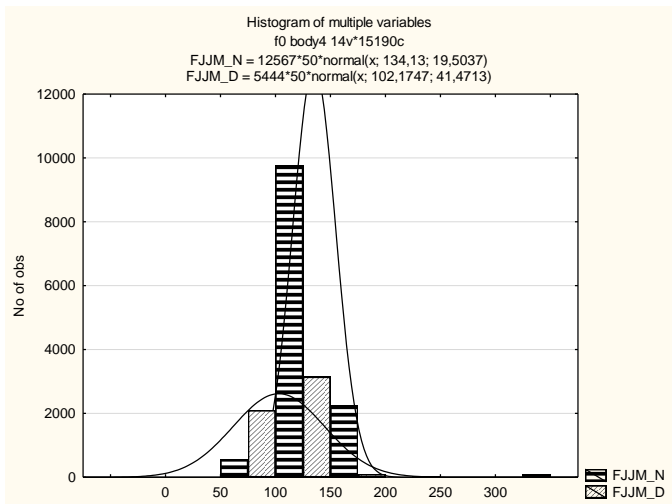
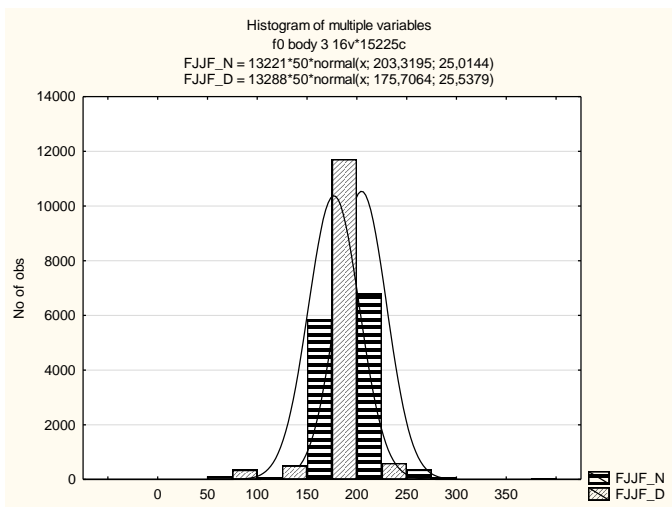
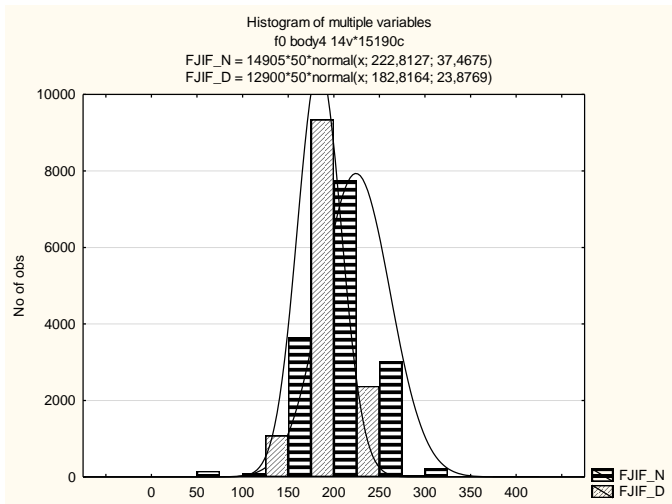
Položka	Mluvčí	Mluvčí	Mluvčí
0. (první zácvičná)	1	2	3
0. (druhá zácvičná)	1	2	3
1.	1	2	3
2.	1	2	3
3.	1	2	3
4.	1	2	3
5.	1	2	3
6.	1	2	3
7.	1	2	3
8.	1	2	3
9.	1	2	3
10.	1	2	3
11.	1	2	3
12.	1	2	3
13.	1	2	3
14.	1	2	3
15.	1	2	3
16.	1	2	3
17.	1	2	3
18.	1	2	3
19.	1	2	3
20.	1	2	3
21.	1	2	3

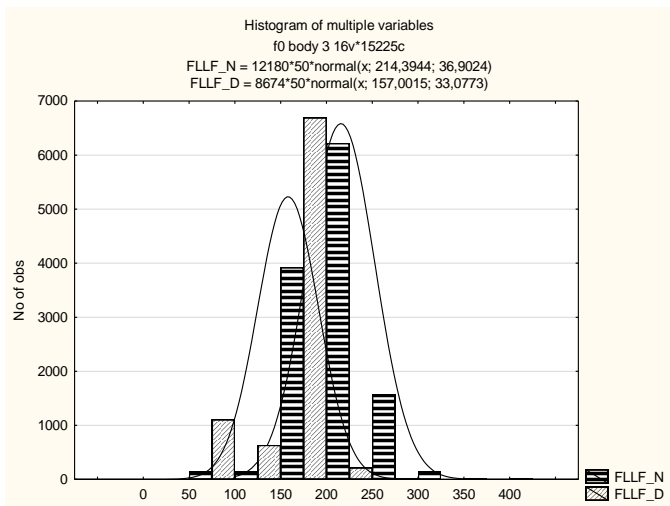
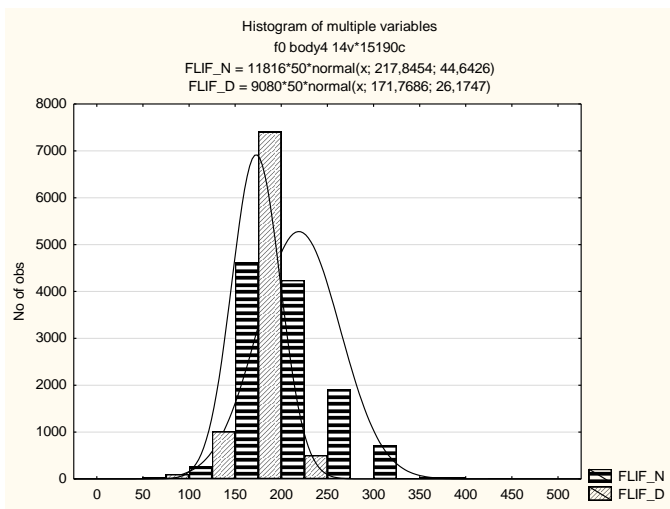
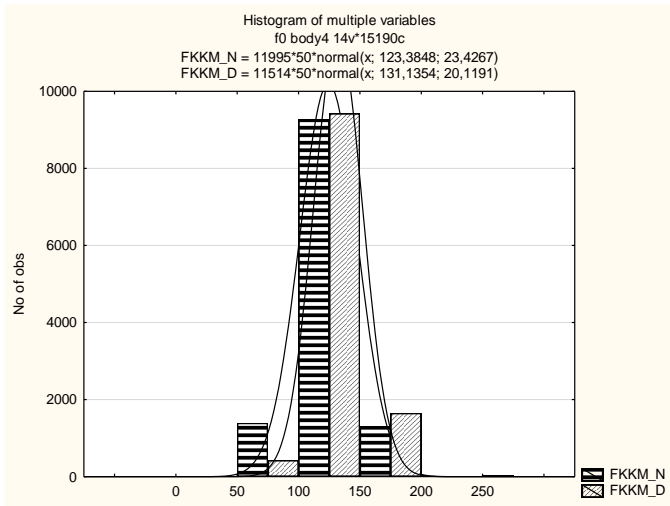
Příloha 6 Histogramy

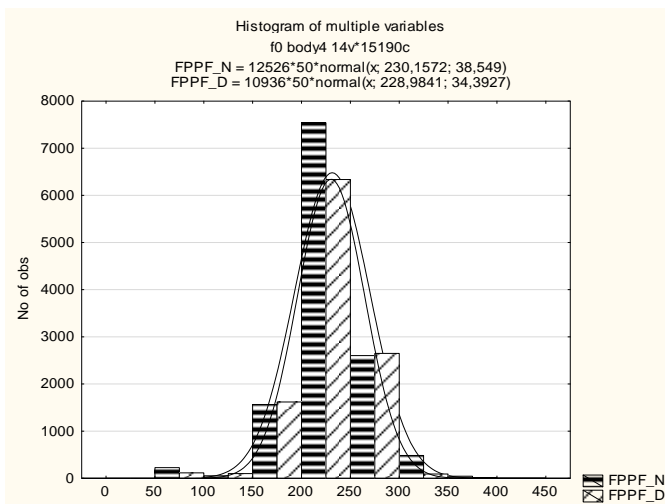
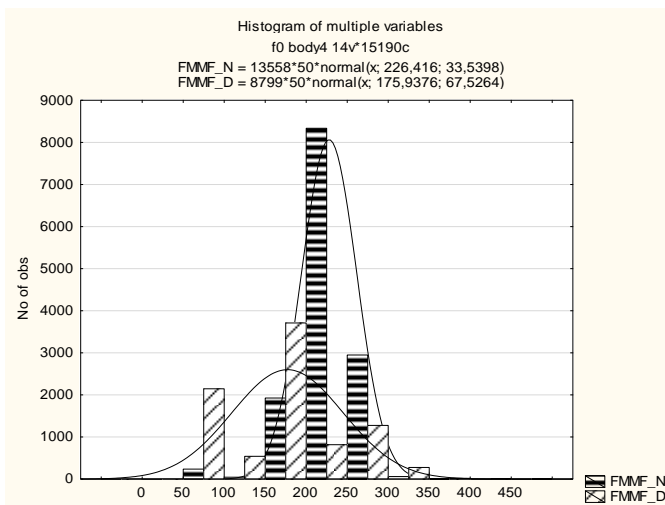
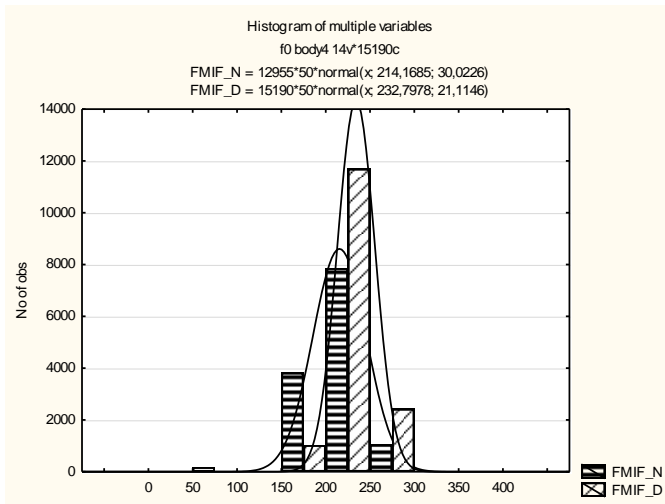
Fonetici



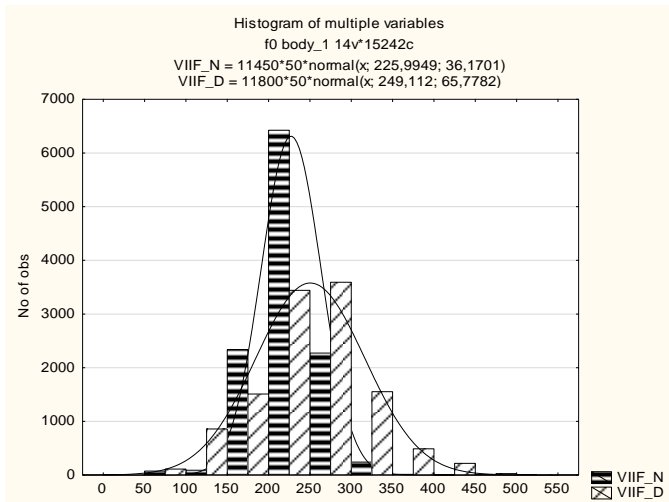
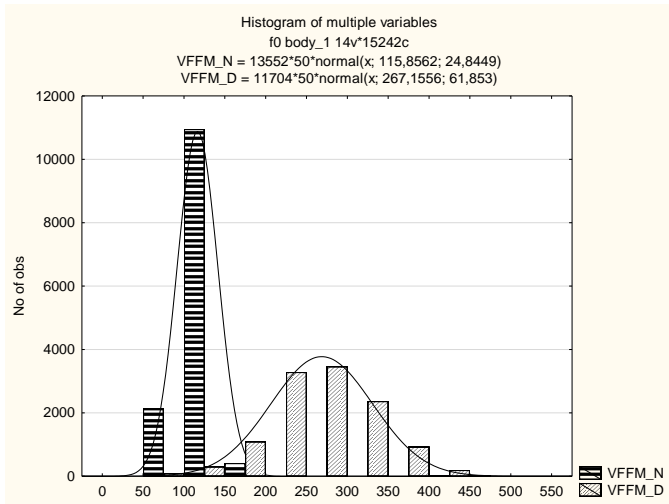
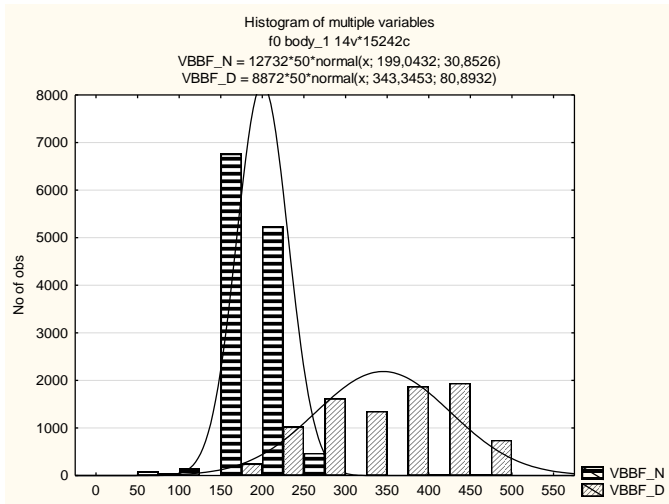


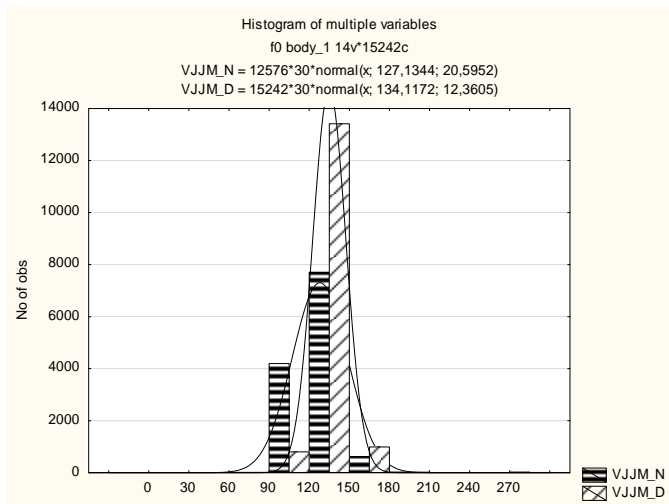
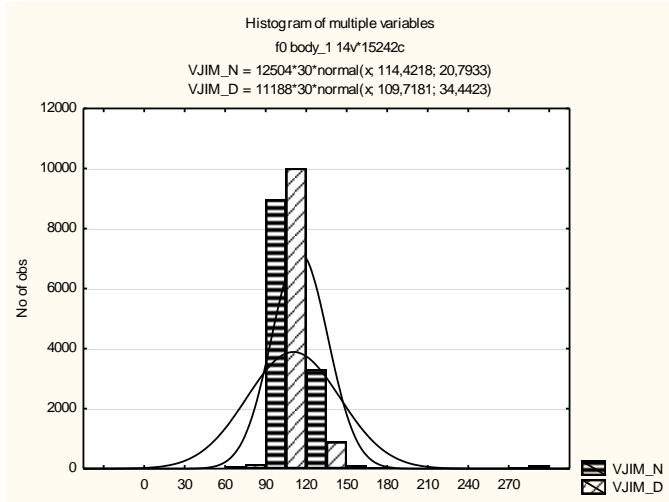
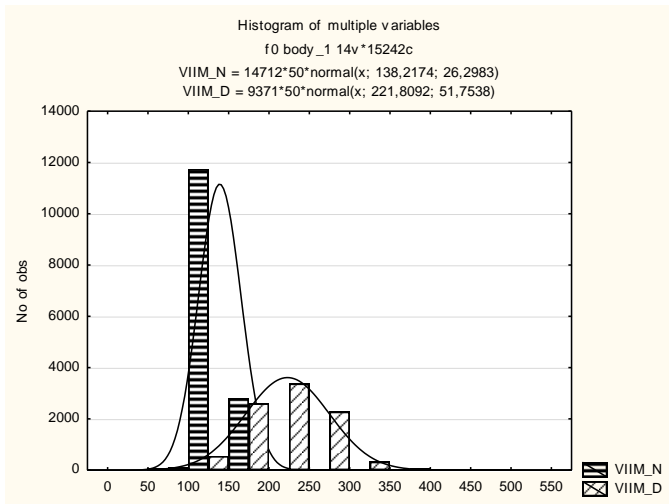


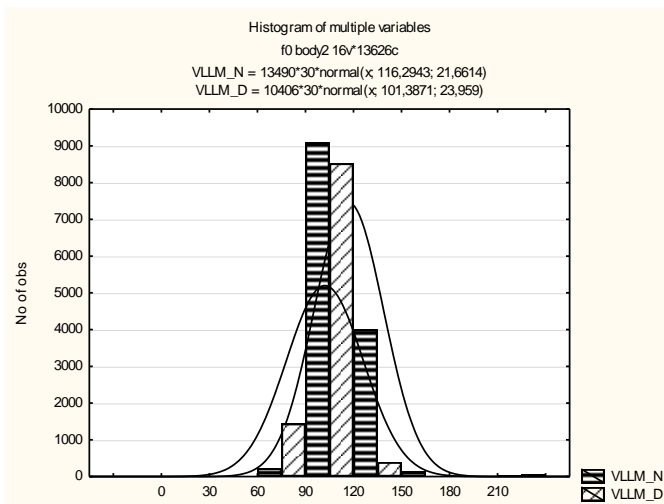
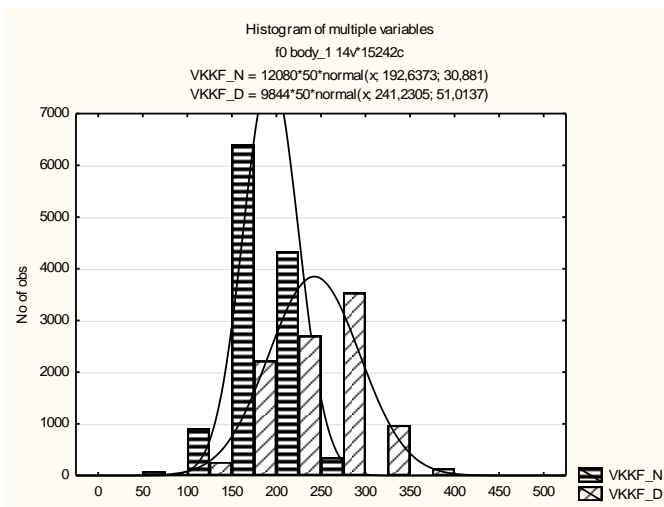
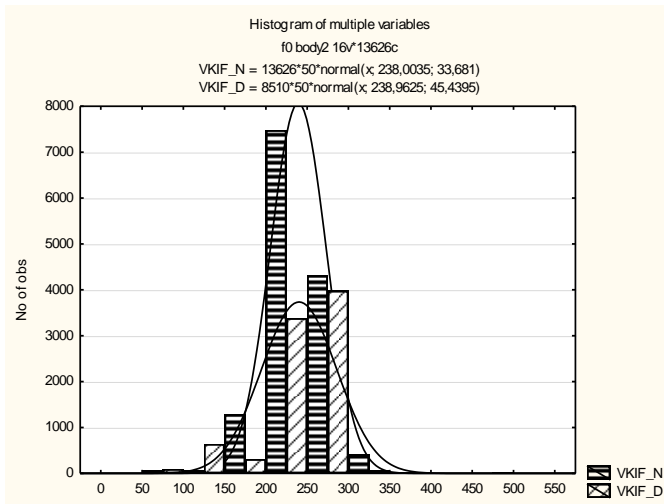


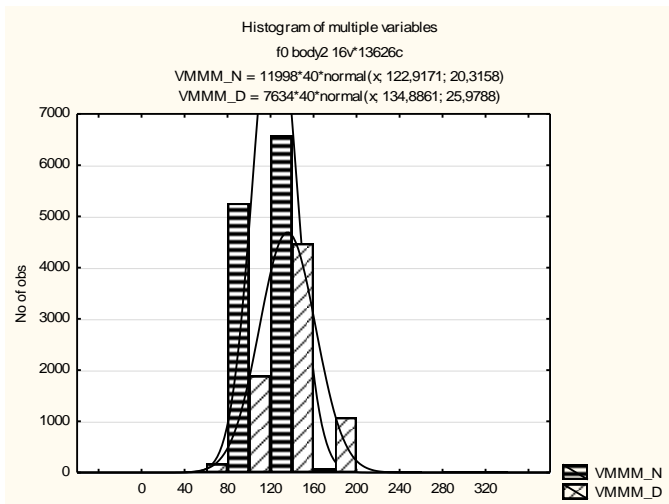
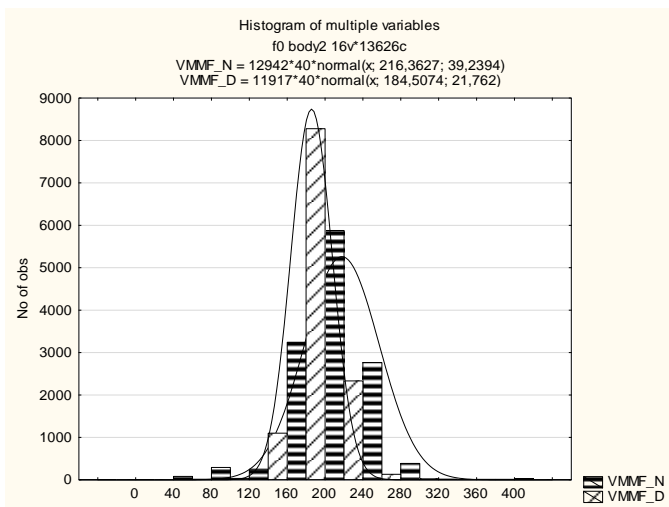
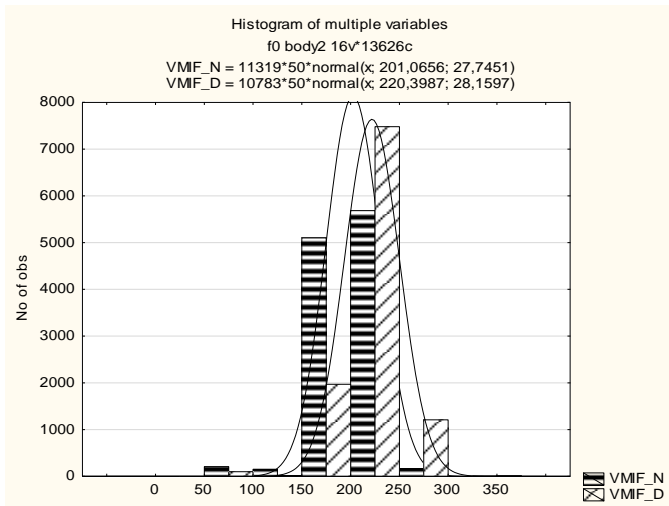


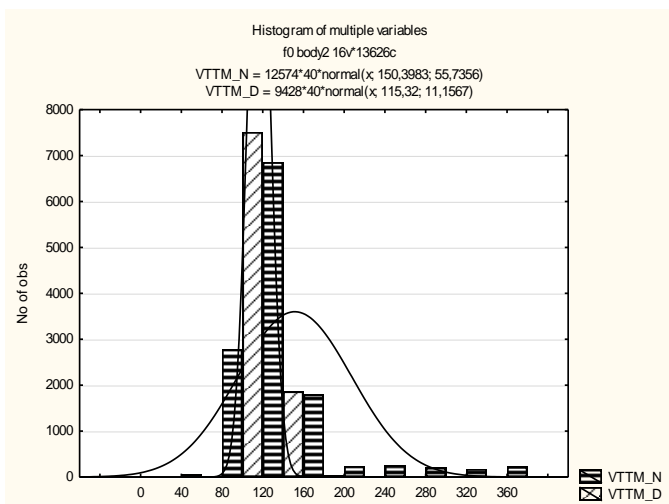
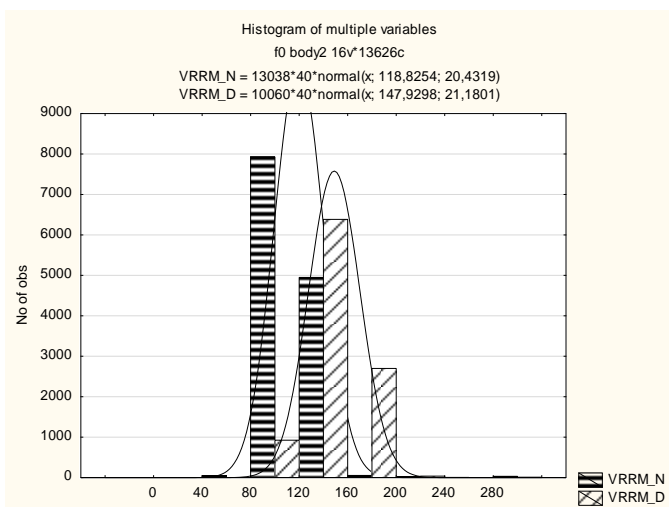
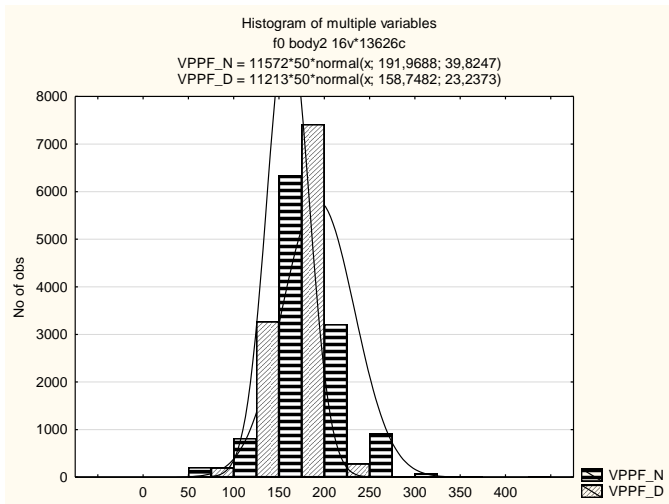
Nelingvisti











Mluvčí bez maturity

