

UNIVERZITA PALACKÉHO V OLOMOUCI
FILOZOFICKÁ FAKULTA

Katedra bohemistiky

KATEŘINA BRAUNOVÁ

česká filologie – anglická filologie

CHARACTERIZING VOWELS. AKUSTICKÁ ANALÝZA
VOKÁLŮ DLE PETERA LADEFOGEDA
(PŘEKLAD S KOMENTÁŘEM)

Characterizing vowels. Acoustic Analysis of Vowels by Peter Ladefoged
(Translation with Comentary)

Bakalářská práce

Vedoucí práce: PhDr. Petr Pořízka
Olomouc 2008

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, pouze s pomocí pramenů a literatury uvedených v bibliografii práce. Souhlasím s tím, aby tato práce byla využita ke studijním a pedagogickým účelům na Univerzitě Palackého v Olomouci.

V Olomouci 11. dubna 2008

Kateřina Braunová

Děkuji PhDr. Petru Pořízkovi za odborné vedení bakalářské diplomové práce, Mgr. Barboře Hedbávné za poskytnutí její diplomové práce, Mgr. Radku, Skarnitzlovi, Ph.D., Mgr. Václavu Jonášovi Podlipskému a Lukáši Crhákovi za ochotu a za pomoc při překladu textu.

1. ÚVOD.....	4
1. 1. PŘEDMĚT A CÍL PRÁCE.....	4
2. REFLEXE ČESKÉ ODBORNÉ LITERATURY.....	6
2. 1 HÁLA	6
2.2 ROMPORTL.....	7
2.3 MLUVNICE ČEŠTINY.....	10
2.4 PALKOVÁ	11
2.5 HEDBÁVNÁ	13
3. TEORIE A ZÁSADY PŘEKLADU.....	15
3.1 SPECIFIKA VĚDECKÉHO A ODBORNÉHO STYLU A PROBLÉMY SPOJENÉ S PŘEKLADEM	17
4. PŘEKLAD S KOMENTÁŘEM	21
<i>Charakteristika samohlásek.....</i>	<i>21</i>
5.1 <i>Spektrogramy.....</i>	<i>21</i>
5.2 <i>Spektra.....</i>	<i>30</i>
5.3 <i>Diagramy samohlásek.....</i>	<i>38</i>
5.4 <i>Nazalizované samohlásky.....</i>	<i>45</i>
5. ZÁVĚR	48
ANOTACE.....	50
6. POUŽITÉ PRAMENY	51

1. Úvod

Akustická fonetika se zabývá studiem výsledného signálu lidské řeči. Pomocí objektivních fyzikálních metod a různých experimentů a experimentálních měření jej popisuje. V poslední době dochází k širokému rozvoji tohoto odvětví fonetiky zejména díky stále pokročilejší výpočetní technice.

Výsledky získané akustickou analýzou dat mají široké použití, přispívají k lepšímu pochopení lidské řeči a popisu její zvukové struktury, využívá se jich též k syntéze řeči a ke komunikaci mezi člověkem a počítačem. Tím však možnosti jejich využití nekončí, jde pouze o ilustrativní výčet.

Akustická fonetika používá specifickou odbornou terminologii, která má leckdy blízko k technickým, resp. přírodovědným oborům. Je to způsobeno tím, že sama používá k analýze řeči fyzikální metody, a dochází tudíž nevyhnutelně k přejímání některých odborných technických výrazů. Tato terminologie však bohužel není v české lingvistické literatuře příliš zažitá, dosud není ustálená a některé termíny se teprve zavádějí. Lingvisté se při popisu řeči opírají tradičně spíše o terminologii artikulační fonetiky.¹

1. 1. Předmět a cíl práce

Cílem této práce je překlad anglicky psaného odborného textu, kapitoly *Characterizing Vowels* z původní monografie *Phonetic Data Analysis* od Petera Ladefogeda (2003) a opatřit text odborným komentářem, v němž jsou popsány problémy terminologické a translatické povahy. V poznámkovém aparátu jsou též vysvětleny pojmy, které jsou českému čtenáři málo známé. Buď proto, že se uvedené jevy v češtině nevyskytují, např. „velarizované“² nebo „retroflexní hlásky“,³ nebo proto, že se dané termíny v české odborné terminologii běžně nepoužívají, např. „FFT spektrum“⁴ nebo „trajektorie formantů“.⁵

¹ Cit.: Krčmová, Marie: Kapitola 6. Akustická a auditivní fonetika. In. *Fonetika*. [online]. 2007. [cit. 2008-04-1]. Dostupné na <<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch06.html>>

² Velarizovaná přední samohláska je tvořená posunutím hřbetu jazyka k měkkému patru (Řeřicha, 1998).

³ Retroflexní modifikované samohlásky se vyskytují v americké angličtině a v jihozápadních anglických dialektech. Jde o samohlásky, po kterých následuje *r*. Jsou jím ovlivněné a téměř po celou dobu svého trvání mají retroflexní kvalitu.

⁴ FFT spektra jsou spektra vytvořená pomocí rychlé Fourierovy transformace. Počítá se korelace sinusových a kosinusových složek se složenou vlnou. Vysoká korelace znamená, že složená vlna a daná sinusovka stoupají a klesají společně. Cit.: Skarnitzl, Radek: *Analýza řečového signálu* [pdf online]. [cit. 2008-03-16]. Dostupné na <<http://fu.ff.cuni.cz/vyuka/akustika/analyza.pdf>>

⁵ Trajektorie formantů jsou grafická znázornění formantů (rezonančních frekvencí) na spektrogramech.

Úvodní část práce se věnuje teorii překladu a problémům spojeným s překladem odborného textu. Tyto poznatky jsem použila při překládání zmíněného textu P. Ladefogeda *Characterizing Vowels*. Snahou byl pokud možno věrný překlad, avšak některé pasáže jsem si dovolila poněkud upravit do podoby odborného, vědeckého textu. Původní monografie má spíše populárně naučný charakter. Autor se v ní snaží čtenáři objasnit a přiblížit danou problematiku srozumitelným jazykem, a proto se nevyhýbá používání poněkud hovorových výrazů a spojení.

Popisované jevy a nastavení počítačových, programů jsou v původní monografii doloženy různými grafy, spektrogramy a spektry,⁶ jež jsou nedílnou součástí původního textu a bez nichž by bylo obtížné sledovat vykládanou problematiku. Tyto obrázky jsem naskenovala a vložila do překladu pokud možno na též místa jako v původní monografii. Ke každému obrázku náleží též odpovídající popis.

Původní text je proložen šesti texty v rámečcích, ve kterých prof. Ladefoged vzpomíná na různá setkání se studenty a na problémy spojené se získáváním a nahráváním dat. Tyto vzpomínkové pasáže jsem nezahrnula do svého překladu, protože jsou ryze osobního charakteru, a tudíž nejsou nezbytné pro samotný výklad dané problematiky a nijak jej z odborného hlediska neochuzují.

Součástí práce je kapitola věnující se reflexi akustického zkoumání vokálů v české odborné literatuře. Je v ní věnována pozornost popisu metod, které se k analýze používají, a popisu jevů v české odborné fonetické literatuře.

⁶ Spektrogram je trojdimenzionální grafické znázornění zvuku, nese informaci o změnách v trvání, frekvenci a intenzitě zvukových vln podél časových os. Spektrum je na rozdíl od spektrogramu dvojdimenzionální znázornění zvuku v určitém okamžiku. Na horizontální ose ukazuje frekvenci a na ose vertikální relativní amplitudy frekvenčních komponentů zvukové vlny.

2. Reflexe české odborné literatury

V následující kapitole se budu věnovat přehledu české odborné literatury, která se věnuje akustickému zkoumání vokálů. Zaměřím se též na popis metod používaných k analýze a popisu jevů. V chronologickém pořadí, v pořadí, ve kterém vyšly, zhodnotím následující tituly: *Akustická podstata samohlásek* od B. Hály (1941), *Základy fonetiky* od M. Romportla (1973), *Mluvnice češtiny I* (1986), *Fonetika a fonologie* od Z. Palkové (1994) a *Diferenciace kvality vokálů v projevu českých mluvčích* od B. Hedbávné (2001).

2.1 Hála

Bohuslav Hála byl jedním z prvních českých fonetiků, kteří se soustavně zabývali akustickým popisem samohlásek. V roce 1941 vyšla jeho kniha *Akustická podstata samohlásek*. Věnuje se v ní akustickému rozboru českých samohlásek, jejich akustickým vlastnostem a akustickému složení. Metody, které se používaly pro akustický popis samohlásek, dělil na metody analytické (subjektivní a objektivní), syntetické a nepřímé.

Mezi subjektivní metody patří přímý poslech nebo analýza samohlásek pomocí různých přístrojů jako jsou rezonátory, ladičky a zařízení pro interferenci.

Rezonátory spočívají na principu spoluznění; rezonátor naladěný na tón znějící před jeho ústím se rovněž rozezvučí, čímž zesílí původní tón. Existují tři typy: velký (pro vyhledávání tónů v oblasti 40 – 6000 Hz), střední (pro vyhledávání tónů v oblasti 200 – 7000 Hz) a malý (pro vyhledávání tónů v oblasti 1600 – 7000 Hz). Pro fonetickou práci je nejvhodnější střední rezonátor.

Ladičky lze používat jako pomůcky pro určování tónů samohlásek při šepotu. Lze je též použít pro tónovou analýzu samohlásek a pro zjišťování vlastních tónů.

U objektivních metod uvádí, jakým způsobem lze získat křivky pro analýzu; tj. pomocí gramofonických, optických, radiofonických zápisů nebo pomocí zápisů oscilografických, fonografických a pomocí zápisů na zvukové filmy. Podstatou oscilografického zápisu je registrování přeměny vzduchových kmitů pomocí mikrofonu na kmity elektrické.

Takto získaná spektra lze analyzovat pomocí matematického rozboru (tzv. Fourierovy řady nebo barycentrický rozbor, který se používá pro zjištění neharmonických složek samohlásky). Křivky bylo možno analyzovat též přímo pomocí různých přístrojů, např. fonoskopů, manometrických Koenigových plamenů nebo elektroautomatických analyzátorů. Výhodou

těchto metod je, že odstraňují počítání harmonické analýzy, protože zvuk je analyzován přímo.

Koenigův analyzátor s manometrickými plameny a 14 rezonátory podává optický obraz analyzovaného zvuku. Fonoskop a strobilion podávají také optický obraz zvuku. Strobilion byl původně zamýšlen pouze jako optická pomůcka při zpěvu, lze ho však použít i pro přímou analýzu samohlásek.

Samohlásky lze též zkoumat pomocí syntetických metod, tzn. umělým vytvořením samohlásek pomocí ladiček a sirén.

Nepřímé metody pro analýzu samohlásek využívají poznatky z oboru telefonie, zkušenosti z foniatrie a také pokusy se změnou rychlosti reprodukcí přístrojů.

Hála ve své práci používal pro zjišťování akustické struktury samohlásek subjektivní metody: rezonátory a ladičky; objektivní metody s registrací zvuku: zápis na fonografu, na zvukovém filmu a oscilografický zápis a metodu přímo analyzující: strobilion. Samohlásky zkoumal též nepřímo pomocí změny rychlosti gramofonové reprodukce.

Cílem práce bylo po všech stránkách popsat akustické složení českých samohlásek. Nejdříve jsou popsány akustické vlastnosti samohlásek a poté se Hála ve své práci věnuje akustickému složení českých samohlásek. Samohlásky jsou charakterizovány svou samohláskovou barvou (vokální timbre) a hlasovou barvou (vokální timbre), které jsou navzájem nezávislé. Hála považuje za důležitější samohláskovou barvu, která je dána několika silnými tóny přibližně stálé výšky. Samohlásku chápe jako periodické opakování souboru resonancí, které mají poměrně stálou výšku. Samohlásky se liší počtem, výškou a místem těchto resonancí.

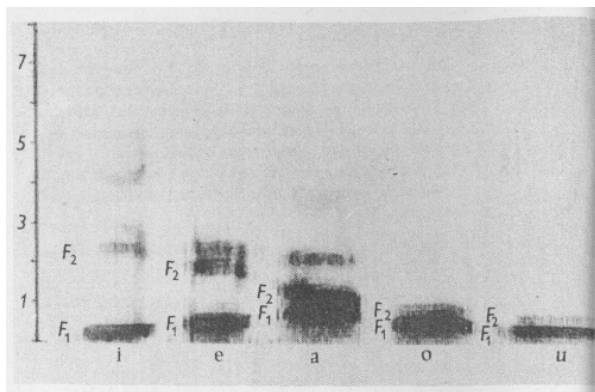
Jednotlivé samohlásky jsou popsány pomocí charakteristiky hlavního formantu (resonance dutiny ústní), vedlejší resonance (dutina hrdelní), resonance střední, resonance horní a vysoké resonance z oblasti 4. oktávy. Uvádí celkový rozsah všech tónů, hlavní oblasti zesílení, průměrné charakteristické oblasti.

2.2 Romportl

V sedmdesátých letech (1973) vyšly *Základy fonetiky* od M. Romportla. Jedná se o vysokoškolská skripta určená nejen pro posluchače Karlovy univerzity v Praze. Jak vyplývá již z názvu, jde o úvod do problematiky fonetiky češtiny, autor se v ní mimo jiné věnuje obecné akustické charakteristice hlásek a popisu českých samohlásek.

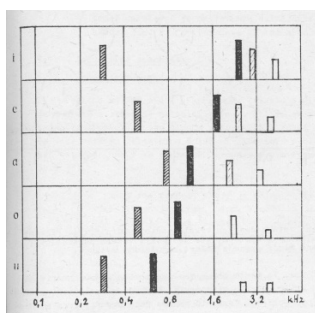
V páté kapitole *Samohlásky* se nejdříve věnuje obecnému popisu samohlásek. Jejich nejdůležitější částí z hlediska akustiky jsou tónové složky. Ty jsou zesilovány nebo zeslabovány při průchodu soustavou rezonátorů (hrdelní, ústní a nosní dutina). Součástí spektra každé samohlásky je základní frekvence, kterou Romportl nazývá základní hlasivkový tón (F_0), a řada vyšších zesílených tónů, kterým se ve fonetické terminologii říká formanty. Nejvyšší důležitost je přikládána F_1 a F_2 . Hodnoty těchto formantů závisí na tvaru dutin (rezonátorů), změny jsou způsobeny pohybem jazyka, rtů a čelisti. Zmíněna je též důležitost dutiny nosní a její vliv na formantové složení samohlásky, při tvorbě českých samohlásek však nehraje důležitou roli.

Při samotném popisu akustické stavby vokálů se použitím různých metod (jež jsou zmíněny v závěrečné kapitole knihy) zjišťují složky charakteristické pro danou samohlásku. Romportl akusticky popisuje samohlásky s pomocí sonagramu,⁷ akustického spektra a grafu F_1 - F_2 .



Sonagram českých vokálů (Romportl, 1973, s. 52).

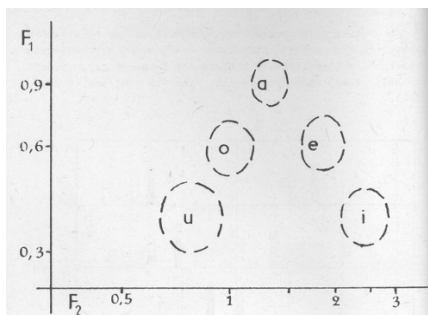
Na sonagramech českých samohlásek prezentuje rozmístění formantů a jejich základní frekvence.



Spektra českých vokálů (Romportl, 1973, s.53).

⁷ Sonagram je starší název pro spektrogram, dnes se tento název již prakticky nepoužívá.

Představuje rovněž akustická spektra českých vokálů jednoho mluvčího, hodnoty jsou převedeny do grafu, který vyjadřuje vztah prvního a druhého formantu.



Vztah českých vokálů podle jejich poměru F1-F2 (Romportl, 1973, s. 54).

Elipsy na obrázku ohraničují oblast nejhojnějšího výskytu. V místech, kde se protínají hodnoty F1 a F2 je uvedeno písmeno odpovídající hlásce. Při obráceném narůstání frekvencí by graf lépe odpovídal zobrazení hlásek na samohláskovém trojúhelníku.

Jednotlivé samohlásky jsou charakterizovány pomocí vzdálenosti formantů a pásma frekvenčních hodnot. Nejsou rozlišovány krátké a dlouhé samohlásky. U některých samohlásek jsou dále uvedeny rysy, v nichž se odlišují.

V závěru publikace jsou uvedeny metody pro výzkum akustické stránky řeči. Romportl metody dělí podle toho, k jakým účelům slouží: registrace řeči, analýza akustických vlastností řeči, akustická syntéza řeči a metody výzkumu percepce řeči. K akustické analýze zvuku se používalo a stále používá několik různých metod, některé z nich jsou univerzální a některé z nich se používají pouze k rozboru určitých vlastností zvuku. Uvedu zde pouze metody pro analýzu zvuku, které se používaly nebo zaváděly v době vzniku této knihy.

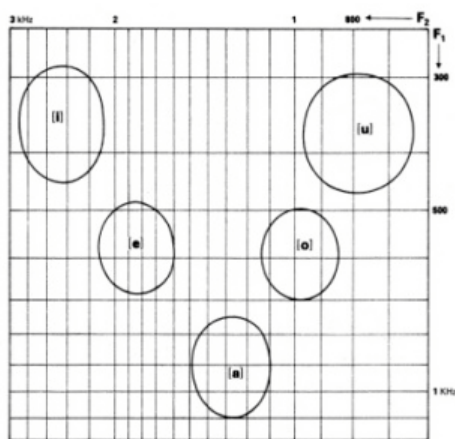
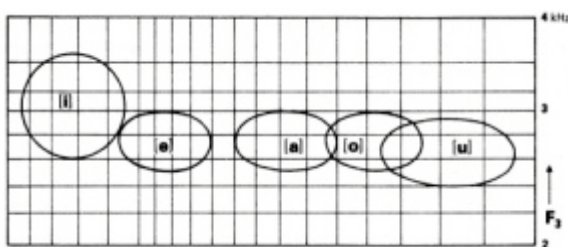
Pro analýzu výšky tónu se používal zapisovač melodie. K zobrazení a měření intenzity se používal zvukoměr. Ke zjištění délky není nutné používat speciální metody, mnohdy je však problém určit hranice zvuku. Největší problém představuje analýza tónu. Pomocí Fourierovy analýzy lze z oscilografického záznamu vypočítat svrchní tóny (formanty). Dříve to znamenalo použít různé poměrně složité výpočty. Technika se však postupně vyvíjela a dnes za nás v podstatě provedou veškeré výpočty počítače. Spektrogram podával výsledky v podobě akustického spektra, zachycoval frekvenci a intenzitu. Obsahoval buď filtry o různé frekvenci, nebo jeden s měnitelnou frekvencí. Pomocí sonagrafu (zvukového spektrogramu) se daly zachytit tři dimenze, čas, frekvence a intenzita, což v té době znamenalo značný pokrok. Na rozdíl od ostatních zařízení zachycoval změnu spektra v čase. Pro zachycení pohybu formantů se používaly filtry o šířce pásma 300 Hz a pro určení frekvencí jednotlivých složek se používal úzký filtr s šířkou pásma 45 Hz.

2.3 Mluvnice češtiny

Mluvnice češtiny je chápána jako stěžejní, reprezentativní dílo české lingvistiky. Vyšla v osmdesátých letech uplynulého století, vznikala však už během let sedmdesátých. V podstatě navazuje na tradiční synchronní popis jazyka, věnuje se zevrubnému popisu českého jazyka a jeho struktury.

V prvním díle *Mluvnice češtiny* je první část věnována fonetice češtiny, její součástí je i kapitola o samohláskách. Ty jsou popisovány jak z hlediska artikulace, tak z hlediska akustické struktury. Nejdříve je obecně popsána jejich struktura. Je zde uvedeno z čeho se samohlásky skládají, jak vznikají formanty a podle čeho je lze rozpoznat na spektrogramech. Je zde také uvedeno, jak se na spektrogramech projevuje změna intenzity a jak se podílí artikulace (pohyb jazyka a rtů) na pohybu a změnách formantů. Jako nejdůležitější jsou z hlediska percepce vnímány první tři formanty.

V části věnované artikulačnímu a akustickému popisu je zdůrazněna souvislost mezi Hellwagovým trojúhelníkem a zobrazením hlásek na grafu znázorňujícím vztah prvních dvou formantů. Je zde také uveden graf poměru druhého a třetího formantu.



Rozptylové pole hodnot prvních tří formantů českých samohlásek (*Mluvnice češtiny* 1, 1986, s. 30).

Jednotlivé samohlásky jsou popsány pomocí jejich artikulace a akusticky jsou charakterizovány minimální vzdáleností prvních dvou formantů. Ze středových hodnot

formantů je určen poměr F1 a F2. Jsou zde také uvedena pásma hodnot frekvencí formantů a jak se liší intenzita obou formantů. Uváděné hodnoty platí jak pro dlouhé tak pro krátké samohlásky. Nejsou zde rozlišovány, protože se údajně liší pouze ve délce trvání.

2.4 Palková

Jednou z modernějších knih věnovaných fonetice je *Fonetika a fonologie češtiny* od Zdeny Palkové, která vyšla v roce 1994 (ve druhém a opraveném vydání r. 1997). V kapitole nazvané *Popis češtiny: samohlásky* se věnuje akustickému popisu českých samohlásek.

Nejdůležitějšími složkami, které určují kvalitu českých samohlásek, jsou podle ní F1 a F2. V několika tabulkách uvádí pro porovnání hodnoty těchto formantů, nejde o její vlastní měření, data jsou převzata ze starší literatury: B. Hála (1941), M. Romportl (1963), akademická Mluvnice češtiny I (1986).

Hodnoty uvedené v tabulkách jsou lehce odlišné. Pouze B. Hála odlišuje krátké a dlouhé samohlásky, uvádí pásma prvních třech formantů a jejich středy. Romportl (spodní tabulka vlevo) ani MČ 1 (spodní tabulka vpravo) nerozlišují krátké a dlouhé samohlásky. Hodnoty F3 uvedené v tabulce podle MČ 1 jsou pouze odečteny z grafu. V souhrnné tabulce jsou uvedena rozložení pásem ve všech variantách popisu.

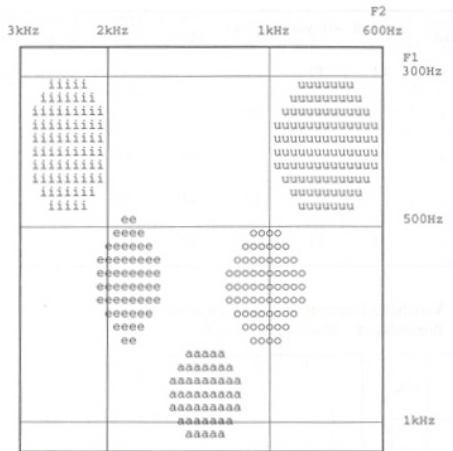
Formant hláska	F1 pásma	Střed pásma	F2 pásma	Střed pásma	F3 pásma	Střed pásma
i	325-390	355	2016-2202	2120	2265-2494	2365
í	294-350	326	2132-2349	2230	2448-2780	2610
e	495-576	572	1566-1742	1660	1952-2190	2070
é	480-575	510	1700-1835	1750	2003-2240	2120
a	699-800	750	1180-1305	1280	1595-1905	1776
á	710-858	795	1105-1266	1175	1360-1740	1577
o	542-603	580	900-1139	982	1463-1617	1617
ó	504-556	530	850-950	900	1632-1960	1792
u	347-408	385	744-918	758	1905-2300	2100
ú	314-397	350	640-720	680	2016-2445	2230

Formant Hláska	F1 pásma	F2 pásma
i, í	300-500	2100-2700
e, é	500-700	1600-2100
a, á	800-1000	1200-1400
o, ó	500-700	900-1200
u, ú	300-500	600-1000

Formanty hlásky	F1 pásma	F2 pásma	F3 pásma
i, í	300-450	2100-2800	2600-3500
e, é	500-700	1600-2100	2500-3000
a, á	800-1100	1100-1500	2500-3000
o, ó	500-700	900-1200	2500-3000
u, ú	300-500	600-1000	2400-2900

Hodnoty formantů (Palková, 1994, s.173-172, dle B. Hála, M. Romportl, akademická Mluvnice češtiny 1.).

Hodnoty uvedené v souhrnné tabulce jsou zobrazeny ve frekvenčním poli formantů. Toto zobrazení odpovídá uspořádání hlásek na Hellwagovu trojúhelníku, který zobrazuje samohlásky podle způsobu jejich artikulace.

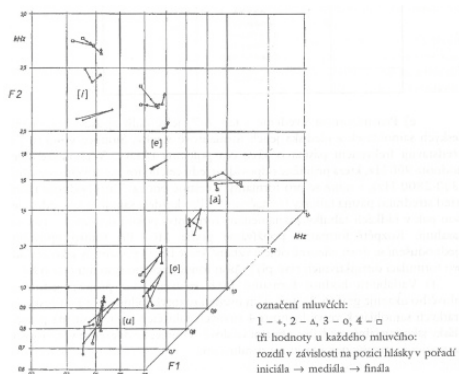


České samohlásky ve frekvenčním poli formantů (Palková, 1994, s. 175).

V poslední tabulce jsou v horní části frekvenční pásma v rozmezí půlky oktávy, každá samohláska má vyznačené formanty v odpovídajících rozmezích. Tento způsob zobrazení přehledně ukazuje, které samohlásky mají frekvenční pásma F1 a F2 blízko sebe.

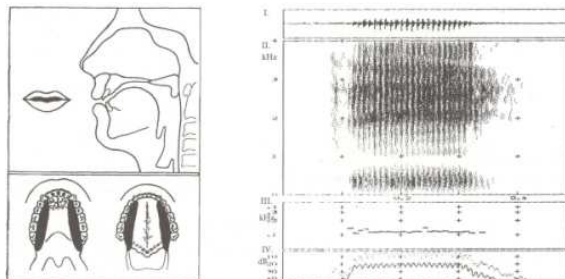
pásmo hláska	260-370	370-550	550-750	750-1100 *	1100-1500	1500-2100	2100-3000
i	F1	F1					F2
e		F1	F1			F2	
a				F1	F2		
o			F1	F1	F2		
u	F1	F1	F2	F2			

Rozložení formantů ve frekvenčních pásmech o šířce půl oktávy (Palková, 1994, s. 176).



Variabilita formantů u čtyřech mluvčích (Palková, 1994, s. 176, dle Borovičková – Maláč).

Graf přejatý z Borovičková – Maláč (1967) poukazuje na variabilitu mluvcích a na variabilitu v projevu jednoho mluvčího. Jsou na něm zobrazeny hodnoty čtyř mluvčích, kteří vyslovovali samohlásky v iniciální, mediální a finální pozici.



Samohláska [i] [i:] (Palková, 1994, s. 181).

Jednotlivé samohlásky Palková charakterizuje pomocí popisu jejich artikulace. Pro lepší pochopení u nich uvádí skiagramy,⁸ tvary retní štěrbiny, palatogramy⁹ a lingvogramy.¹⁰ Co se týká jejich akustických vlastností, zobrazuje pouze záznamy zvuku, spektrogramy a průběhy základního tónu a dynamiky.

2.5 Hedbávná

Hedbávná se věnuje akustické charakteristice samohlásek ve své diplomové a dosud nepublikované práci *Diferenciace kvality vokálů v projevu českých mluvčích* (2001). Cílem této práce bylo získat údaje a porovnat stabilitu a rozsah variability v projevu jednoho mluvčího a skupiny mluvčích. Bylo porovnáváno osm mluvčích, materiály pro rozbor byly pořízeny z rozhlasových nahrávek. V odlišných pozicích, v iniciální, mediální a ve finální, byly analyzovány jak dlouhé, tak krátké samohlásky. Naměřené hodnoty F1 a F2 byly porovnávány s hodnotami ze starších fonetických prací. Dále se zjišťovalo, které samohlásky jsou stabilní a které mají tendenci se lišit v naměřených hodnotách F1 a F2, jak mluvčí diferencuje jednotlivé samohláskové typy a jaký je vliv trvání na formantové složení

⁸ „Skiagram“ znamená rentgenový snímek tělesných orgánů.

⁹ Palatogram je fotografie paterní klenby nebo její kresba s obrazem dotyku.

Cit.: Krčmová, Marie: 5.5 Metody poznávání artikulace. In. *Fonetika*. [online]. 2007. [cit. 2008-04-1]. Dostupné na <<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch05s05.html>>

¹⁰ Lingvogram je fotografie jazyka v ústech nebo jeho kresba ukazující, kterou částí se artikulace provádí.

Cit.: Krčmová, Marie: 5.5 Metody poznávání artikulace. In. *Fonetika*. [online]. 2007. [cit. 2008-04-1]. Dostupné na <<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch05s05.html>>

samohlásek. Byly vyhodnoceny rozdíly a shody mezi mluvčími a rozmístění center samohláskových typů.

V úvodní části práce autorka popisuje artikulační a akustickou podstatu samohlásek. Jde v podstatě o souhrn dosavadního – jak českého, tak světového – výzkumu, jsou zde porovnávány jednotlivé přístupy a jejich výhody i nevýhody. Dále je zhodnocena důležitost jednotlivých formantů pro rozlišení a percepci hlásek, je zde také popsána variabilita formantových frekvencí a to, čím je způsobena. V souhrnu je uvedena základní charakteristika jednotlivých samohlásek češtiny a příčiny její variability.

V kapitole *Výběr a analýza materiálu* jsou představena kritéria pro výběr mluvčích a materiálu, možné postupy akustické analýzy a metody, které byly v práci použity.

Obecné metody, které zmiňuje Hedbávná zde nebudu uvádět, to je předmětem následující kapitoly. Představím zde pouze metody, které byly použity pro analýzu samohlásek v této práci.

Účelem této práce, bylo zjistit hodnoty frekvencí ve středu samohlásky a změny v průběhu formantů nebyly důležité. (Změny lze nejlépe pozorovat na spektrogramech). Pro tento účel se autorka rozhodla použít FFT spektra z programu Cool Edit verzí 96 a 2000. Hodnoty byly získány pomocí měření kurzorem myši přímo na spektru samohlásky. Někdy bylo problémem správně určit stabilní část samohlásky, v těchto případech pomohla vlna oscilogramu a sluchová kontrola. Program, který autorka používala nabízel výběr šesti typů okna FFT, po uvážení jejich výhod a nevýhod se rozhodla použít typ *Blackmann-Harris*. Při zobrazení FFT spekter bylo na ose x použito logaritmické měřítko, které více odpovídá sluchovému vjemu. Výhodou FFT spekter je, že ukazují jednotlivé harmonické, můžeme sami uvážit, kde je střed formantového pásma, problém však může nastat, jestliže jsou harmonické přibližně stejné amplitudy blízko sebe nebo když jde o shluk více harmonických.

3. Teorie a zásady překladu

Cílem překladu je převést text z výchozího jazyka do jazyka cílového, dochází tedy k překódování z jednoho jazyka do druhého. Za základní přístup k překladu se považuje tzv. funkční ekvivalence. Při překladu nelze používat tytéž prostředky, což vzhledem k odlišnému charakteru obou jazyků ani není možné, ale měly by plnit stejnou funkci, a to jak denotační,¹¹ tak konotační¹² a pragmatickou.¹³ Při překladu bychom neměli měnit charakter sdělení, záměr autora a typ adresáta, záleží však na účelu textu a překladu.

Problémem při překladu jakéhokoliv textu je hledání odpovídajících ekvivalentů (překladových protějšků). Setkáváme se zejména s těmito možnostmi: existuje ekvivalent buď úplný (absolutní) nebo částečný, anebo existuje více možných ekvivalentů. Může však nastat i případ, že neexistuje žádný ekvivalent.

O úplných, absolutních, protějšcích můžeme mluvit v případech, kdy si výrazy výchozího a cílového jazyka odpovídají jak po stránce významové, tak stylistické, mají stejné denotační a konotační složky. Patří mezi ně výrazy z centra slovní zásoby týkající se lidské činnosti, mentálních procesů, percepce a modálních postojů. Lze mezi ně řadit i slova s více protějšky, která však mají význam pevně daný kontextem, slovesu „get“ odpovídají „dostat“, „přinést“, „připravit“, „mít“, „pochopit“ atd.). Málo úplných protějšků nalezneme zejména mezi slovesy: je to dáno tím, že česká slovesa obsahují ve srovnání s anglickými více informací a čeština má též kategorii vidu. Čeština je v podstatě verbální jazyk, zatímco angličtina nominální.¹⁴ Anglické sloveso plní ve větě funkci spojovacího výrazu nominálních frází, které nesou význam. Při překladu bychom na to neměli zapomínat, a měli bychom se doslovným překladům těchto frází vyhnout. Již zmíněnou verbo-nominální frázi „take a rest“ nelze překládat jako „vzít si oddech“, ale jako „odpočinout si“. U adjektiv nastává úplná ekvivalence, jestliže se jedná o popis objektivních vlastností, jako jsou „long“ (č. dlouhý) nebo „empty“ (č. prázdný) (viz Knittlová, 1995, s. 16). Absolutní ekvivalenty si většinou odpovídají po stránce formální. V případě, že je cizí protějšek víceslovný, tak je i domácí

¹¹ Ekvivalenty by si měly odpovídat po stránce denotační, tzn. že by vztahy jejich formy (znaků) ke třídě objektů měly být totožné (viz Čermák, 2004, s. 231).

¹² Měly by si odpovídat i po stránce konotační, tzn. že jejich sekundární nekognitivní rysy především asociativní povahy by měly být shodné (viz Čermák, 2004, s. 253).

¹³ Ekvivalenty by se měly shodovat též po stránce pragmatické, zkušenostní. Neměly by být změněny funkce znaků a jejich účinek na adresáta by měl zůstat stejný (viz Čermák, 2004, s. 273).

¹⁴ Zatímco angličtina má tendenci se vyjadřovat jmény, např. „take a rest“, čeština používá slovesa se specifickým významem „odpočinout si“. Angličtina je analytický jazyk a čeština jazyk syntetický. Výrazy „nominální“ a „verbální jazyk“ poukazují na tyto tendence.

ekvivalent víceslovný, např. „make notes“ (č. dělat si poznámky). Korespondence se nachází i v motivovanosti, např. „crossroads“ (č. křižovatka) (viz Knittlová, 1995, s. 16).

Ve většině případů však nenajdeme úplné ekvivalenty, ale pouze ekvivalenty částečné. Je to způsobeno také odlišnou typologickou charakteristikou obou jazyků. Rozdíly mezi ekvivalenty mohou být po stránce formální, významově denotační, významově konotační a pragmatické, mohou se však i kombinovat.

Angličtina je jazyk izolační, analytický a dává přednost víceslovným výrazům, oproti tomu čeština je jazyk flexivní, syntetický a upřednostňuje jednoslovná vyjádření. Modifikující člen může být vyjádřen slovotvorným formantem, např. „apple tree“ (č. jabloň) (viz Knittlová, 1995, s. 17). Patří sem i anglicky sémanticky chudá slovesa, která se pojí s nominální nebo předložkovou frází, v češtině jim většinou odpovídají jednoslovná slovesa s předponou, např. „get up“ (č. vstát) nebo „go away“ (č. odejít). Čeština používá analytické výrazy, jestliže neexistuje protějšek k anglickému slovesu, které vzniklo konverzí substantiva, např. „trapped“ (č. lapený do pasti), nebo v případě termínů, u kterých nedošlo k univerbizaci, např. „cog“ (č. ozubené kolečko) (viz Knittlová, 1995, s. 18).

Víceslovná pojmenování bývají obvykle explicitnější. Jestliže se při překladu snažíme vyjádřit stejnou míru explicitnosti, můžeme vyjádřit informace, které jsou v původním výrazu obsaženy implicitně, např. „cistern“ (č. nádržka na vodu) (viz Knittlová, 1995, s. 18).

U částečných ekvivalentů narazíme na problém v rozdílech významové denotace. České protějšky většinou oproti anglickým obsahují navíc významovou složku, např. slovesa pohybu jako jsou „go“, „come“, „arrive“ nevyjadřují na rozdíl od českých ekvivalentů prostředek pohybu. Slovesu „go“ odpovídají české protějšky „jít“ a „jet“ (viz. Knittlová, 1995, s. 21). Méně často dochází k tomu, že je některá složka potlačena a specifický výraz je nahrazen výrazem generalizovaným, např. „store bread“ (č. chleba) (viz Knittlová, 1995, s. 29).

Částečná ekvivalence může být též způsobena konotačními rozdíly. Slova jsou většinou konotačně neutrální, někdy však v kontextu získávají různé stylistické či expresivní konotace, např. slovu „star“ může odpovídat zdobně protějšek „hvězdička“ (viz Knittlová, 1995, s. 32).

Částečné protějšky jsou též způsobeny odlišnými mimojazykovými zkušenostmi mluvčích. Záleží na překladateli, jestli se rozhodne upřesnit a přidat informace, které jsou českému adresátovi neznámé, např. „Atlantic Monthly“ lze přeložit jako „Atlantic Monthly“, nebo vypustit informace, které by mohly být považovány jako redundantní, např. „academic building“ (čes. ekv. škola) (viz Knittlová, 1995, s. 49).

Mnohdy nastane případ, že v cílovém jazyce neexistuje protějšek. Může to být způsobeno tím, že daná skutečnost je uživatelům cílového jazyka neznámá. Daný termín z výchozího jazyka lze buďto přejmout do jazyka cílového, tj. zde jeho počestěním, např. termín „formant tracks“, ke kterému neexistuje protějšek, lze přeložit jako „trajektorie formantu“.

Angličtina je velmi polysémní jazyk, tudíž některá slova mohou mít velké množství odlišných významů. V případě existence více ekvivalentů se musíme rozhodnout podle jazykového kontextu a zvolit tu možnost, která významově odpovídá a je spojitelná s ostatními lexikálními jednotkami. Výraz „subject“ má několik možných ekvivalentů, např. „téma“, „předmět“, „předloha“, „subjekt“ atd. Jestliže se vyskytuje ve větě, která popisuje školní prostředí, odpovídá mu protějšek „předmět“ (viz Knittlová, 1995, s. 52).¹⁵

3.1 Specifika vědeckého a odborného stylu a problémy spojené s překladem

Vědecký styl má odborně sdělnou funkci. Jeho hlavním cílem je přesné, výstižné a úplné sdělení myšlenky. Vysvětluje a objasňuje danou problematiku. Je pro něj charakteristické logické členění textu, ne vždy je sdělovaný obsah zcela zřejmý, tudíž by se měl autor vyvarovat obtížných a těžko srozumitelných konstrukcí. Tento styl je objektivní a vyhýbá se používání expresivních a subjektivních výrazů. Velmi často se opakují některé výrazy a důsledné je užívání odborné terminologie.

Pro vědecký styl je charakteristická neosobnost, osoba autora je potlačena do pozadí. Proto se ve vědeckém stylu velmi často používá pasivum, a to tam, kde je požadavek implikovat neurčitý všeobecný konatele nebo jestliže konatel není podstatný a cílem je zdůraznit zasažený objekt. V angličtině se používá též všeobecný zájmenný podmět „we“, který odkazuje na autora. V češtině se neosobnost vyjadřuje používáním 1. osoby plurálu. Pomocí pasiva lze také rematizovat¹⁶ konatele, a to tím, že je umístěn na konci věty a vyjádřen předložkou „by“ např. „The story was written by John.“ (čes. Tu povídku napsal John.).

Neosobnost se v angličtině vyjadřuje též speciálními konstrukcemi jako „It is easy to do sth.“ (č. Je snadné něco ...) nebo „It is likely that...“ (č. Je pravděpodobné, že...) (viz Knittlová, 1995, s. 83-84). Tyto vazby v podstatě nenesou sémantickou informaci, ale slouží jako spojovníky v textu nebo jako úvodní věty odstavce. Neosobní věty jsou ve vědeckém stylu velmi časté, protože autor uvádí různé hypotézy a domněnky, v podstatě je nereprezentuje

¹⁵ Knittlová (1995).

¹⁶ „Rematizovat“ znamená umístit větný člen do pozice, která odpovídá rématu (jádro) výpovědi, čímž tento člen získá nejvyšší stupeň dynamičnosti v dané výpovědi.

přímo, ale s jistou opatrností. V neosobních větách se často používají modální slovesa, jejich lexikální význam je oslaben a slouží spíše k vyjádření hypotetického děje a oslabení obecné platnosti výroku.

Anglický odborný styl velmi často používá kauzativní vazby typu: „This makes the problem easy.“ Tyto vazby by se do češtiny neměly překládat doslovně, je nutné konstrukci změnit nebo situaci volně opsat, např. čes. „Tímto se problém zjednoduší“ (viz Knittlová, 1995, s. 85). K těmto typům vazeb patří také konstrukce např. s „enable“ (č. umožnit), „allow“ (č. povolit), „permit“ (č. dovolit), které se pojí s infinitivem, tyto vazby se obvykle překládají vedlejšími větami, např. „Safety valves allow the metal to cool slowly.“ (č. Bezpečnostní ventily umožňují, aby kov pomalu chladl.) (viz Knittlová, 1995, s. 86).

Dalším rysem vědeckého stylu je přesná formulace a logický sled vět. Často se používají navazovací formule, které by se při překladu neměly opomíjet. Slouží k členění textu a k navazování vět, např. „however“ (č. avšak) nebo „moreover“ (č. kromě toho) (viz Knittlová, 1995, s. 86).

Logicky uspořádány jsou též větné členy a věty uvnitř souvětí. Věty postupují od tématu, od toho, co je známé, k rématu, nové informaci. Odhalit téma a réma je při překládání velmi důležité. Čeština je flexivní jazyk, a proto má volnější slovosled než angličtina – to znamená, že v češtině není problémem umístit různé větné členy na konec věty. V případě, že je v angličtině potřeba rematizovat podmět, je jednou z možností pasivizace, např. „The letter was written by Peter“. Čeština na rozdíl od angličtiny může umístit podmět na konec věty „Dopis napsal Petr“. Réma bývá v češtině vyjádřeno také lexikálními prostředky nebo speciálními syntaktickými strukturami. Oproti tomu má angličtina pevný slovosled, který má gramatickou a sémantickou funkci. Ke zdůraznění rématu tudíž využívá odlišné prostředky, čehož by si měl být překladatel vědom.

Jednou z možností je postavit réma za přísudek „to be“ (č. být), „to seem“ (č. zdát se), např. „There is a problem.“ (č. Je v tom problém.). Réma se může objevit na konci výpovědi nebo na jejím začátku pomocí předrážky „It is...“ – jejím prostřednictvím může být vytčen kterýkoli větný člen kromě určitého slovesného tvaru, např. „It is this interaction that causes the absorption.“ (č. Právě tato interakce způsobuje absorpci.) (viz Knittlová, 1995, s. 87). Réma lze též signalizovat pomocí pasivních konstrukcí, ať už s „by“, např. „The letter was written by Peter.“ (č. Ten dopis napsal Petr.) nebo bez něj, např. „The deer was killed.“ (č. Toho jelena zabili.). Lexikálně lze poukázat na réma pomocí výrazů jako jsou „only“ (č. pouze), „merely“ (č. prostě) a „just“ (č. právě); např. „Only a few problems in quantum mechanics can be solved exactly.“ (č. Přesně lze v kvantové mechanice řešit pouze několik

problémů.) (viz Knittlová, 1995, s. 87). Réma lze také vyjádřit pomocí neurčitých členů nebo nulových členů (v plurálu u nepočítatelných a u abstrakt). „There was a film going on in the large room.“ (č. Ve velké místnosti běžel nějaký film.). Na členy si však musíme při překladu dát pozor, protože ne každý neurčitý člen signalizuje réma.

Logické uspořádání platí též pro strukturu souvětí – uspořádání vět v textu. Souvětí jsou v odborném textu nejčastěji ve vztahu podmínkovém, objevuje se též vztah časový, důsledkový a důvodový. Pro odborný styl je primární účelnost, překladatel tudíž může po uvážení rozbít větné celky a přemístit věty, obsah však musí zůstat nezměněný.

Anglický odborný styl je charakteristický svou hutností, je velmi kondenzovaný. Často se používají participiální a gerundiální vazby a infinitivní konstrukce, např. „To obtain tolerance levels by applying Eq. (13) to certain specific aberration types, we must determine C1 and C2 so that the left-hand side is minimum, this being equivalent to taking the brightest part of the spectrum line at the best plane of focus“. Do češtiny se tyto konstrukce překládají pomocí vedlejších vět: (č. Abychom zjistili toleranční hranice tím, že budeme aplikovat rov. (13) na jisté typické druhy aberací, musíme určit C1 a C2 tak, aby levá strana nabyla minimální hodnoty, což je ekvivalentní tomu, že nejjasnější část spektrální čáry bereme v podmínkách nejlepšího zaostření.) (viz Knittlová, 1995, s. 89). Avšak i čeština používá kondenzovaná vyjádření, konstrukce s podstatným jménem slovesným a s infinitivem a zpřídavnělých přechodníků a přičestí, např. „V tom případě nelze použít aproximací používaných k vyhodnocování rozdělení iradiance při malých chybách fáze.“ (viz Knittlová, 1995, s. 89).

Pro angličtinu je charakteristické používání rozvitých determinačních skupin, což může leckdy znamenat pro překladatele problém. Většinou se při dekódování těchto skupin postupuje zprava doleva, tento vztah však může být složitý a různorodý. Takováto vyjádření jsou často nejednoznačná, čeština dává přednost explicitnějšímu vyjadřování, např. slovní spojení „steam corrosion inhibition“ lze překládat jako „zamezení koroze způsobené parou“ (viz Knittlová, 1995, s. 90).

S odborným stylem souvisí potlačování emocionality a expresivity. Termíny jsou jednoznačné, mají úzký význam a omezené pole použití. Vědecká terminologie směřuje k internacionalizaci, nejnáze se převádějí termíny z latiny a řečtiny. Překladatel si však při převádění internacionalismů musí dát pozor, protože např. „some techniques“ neznámá „techniky“, ale „metody“ (viz Knittlová, 1995, s. 92).

Pro češtinu představuje problém zejména překlad termínů, které jsou ve formě již zmiňovaných substantivních řad. Výsledkem bývají popisné překlady, které jsou však velmi dlouhé. Při snaze o doslovný překlad může vzniknout ekvivalent, který se příliš nehodí pro

běžné používání: např. „wing-sweep system“ (č. soustava ovládní úhlu šípu křídla) (viz Knittlová, 1995, s. 93). V ostatních jazycích je tendence k ekonomičnosti a zkracování takovýchto výrazů, v češtině však tato tendence není příliš znatelná. V současné době se ve vědecké terminologii v souvislosti se snahou o ekonomičnost objevují také svérázné metafory, čeština však takovýmto výrazům není příliš přístupná.

V technických oborech panuje mezi termíny většinou těsná korespondence, někdy však překladatel musí termín vyjádřit opisem, který je explicitnější, např. „jet pilot“ lze přeložit jako „pilot tryskové stíhačky“ (viz Knittlová, 1995, s. 18). Některé termíny nejsou příliš ustálené nebo se případně v jazyce, do jehož se překládá, nepoužívají. Musíme tudíž danou skutečnost popsat. Mnohdy vznikají i řady ekvivalentních termínů, které lze chápat jako synonyma, např. pro termín „creaky voice“ lze použít buď ekvivalentu „stlačená fonace“ nebo „třepeň fonace“ záleží pouze na překladateli, který z nich upřednostní.¹⁷

¹⁷ Knittlová (1995).

4. Překlad s komentářem

CHARAKTERISTIKA SAMOHLÁSEK

Nejlepší způsob jak charakterizovat samohlásky nespočívá v popisování jejich artikulace, ale na základě popisu jejich akustických vlastností. Úvodní popisy akustiky řeči se nacházejí v několika textech věnovaných obecné fonetice, stejně tak jako ve specializovaných knihách, některé z nich jsou zmíněny na konci předchozí kapitoly. Tyto knihy si všímají, že nejdůležitější akustickou vlastností samohlásek jsou jejich formanty,¹⁸ které lze snadno rozpoznat na spektrogramech. Učebnice vám však většinou neřeknou, jak vytvářet dobré spektrogramy a jak analyzovat a prezentovat údaje o akustické struktuře samohlásek. To je předmětem této kapitoly.

5.1 SPEKTROGRAMY

Řekněme, že chcete popsat samohlásky daného jazyka a nahráli jste podle první kapitoly sady slov. Nejdříve uděláte spektrogramy¹⁹ tohoto materiálu, abyste získali celkovou představu o tom, kde se nacházejí formanty. Než začnete měřit, měli byste se podívat na spektrogramy kompletní sady samohlásek, abyste zjistili, zda nejde o diftongy a viděli, kde by bylo v každém vokálu nejvhodnější místo pro měření formantů. V krátkých jednoslabičných slovech jako jsou *bead, bid, bed, bad...*, ve kterých nejsou diftongy, je většinou vhodný interval blízko prostředku samohlásky. Jestliže jde zřejmě o diftongy, měli byste provádět měření ve dvou bodech, prvním blízko začátku, ale ne tak blízko, aby byl součástí souhláskového tranzientu (consonant transition)²⁰ a ve druhém blízko konce, ale zase dostatečně daleko od vlivů souhlásky.

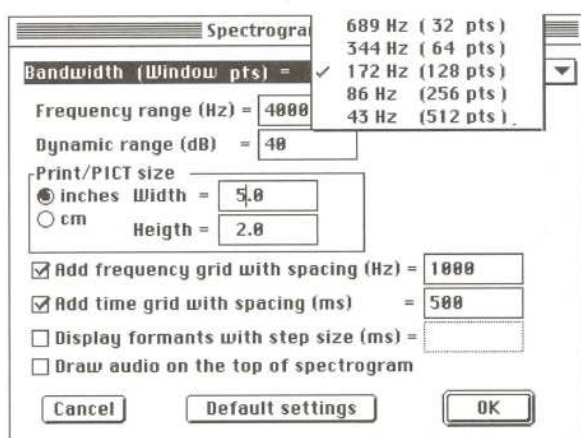
¹⁸ Formant je rezonanční frekvence dutiny hrdelní (laryngální), ústní (orální) a nosní (nazální). Frekvence formantů jsou v zápisech reprezentovány velkým písmenem F, za nímž následuje číslo odpovídající pořadí formantu. První formant se značí F1, druhý F2, atd. Obvykle se uvádí, že formant F1 odpovídá rezonanci dutiny hrdelní a F2 rezonanci ústní dutiny. Tyto dva formanty jsou při popisu samohlásek považovány za nejdůležitější. V některých případech, např. u labializovaných samohlásek, se bere v potaz též F3, ostatní, vyšší formanty samohlásek vypovídají spíše o zabarvení hlasu jednotlivého mluvčího.

¹⁹ Spektrogram bývá ve starších pracích (Hála, 1941; Romportl, 1973) též nazýván jako sonagram. Jde o trojdimenzionální grafické znázornění zvuku na dvojdimenzionální rovině, které nese informaci o změnách v trvání, frekvenci a intenzitě zvukových vln podél časových os. Na vertikální ose je zobrazena frekvence, na horizontální časový průběh, intenzitě odpovídá míra zabarvení.

²⁰ Souhláskový tranzient je pohyb formantu při přechodu z jedné hlásky ke druhé, tj. pohyb formantů v okrajových fázích samohlásek.

Samohlásky lze většinou dobře charakterizovat právě pomocí frekvencí prvního a druhého formantu, ale u vysokých předních a retroflexních (r-colored vowels)²¹ samohlásek byste však měli měřit i třetí formant. Bylo by ideální provádět měření ve stejný čas, kdy jsou všechny tři formanty, nebo alespoň první dva relativně stálé. Velmi často se však stává, že nejkonstantnější (steady-state)²² čas prvního formantu není nejlepším časem pro měření druhého formantu. To znamená, že nelze jednoduše definovat interval, který nejlépe charakterizuje samohlásky. Musíte brát v úvahu všechna data a vybrat čas, kdy budete moci provádět důkladná měření všech formantů. Napište si všechny postupy, které používáte při výběru bodu měření, jinak je nebudete schopni zopakovat, až budete později provádět další měření. Než začnete měřit, pořádně se podívejte na data. Ušetří vám to problémy, které se objeví, až zjistíte, že nemůžete provádět shodná měření, protože některé samohlásky nemůžete měřit v časech vybraných pro zbytek samohlásek.

Ve čtvrté kapitole jsme řešili problémy spojené s přenosem hlásek do počítače. Nyní musíme zvážit volby (options)²³ pro tvorbu vhodných spektrogramů, na kterých budeme zjišťovat formanty těchto hlásek. Na obrázku 5.1 jsou znázorněny volby Macquiner/Pcquiner systému, ostatní systémy analýzy mají podobné volby.



Obrázek 5.1 Volby, které musíte zvážit při tvorbě spektrogramů.

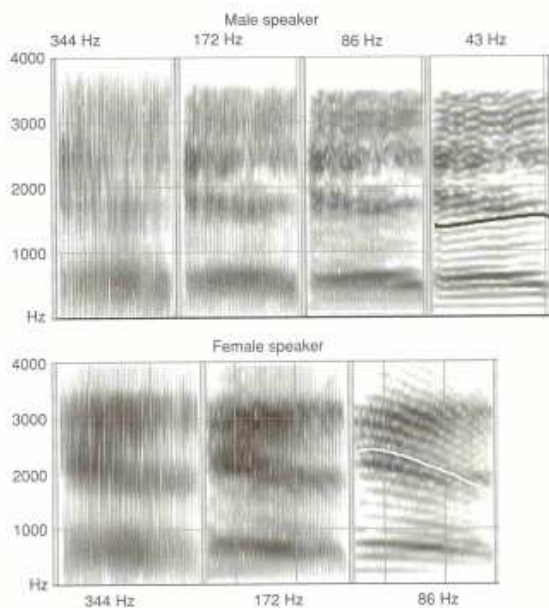
²¹ Retroflexní modifikované samohlásky se vyskytují v americké angličtině a v jihozápadních anglických dialektech. Jde o samohlásky, po kterých následuje **r**. Jsou jím ovlivněné a téměř po celou dobu svého trvání mají retroflexní kvalitu. Týká se to především zadních a středových samohlásek, ve kterých má přední část jazyka prostor pro změnu tvaru a během artikulace se pohybuje vzhůru.

Cit.: Roach, Peter: *A Little Encyklopaedia of Phonetics* [pdf online]. 2002. [cit. 2008-03-16]. Dostupné na <<http://www.personal.reading.ac.uk/~llsroach/encyc.pdf>>

²² V originále „steady state“: je to ta část samohlásky, kdy jsou formanty poměrně stabilní, konstantní, neměnné, nejvíce ustálené. Nedochozí v nich k žádným výrazným změnám, jsou tedy nedynamické.

²³ Výraz „options“ se používá v počítačové terminologii a lze jej překládat jako „možnosti“, „nastavení“ nebo „volby“.

Nejdříve musíte uvážit šířku pásma (bandwidth),²⁴ nebo délku okna (window length), která je zvýrazněná nahoře na obrázku. Na obrázku 5.2 je znázorněno, jaký účinek má změna šířky pásma při zkoumání samohlásky ve slově *head* vysloveném mluvčím mužského a ženského pohlaví.²⁵ Šířka pásma okolo 300 Hz se někdy nazývá širokopásmový spektrogram (wideband spectrogram). V každé řadě na obrázku 5.2 je nalevo širokopásmový spektrogram se šířkou pásma 344 Hz. (Konkrétní čísla používaná při analýze závisí na vzorkovací frekvenci (sampling rate). Matematické postupy vyžadují, aby počet bodů v okně analýzy byl násobkem 2 (32, 64, 128, 256, atd.). Šířka pásma je výsledkem dělení vzorkovací frekvence počtem bodů v okně analýzy, takže 344 je výsledkem $22\ 000/64$.) Na těchto spektrogramech se objeví první čtyři formanty mužského mluvčího a první tři formanty ženské mluvčí jako široké pásy. Při šířce pásma 43 Hz (= $22\ 000/512$), jako je na spektrogramu mužského mluvčího, který je napravo, vznikne úzkopásmový spektrogram (narrowband spectrogram), na kterém jsou zřejmé jednotlivé harmonické (harmonics).²⁶ U ženské mluvčí jsou harmonické viditelné při šířce pásma 86 Hz (= $22\ 000/256$).



Obrázek 5.2 Spektrogramy rozdílných šířek pásma samohlásek ve slově *head* vyslovené mluvčím mužského a ženského pohlaví. Na úzkopásmových spektrogramech napravo je vyznačena desátá harmonická u mužského mluvčího černou čarou a u ženské mluvčí bílou čarou.

²⁴ Šířka pásma znamená rozsah zobrazených frekvencí.

²⁵ Je obtížné vyjádřit, že se jedná o mluvčí obou pohlaví. V případě, že se jedná pouze o mluvčí ženského pohlaví, lze tuto skutečnost vyjádřit spojením „ženská mluvčí“. Avšak říci „vyslovené mužským a ženskou mluvčí“ se v tomto případě nezdá příliš vhodné. Navrhují dvě možnosti řešení, je možné nevyjádřit rozdíl v pohlaví gramatickými koncovkami a říci „vyslovené mužským a ženským mluvčím“ anebo lépe „mluvčím ženského a mužského pohlaví“.

²⁶ Jednotlivé harmonické jsou násobky základního tónu, základní frekvence. V odborné literatuře se pro ně užívá též pojmů svrchní, parciální, alikvotní nebo částkové tóny.

Když jsem v předchozí kapitole vysvětloval způsob analýzy výšky, zmínil jsem možnost použití spektrogramů. Nyní vidíme, že úzkopásmové spektrogramy které ukazují jednotlivé harmonické dobře informují o výšce. Na obrázku 5.2 vidíte, že všechny harmonické mužského mluvčího stoupají. Je obtížné přesně určit nárůst základního tónu (fundamental frequency),²⁷ ale desátá harmonická, která je vyznačená černou čarou stoupá přibližně z 1400 Hz na 1600 Hz. To znamená, že základní tón musí stoupat o desetinu, ze 140 Hz na 160 Hz. Ženská mluvčí má klesavou výšku hlasu, desátá harmonická (která je dobře vidět, v tomto případě označená bílou čarou) klesá přibližně z 2300 Hz na 1600 Hz, tudíž můžeme říci, že výška klesá z 230 Hz na 160 Hz.

Nejlépe lze pozorovat formanty na spektrogramech, na kterých je šířka pásma dostatečně široká, aby neukazovala jednotlivé harmonické. Hardwarové systémy pro analýzu mluvené řeči poskytovaly dříve výběr dvou šířek pásma, 45 Hz a 300 Hz. Stále převažuje tendence domnívat se, že šířka pásma 300 Hz („široké pásmo“) je nejvhodnější pro pozorování formantů a že šířka pásma 45 Hz („úzké pásmo“) je nejlepší pro pozorování zobrazení, která ukazují jednotlivé harmonické. Nyní však máme více možností a můžeme si lépe vybrat. Obecně řečeno, šířka pásma okolo 200 Hz je vhodná pro tvorbu spektrogramů, které ukazují formanty mužského hlasu a šířka pásma blíž 300 Hz je vhodnější pro většinu ženských mluvčích. Dětské hlasy vyžadují dokonce větší šířku pásma. Šířka pásma 172 Hz na obrázku 5.2 je vhodná pro mužského mluvčího a je pravděpodobně nejvhodnější i pro ženskou mluvčí, ačkoliv v jejím případě jsou jednotlivé harmonické zřetelné ke konci samohlásky. Jestliže neposoudíte jak 300 Hz tak i 172 Hz spektrogram, můžete získat špatný dojem o pohybu formantů u ženské mluvčí. Dokonce i malé změny základního tónu ovlivní vzhled formantů. V případě větších změn, které se objevují při nahrávání tónových jazyků nebo emocionálních forem řeči (emotional forms of speech),²⁸ nemusí být možné vytvořit dobře vypadající spektrogramy, které by zřetelně ukazovaly formanty.

Abyste získali nějakou zkušenost s umístěním formantů, můžete si udělat kopii obrázku 5.2 a pokusit se udělat čáru středem každého ze tří formantů na spektrogramu se šířkou pásma 172 Hz. Později uvidíme, že to počítačové programy dobře zvládají, ale občas dělají chyby.

Časové rozlišení (time resolution) spektrogramu závisí na šířce pásma. Přibližně řečeno, šířka pásma 200 Hz oddělí události, které jsou od sebe vzdáleny 1/200 sekund a šířka pásma 50 Hz oddělí ty, které jsou od sebe 1/50 sekund. Vibrace hlasivek, které mají frekvenci 200 Hz nebo

²⁷ Základní tón neboli také základní frekvence je frekvence vibrací hlasivek.

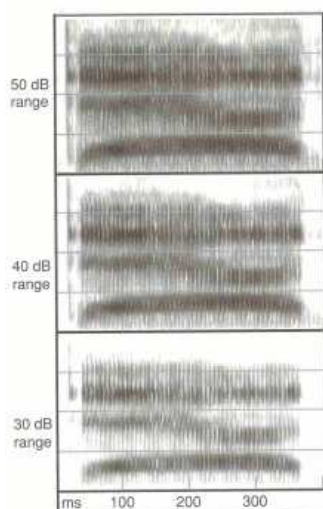
²⁸ Při překladu slovního spojení „emotional forms of speech“ bylo potřeba vyjádřit adjektivum odvozené od slova „emoce“. Nabízejí se tyto dvě možnosti: „emoční“ nebo „emocionální“. Rozhodla jsem se použít výraz „emocionální formy řeči“.

nižší (a které jsou od sebe nejméně 1/200 sekund) budou rozlišeny na spektrogramu šířky pásma 200 Hz. Jestliže je šířka pásma jen 50 Hz, analýza bude přesněji ukazovat frekvence oddělující jednotlivé harmonické (dokud bude výška nad 50 Hz), ale nebude ukazovat přesný čas jednotlivých vibrací hlasivek.

Další číslo, zobrazené v první volbě je délka okna, která je měřena na základě počtu bodů užitých při výpočtu každého bodu spektra. Jak jsme již viděli, při změně šířky pásma měníme toto číslo.

Další dvě volby na obrázku 5.1 jsou kmitočtový a dynamický rozsah (dynamic range).²⁹ Zatímco pro mužské mluvčí je pro zjišťování formantů vhodný kmitočtový rozsah 4000 Hz, pro ženské mluvčí je vhodnější rozsah 5000 Hz, to proto, aby byly vidět i vyšší formanty. Dynamický rozsah v dB ovlivňuje rozsah kontrastů, které jsou na spektrogramu, jako je to zobrazeno na obrázku 5.3. Dynamický rozsah 40 dB je vhodný pro většinu účelů, ale při zmenšení rozsahu na 30 dB se lépe ukážou formanty. Při rozsahu 50 dB je často příliš tmavé pozadí.

Spektrogramy na obrázku 5.3 představují amerického mluvčího vyslovujícího slovo *dad*. První formant je během hlavní části vokálu poměrně konstantní, ale druhý formant se zřetelně mění. Zcela určitě jde o diftong a nejlépe by se charakterizoval zaznamenáním formantů ve dvou bodech, řekněme na 120 ms a na 300 ms.

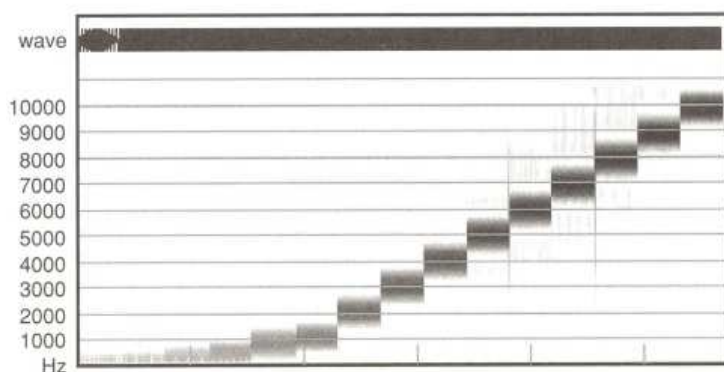


Obrázek 5.3 Změna dynamického rozsahu z 50 na 30 dB na spektrogramu mužského mluvčího, který vyslovuje slovo *dad*.

²⁹ Dynamický rozsah je poměr mezi největší a nejmenší hodnotou, je to bezrozměrná hodnota a v případě velkého rozsahu se používá logaritmická škála a vyjadřuje se v decibelech (dB).
Cit.: Skarnitzl, Radek: *Dynamický rozsah* [online]. 2007-11-07. [cit. 2008-03-16]. Dostupné na <http://cs.wikipedia.org/wiki/Dynamick%C3%BD_rozsah>

Zbývající volby na obrázku 5.1 se týkají přidání stupnice frekvence a času a možnosti zobrazení trajektorií formantů (formant tracks)³⁰ s určitou velikostí kroku, ale o tom bude řeč později.

Jsou ještě dva body, které bychom měli zmínit než opustíme tvorbu spektrogramů. Za prvé to že, na spektrogramech jsou většinou zvýrazněny (have boost to the intensity) vyšší frekvence. Jestliže se objeví stejně tmavé pásy v místech s odlišnými frekvencemi, tak to neznamena, že tyto frekvence mají stejnou energii. Obrázek 5.4 ukazuje, co se stane, když se zobrazí série vln s rozdílnou frekvencí, ale se stejnou amplitudou. V horní části obrázku je zobrazen oscilogram (waveform),³¹ ale protože je komprimovaná (compressed)³² stupnice času, vibrace jsou vidět odděleně pouze při nejnižší frekvenci, při 50 Hz. Je zřejmé, že všechny vlny mají stejnou amplitudu. Na spektrogramu jsou však velmi nízké frekvence stěží viditelné. Pásy představující rozdílné frekvence tmavnou, když stoupá frekvence.



Obrázek 5.4 Spektrogram sady vln, které mají stejnou amplitudu s frekvencemi 50, 100, 200, 400, 800, 1000 Hz a potom každých 1000 Hz až do 10 000 Hz. Samotná vlna je v horní části obrázku.

Poslední bod se týká svislých čar, které jsou vidět na širokopásmových spektrogramech. Každé otevření a zavření hlasivek vytváří glotální puls (glottal pulse),³³ krátký zvuk, který se jeví jako svislá čára. Jakýkoliv obrázek na obrazovce počítače nebo vytvořený tiskárnou se však také skládá z blízké rozmístěných vertikálních (a horizontálních) čar. Někdy je vzdálenost mezi vertikálními čarami pulsů hlasivek těsně spojena s počtem čar vytvořených programem. Výsledkem je interferenční obrazec s některými skupinami čar pulsů hlasivek tmavšími než ostatní, jako je to na obrázku 5.5.

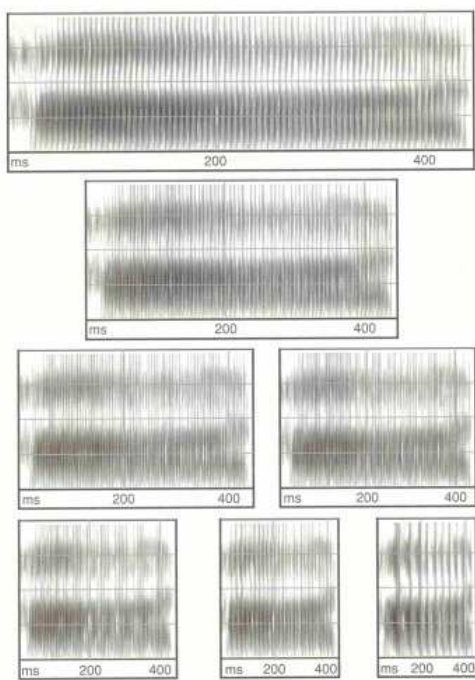
³⁰ Skarnitzl mi navrhl termín „trajektorie formantu“. Je možné též použít opisu „grafické zvýraznění formantu“ či sousloví „stopa formantu“.

³¹ Oscilogram neboli také křivka zvukové vlny je zobrazením časového průběhu vlny.

³² K výrazu „komprimovaná“ lze použít český protějšek „stlačená“.

³³ Lze též použít domácí výraz „hlasivkový puls“.

Obrázek 5.5 představuje sedm spektrogramů té samé hlásky. Liší se pouze šířkou, počtem pixelů alokovaných každému pulsu hlasivek na obrazovce počítače a výtisku. Na vrchním spektrogramu jsou vidět všechny detaily hlásky, lze na něm rozpoznat každý puls hlasivek. Spektrogram ve druhé řadě působí dojmem, že pulsy hlasivek jsou nepravidelně rozmístěny. Na spektrogramech ve třetí řadě, hlavně na začátku toho nalevo a na konci toho napravo se zdá, jako kdyby byly pulsy hlasivek ve shlucích. První spektrogram v poslední řadě pokračuje v tomto trendu, ale v prostředním spektrogramu nastane něco jiného, zdá se, že výška klesla více než na polovinu. Vypadá to, že je tam před značkou 200 ms dvanáct stejnoměrně rozmístěných pulsů, oproti tomu samému intervalu vrchního spektrogramu, kde je třicet jedna pulsů. Konečně, na posledním spektrogramu se zdá, že tato samohláska má pouze devět pulsů, jako kdyby šlo o třepený typ fonace (creaky voice).³⁴ Na všech spektrogramech je však stále ta samá hláska s pravidelnými vibracemi, jako je zobrazena na vrchním spektrogramu.



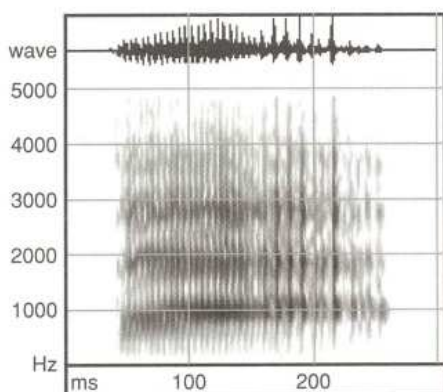
Obrázek 5.5 Spektrogramy toho samého oscilogramu (samohlásky ve slově *hod* vyslovené ženskou mluvčí) s různým stupněm komprese stupnice času.

Tento jev patří ke způsobu, jakým jsou spektrogramy vytvářeny, zobrazovány nebo tištěny. Problémem je to pouze v případě hlásek s poměrně nedynamickým základním tónem. V těchto případech byste si měli dávat pozor při interpretaci shluků vibrací hlasivek, které

³⁴ Slovní spojení „creaky voice“ se nejčastěji překládá jako „třepená“ nebo „stlačená fonace“. Při tomto typu fonace jsou hlasivkové vazy v klidové poloze přitisknuty pevně k sobě, tudíž je amplituda kmitání hlasivkových vazů daleko menší než u běžné fonace. Cit.: Skarnitzl, Radek: *Filtrová teorie produkce řeči* [pdf online]. [cit. 2008-03-16]. Dostupné na < http://fu.ff.cuni.cz/vyuka/akustika/source_filter.pdf>

mají zjevně stoupající a klesající amplitudu. Ve skutečnosti možná v činnosti hlasivek není vůbec žádná změna. Jestliže při zvýšení časového rozlišení zmizí změny amplitudy, jde pouze o výsledek zobrazení. Změny amplitudy však mohou být i skutečné. Na obrázku 5.6 je zobrazen spektrogram, nad kterým je oscilogram. Pravá strana spektrogramu vypadá jako poslední spektrogram na obrázku 5.5. Vlna v horní části obrázku však naznačuje, co skutečně nastalo, v poslední části samohlásky výška klesala, protože mluvčí vyslovil samohlásku poněkud třepeným typem fonace.

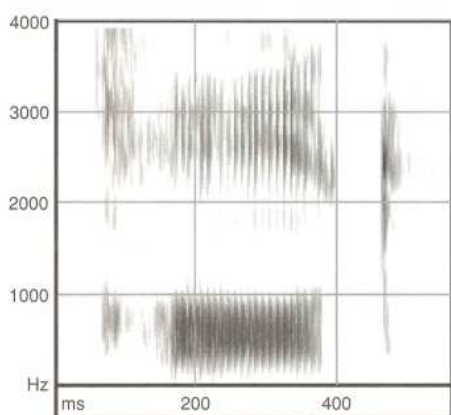
Dobré spektrogramy nám velmi pomáhají při určování formantů. Ne vždy je to však snadné. Dříve než začnete, musíte vědět, kde byste měli formanty hledat. Nejlepší metodou je hledat jeden formant na každých 1000 Hz. Například samohláska *a* mužského mluvčího má formanty okolo 500, 1500 a 2500 Hz (všechny jsou o trochu vyšší u ženské mluvčí). Ostatní samohlásky budou mít formanty nahoru nebo dolu od tohoto středního rozsahu. Existují však výjimky od obecného pravidla jeden formant na 1000 Hz. Bylo by přesnější tvrdit, že na každých 1000 Hz je průměrně jeden formant. Zadní vokály mohou mít dva formanty pod 1000 Hz, ale žádný mezi 1000 a 2000 Hz a třetí formant někde mezi 2000 a 3000 Hz. Jestliže víte, že analyzujete nízký zadní vokál, nesmíte být překvapeni, když naleznete na spektrogramu jeden široký pás, který ve skutečnosti odpovídá dvěma formantům blízko u sebe, které jsou pod 1000 Hz. Obrázek 5.7 představuje spektrogram mé výslovnosti slova *caught*, ve kterém jsou první dva formanty velmi blízko u sebe. Jestliže jsou na tomto spektrogramu tři formanty pod 3000 Hz, musí být dva pod 1000 Hz.



Obrázek 5.6 Oscilogram a spektrogram slova *bat*, které mluvčí na konci samohlásky vyslovil třepeným typem fonace.

Někdy není hned zcela zřejmé, zda obzvláště široký pás představuje jeden nebo dva formanty. Na obrázku 5.8 je zobrazen spektrogram slova *bud*, které je vyslovené ženou, pro niž je charakteristický kalifornský akcent. Pod 1000 Hz je široký pás, ale jde v tomto případě o

jeden nebo o dva formanty jako na obrázku 5.7? Všimněte si, že na obrázku 5.8 je jeden formant zřetelně okolo 1500 Hz a další formant (additional formant)³⁵ výše, musíme to brát tak, že pod 1000 Hz je pouze jeden formant. Zdá se, že je tam blízko prvního formantu něco jako vedlejší formant (extra formant),³⁶ který způsobuje tento tmavý pás širším. Podle této jediné samohlásky není možné říci, zda je dodatečná energie pod nebo nad prvním formantem. Další analýza hlasu tohoto mluvčího ukázala, že bez ohledu na to o jakou šlo samohlásku, byla energie často v oblasti okolo 1000 Hz. Falešný formant nesouvisí s kvalitou samohlásky, ale je charakteristickou součástí hlasu určitého mluvčího. Dobře nám to ukazuje, že je nutné se podívat se na reprezentativní vzorek hlasu mluvčího, dříve než začneme s měřením.



Obrázek 5.7 Spektrogram slova *caught*, ve kterém jsou první dva formanty velmi blízko u sebe.

S umístováním formantů vám může pomoci zobrazení trajektorií formantů, jak je určí počítač (jde o předposlední volbu na obrázku 5.1). Počítač určí formanty tak, že vyhledá vrcholy spektra, o němž bude řeč v následující části. Obrázek 5.9 představuje spektrogram sady samohlásek ásámštiny, které jsou obzvláště zajímavé kvůli poslední samohlásce, ve které je jazyk v pozici samohlásky v anglickém slově *pot*, ale ve skutečnosti má téměř maximální zaokrouhlenost. Takovou samohlásku jsem ještě neslyšel v žádném jiném jazyce.

Detektor formantů (formant tracker)³⁷ našel správně formanty pro první dvě samohlásky, ale ve třetí samohlásce vyznačil mezi prvním a druhým formantem jeden formant navíc. V této oblasti je nějaká energie, ale když víme, že tato samohláska zní jako **ɛ**, musíme zvážit, zda nejde o falešný formant. Když se podíváme na spektrogram, je zřejmé že tento pás není zcela

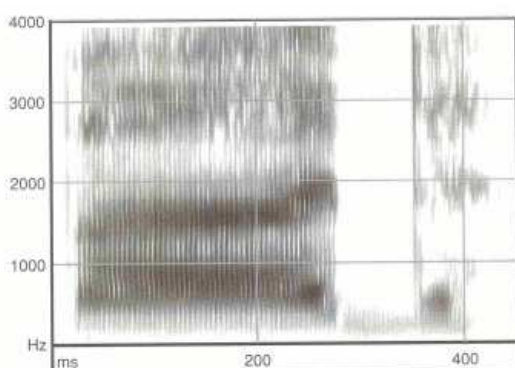
³⁵ V tomto případě nejde o terminologizovaný fonetický výraz. Lze tedy použít různé ekvivalenty jako jsou „další“, „doplňkový“ nebo „přídavný formant“.

³⁶ Slovní spojení „extra formant“ lze přeložit jako „vedlejší“, „mimořádný“ nebo „dodatečný formant“.

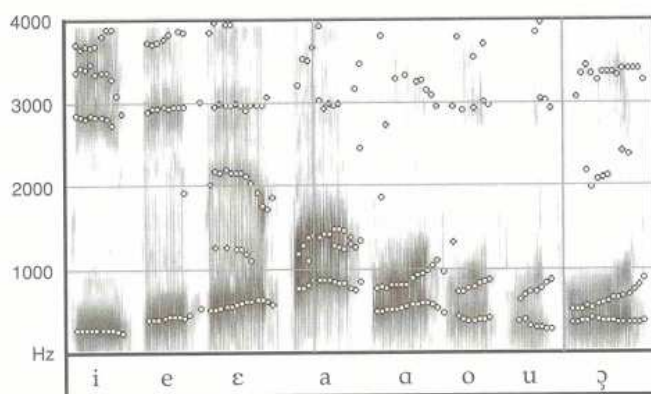
³⁷ Výraz „formant tracker“ lze překládat jako „detektor trajektorie formantů“. Toto slovní spojení lze zkrátit na „detektor formantů“.

přesně definovaný jako v případě ostatních formantů. V samohlásce **a** jsou chyby ve druhém formantu. Ve zbylých samohláskách určil detektor formanty poměrně správně, obzvlášť třetí formant, který je velmi obtížně vidět. Já bych přesto nevěřil určení čtvrtého formantu ve většině těchto samohlásek.

Některé detektory formantů mají volby, které vám dovolí stanovit minimální hodnotu amplitudy formantu požadovanou pro zobrazení trajektorie formantů. Jestliže detektor formantů najde formant s šířkou pásma větší než 400 Hz, bude mít nízkou amplitudu a tato část trajektorie formantů by měla být vynechána.



Obrázek 5.8 Spektrogram slova *bud*, vysloveného ženou, pro kterou je charakteristický kalifornský akcent angličtiny.



Obrázek 5.9 Spektrogram sady ásámských samohlásek ve slabikách tvaru **p_t** s přidanými trajektoriemi formantů, které jsou vyznačené malými bílými kroužky.

5.2 SPEKTRA

Detektory formantů nejsou jediným způsobem, jak měřit formanty samohlásky. Lze je měřit též pomocí kurzoru myši přímo na spektrogramu, tento způsob je však pouze přibližný a

vhodný pouze pro rychlý odhad. Lepší metodou je výpočet a zobrazení spektra (spectrum),³⁸ amplitudy každého komponentu frekvence v daném místě a čase. Spektrum může být spočítáno různými způsoby, z každého z nich získáme poněkud odlišný dojem o umístění formantů. Na obrázku 5.10 je zobrazeno typické menu, které ukazuje možnosti týkající se tohoto problému.



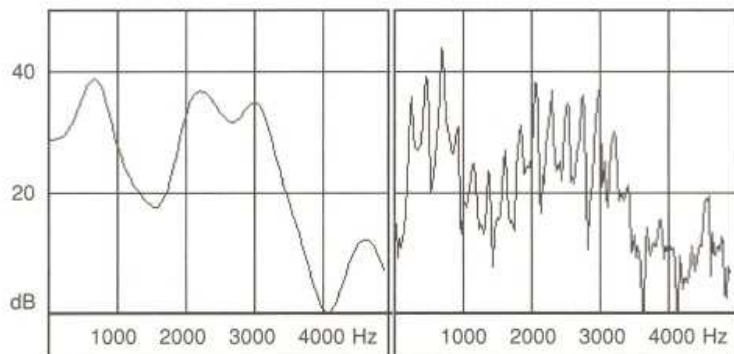
Obrázek 5.10 Volby, které musíte uvážit při tvorbě spektra.

První volbou je šířka pásma pro analýzu. Jestliže chcete přesně vědět z jakých frekvencí se hláska skládá, musíte zkoumat poměrně dlouhý úsek zvukové vlny. Jde o ten samý typ volby, jaký jsme řešili, když jsme uvažovali o nastavení spektrogramu znázorněném na obrázku 5.1. Spektrum v levé části obrázku 5.11 bylo vytvořeno užitím šířky pásma 344 Hz. Pro výpočet přítomných frekvencí používá jen 32 bodů samplované vlny (sampled wave).³⁹ Tato hláska byla samplována na 11 000 Hz, ne na 22 000 Hz jako na spektrogramech uváděných dříve ($11\,000/32 = 344$). Tímto je spektrum poměrně přesné ve sféře času, je to však na úkor neoddělování jednotlivých harmonických. Jsou zobrazeny dohromady stejně jako na širokopásmovém spektrogramu na obrázku 5.2. Úzkopásmové spektrum (narrowband

³⁸ Spektrum je dvojdimenzionální graf, který ukazuje na horizontální ose frekvenci a na ose vertikální relativní amplitudy frekvenčních komponentů zvukové vlny. Na rozdíl od spektrogramu nezobrazuje časový průběh, protože udává hodnoty v konkrétním okamžiku.

³⁹ Anglické slovo „sample“ znamená „vzor“, „vzorek“. Výraz „smplovaná“ je tudíž možno přeložit jako „vzorkovaná vlna“, odpovídající sloveso je „vzorkovat“.

spectrum) v pravé části obrázku 5.11, které má šířku pásma 21 Hz a okno s 512 body, odděluje harmonické složky hlásky.

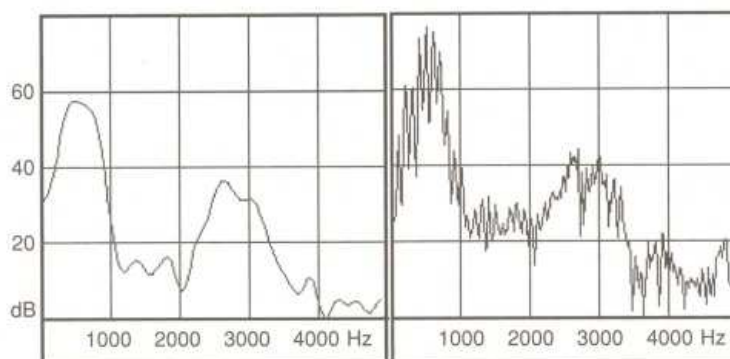


Obrázek 5.11 Spektrum vytvořené v prostředí samohlásky ve slově *head* vysloveném ženskou mluvčí. Nalevo je spektrum, které má šířku pásma 344 Hz, a napravo to, které má šířku pásma 21 Hz.

Nelze říci, které ze dvou spekter na obrázku 5.11 je vhodnější pro určení polohy formantů a jak uvidíme později, žádné z nich není vhodnější pro soustavné měření frekvencí formantů. Na širokopásmovém spektru, které je nalevo, naleznete v této samohlásce vrcholy formantů okolo 650, 2250 a 3000 Hz. V jiných samohláskách může být první formant blízko základního tónu, takže není jisté, který vrchol odpovídá prvnímu formantu, nebo se může stát, že dva formanty jsou příliš blízko u sebe. Úzkopásmové spektrum napravo ukazuje harmonické, což může být někdy užitečné, ale abyste určili vrcholy formantů, tak si však musíte představit křivku plynule probíhající okolo nejvyšších harmonických. V tomto případě by taková křivka měla ukázat, že vrchol formantu je mezi druhou a třetí harmonickou (dvěmi nejvyššími harmonickými) ale o něco blíže třetí (nejvyšší) harmonické.

Když jsou formanty blízko u sebe, jako na spektru samohlásky na obrázku 5.12, ani široko- ani úzkopásmové spektrum nevypovídá dobře o frekvencích formantů. (Jde o tu samou samohlásku jako na spektrogramu na obrázku 5.7.) První dva formanty se jeví jako jeden vrchol pod 1000 Hz. Z tohoto spektra nelze určit jejich frekvence.

Existuje však i jiný způsob, jak určit formanty. Nemusíte se pokoušet umístit vrcholy na široko- nebo úzkopásmovém spektru, použitím počítače lze vypočítat frekvence formantů z tzv. LPC spektra. Tato volba je ukázána ve spodní části obrázku 5.10 „Zobraz LPC spektrum“.



Obrázek 5.12 Spektrum vytvořené v prostředku samohlásky ve slově *caught* vysloveném mužským mluvčím. Nalevo je spektrum, které má šířku pásma 173 Hz, a napravo to, které má a šířku pásma 21 Hz.

Spektra, o nichž už byla řeč, byla podrobena Fourierově analýze (Fourier analysis).⁴⁰ Cílem této analýzy není odhadnout počet přítomných formantů. Určí pouze množství energie každé rozdílné frekvence a nechá na uživateli, aby sám určil, které vrcholy spektra odpovídají vrcholům formantů. Alternativní metodou při určování frekvencí formantů je LPC (lineární prediktivní kódování)⁴¹ pracující na základě počtu vrcholů (označovaných jako póly), které odpovídají formantům. Předpokládá, že vlna může být popsána jako součet několika formantových pólů (formant poles),⁴² a pak určí, která sada pólů (frekvencí formantů a amplitudy) by s co možná nejmenší odchylkou odpovídala této vlně. (Nejde o popis matematického postupu, ale o výsledek toho, co se stane. Jestliže chcete vědět více, podívejte se do druhé edice mých *Elements of Acoustic Phonetics*.)

Na obrázku je zobrazeno 5.13 to samé spektrum jako na obrázku 5.11 (prostředek samohlásky ve slově *head* vyslovené ženskou mluvčí), je však proložené (superimposed)⁴³ LPC spektrem. Všimněte si, že LPC spektrum, narozdíl od širokopásmového spektra nalevo, poskytuje lehce odlišný obrázek toho, kde se vyskytují formanty. První formantový pól na LPC spektru je nepatrně vyšší, trochu blíže nejvyšší harmonické na úzkopásmovém spektru nalevo a druhý

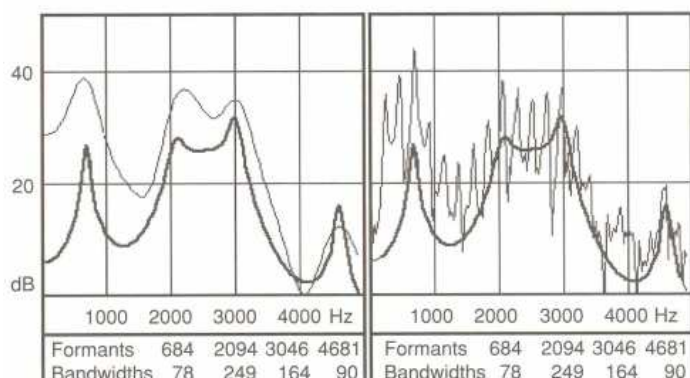
⁴⁰ Fourierova analýza je matematická technika založená na dekompozici signálu do sinusoid. Je pojmenována po Jean Baptiste Joseph Fourierovi, jenž žil v období 1768–1830. Působil ve Francii jako matematik a fyzik. Cit.: Vlček, Vítězslav Vít: *Dvojměrná frekvenční a směrová filtrace - pomocí diskrétní Fourierovy transformace* [pdf online]. 2002. [cit. 2008-03-16]. Dostupné na <<http://home.zcu.cz/~vsoft/dft/10konfvlcek.pdf>>

⁴¹ LPC je anglická zkratka pro Linear Predictive Coding. Lze též použít český protějšek lineární predikce. Cit.: Skarnitzl, Radek: *Analýza řečového signálu* [pdf online]. [cit. 2008-03-16]. Dostupné na <<http://fu.ff.cuni.cz/vyuka/akustika/analyza.pdf>>, ale nejčastěji se používá pouze zkratka LPC.

⁴² Je otázkou, jak tento termín přeložit, nabízejí se zejm. výrazy formantové póly, formantové vrcholy, póly formantů. Termín „pól“ je převzat z technické literatury. Formantové frekvence se někdy nazývají póly (Hayward, 2000, s. 82).

⁴³ Výraz „superponovat“ se v české fonetické terminologii příliš nepoužívá, a proto překládám výraz „superimposed“ jako „proložené“. Nabízí se též ekvivalent „překryté“.

formantový pól je o něco níže, opět blízko nejvyšší harmonické na úzkopásmovém spektru. Který způsob analýzy je správný, ten, kterým určíme formanty na širokopásmovém spektru, nebo ten, kterým je určíme na LPC spektru?



Obrázek 5.13 To samé spektrum jako na obrázku 5.11 ale je proložené LPC spektrem (tlustá čára) a s formanty vypočítanými LPC, které jsou uvedeny ve spodní části.

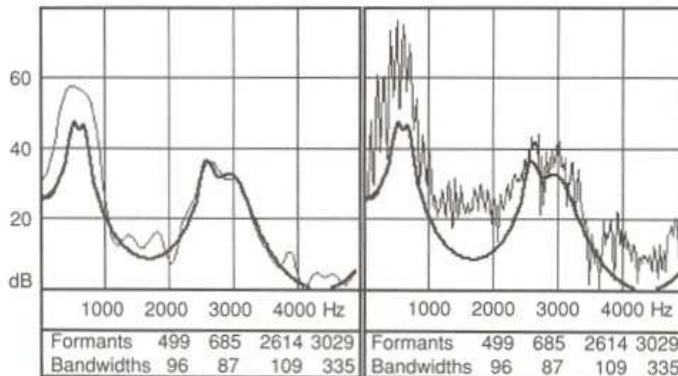
Není snadné říci, který způsob analýzy je správný. Každá analýza bere při výpočtu křivky spektra rozdílným způsobem v potaz energii základního tónu a vliv falešných formantů. Jestliže je cílem analýzy určit co nejvíce o pohybu jazyka a rtů, mělo by být minimalizováno spektrální příspěvek základního tónu a falešných formantů. Množství energie základního tónu je komplexní záležitostí, záleží na blízkosti F1 základnímu tónu a stavu hlasivek. Jestliže hlasivky volně vibrují, produkují dyšnost (breathiness), základní tón bude mít větší intenzitu blízkou vyšším frekvencím a bude také přidána tracheální rezonance (tracheal resonances).⁴⁴ (O dyšném typu fonace (breathy voice)⁴⁵ bude řeč později v kapitole 7.) A navíc, mírná nazalizace ovlivní spektrum uvedením rezonancí, které souvisejí s nosní dutinou. Jestliže se zaměříme na tvar hlasového ústrojí reflektovaný pohybem jazyka a rtů, potřebujeme najít frekvence formantů spojené s těmito pohyby. Není však zcela jasné, který typ analýzy lépe vyhovuje.

Je však jisté, že když necháte počítač najít formanty na LPC spektru, získáte možná více odpovídající výsledky, než když se sami snažíte najít vrcholy na spektru. Počítačový algoritmus poskytuje hodnoty frekvencí formantů a šířek pásma (které dohromady tvoří tzv. kořeny LPC rovnice). Je to vhodné zejména, když jsou dva formanty blízko u sebe. Obrázek 5.14 představuje to samé spektrum jako obrázek 5.12, prostředek samohlásky ve slově *caught*,

⁴⁴ Jedná se o rezonanci průdušnic.

⁴⁵ Při tomto typu fonace jsou hlasivky, přesněji hlasivková štěrbina, v klidové pozici alespoň částečně otevřené, stále mezi nimi proudí vzduch, který vytváří slyšitelný zvuk. Cit.: Skarnitzl, Radek: *Filtrová teorie produkce řeči* [pdf online]. [cit. 2008-03-16]. Dostupné na <http://fu.ff.cuni.cz/vyuka/akustika/source_filter.pdf>

širokopásmové spektrum je vlevo a úzkopásmové vpravo. Obě FFT⁴⁶ spektra jsou proložena tím samým LPC spektrem (které je vyznačeno tlustou čarou). První dva formanty jsou blízko u sebe, což ale LPC analýza jasně rozlišila, stejně tak jako třetí i čtvrtý formant.

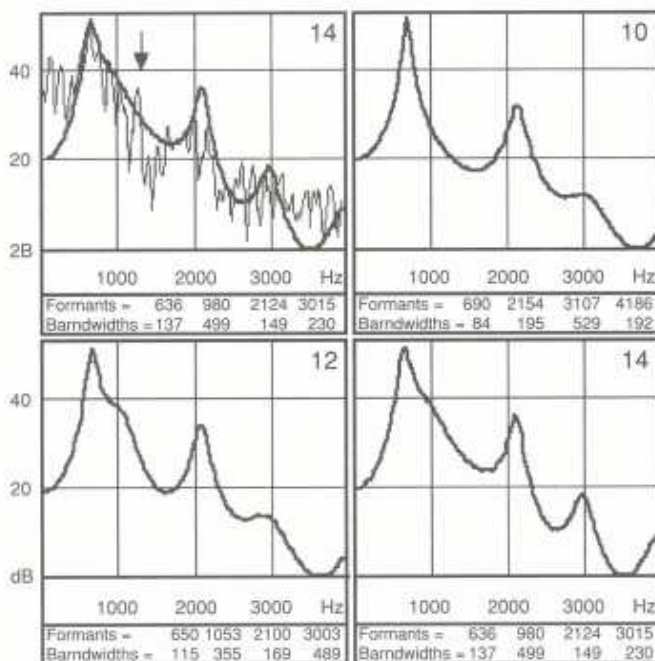


Obrázek 5.14 To samé spektrum jako na obrázku 5.12 ale je proloženo LPC spektrem (tlustá čára) a s formanty vypočítanými LPC.

Ne vše, co počítačový algoritmus určí jako sadu formantů, bychom my vybrali pomocí našich znalostí, kde hledat formanty. Pro lepší pochopení musíme prostudovat koncepty LPC analýzy. Jedna z voleb na obrázku 5.10 stanovuje počet koeficientů použitých při výpočtu LPC. Určuje, kolik je vypočítáno formantových pólů (ne všechny z nich jsou zobrazeny). Obecně řečeno, pro každý formant potřebujete dva koeficienty a další dva pro výpočet vyšších formantů. (Někdo navrhuje, že potřebujete další dva koeficienty pro vypočítání vrcholku spojeného s glotálním zdrojem (glottal source)).⁴⁷ Při vzorkovací frekvenci 11 200 Hz se rozsah výpočtu zmenší přesně na polovinu (5600 Hz). Můžete tudíž očekávat, že naleznete 5 nebo 6 formantů a standard pro LPC výpočet (bez zřetele na vrchol způsobený glotálním zdrojem) je stanoven na 14 koeficientů ($6 \times 2 + 2$). Abychom viděli, jak to funguje podíváme se na analýzu třetího formantu na obrázku 5.9 (jde o ásámskou samohlásku s falešným formantem).

Obrázek 5.15 představuje výsledek použití rozdílného počtu koeficientů při analýze ásámské samohlásky ϵ . Levý horní panel je standardní analýzou, je na něm zobrazeno jak FFT tak LPC spektrum. Na FFT spektru vidíte lokální vrchol (označený šipkou), který je spojován s tím, co označujeme jako falešný formant. Je však otázka, zda LPC analýza určí tento vrchol jako formant.

⁴⁶ „FFT“ je anglická zkratka pro „Fast Fourier Transformation“, česky „rychlá Fourierova transformace“. Počítá se korelace sinusových a kosinusových složek se složenou vlnou. Vysoká korelace znamená, že složená vlna a daná sinusovka stoupají a klesají společně. Cit.: Skarnitzl, Radek. *Analýza řečového signálu* [pdf online]. [cit. 2008-03-16]. Dostupné na <<http://fu.ff.cuni.cz/vyuka/akustika/analyza.pdf>>



Obrázek 5.15 Vliv změny počtu koeficientů (který je vyznačený v pravém horním rohu každého spektra) v LPC analýze. Šipka v horním levém panelu označuje lokální vrchol na LPC spektru.

Na ostatních panelech je pro zjednodušení zobrazeno pouze LPC spektrum. Když má 10 koeficientů, výpočet nalezne 5 formantů. Panel vpravo nahoře představuje vygenerované LPC spektrum. V zobrazeném rozsahu frekvence jsou čtyři formanty a pravděpodobně další formant vyšších frekvencí. Spektrum je přijatelné pro tuto samohlásku a formanty uvedené pod spektrem mají hodnoty, které odpovídají ϵ samohlásce. Jestliže tam má 12 koeficientů, tak jako v spodním levém panelu, výpočet předpokládá, že pod 5600 Hz je šest formantů a na zobrazeném spektru se objeví malý ohyb (bend).⁴⁸ Frekvence a šířka pásma odpovídající tomuto ohybu jsou udané pod spektrem a ukazují, že druhý formant má frekvenci 1053 Hz, hodnotu, která je nepochybně vyloučena pro samohlásku znějící jako ϵ . Spektrum, které užívá 14 koeficientů a které není proloženo úzkopásmovým FFT spektrem je vpravo dole. Lehký ohyb křivky, který odpovídá falešnému formantu, je méně zřejmý, ale uvedená frekvence znovu ukazuje, že druhý formant je velmi blízko prvnímu formantu, což víme, že není pravda. Nelze říci, kterou z těchto analýz bychom měli použít a které z hodnot formantů jsou správné. Analýza s 10 koeficienty dobře vypadá, protože ukazuje právě čtyři formanty (s dalším vyšším formantem, který je mimo zobrazovaný rozsah). Avšak když omezíme analýzu na malý počet formantů, první vrchol by byl pólem, který zahrnuje veškerou energii této oblasti

⁴⁷ „Glottal source“ znamená „glotální“ neboli „hlasivkový zdroj“.

⁴⁸ Na spektru je v podstatě lehké zakřivení, které však neodpovídá formantu.

nad a pod opravdovým prvním formantem. Analýza používající 12 a 14 koeficientů udává výrazně nižší hodnoty prvního formantu. Většinou je nejlepší užít standardní počet koeficientů a po uvážení charakteru samohlásky a vzhledu proloženého úzkopásmového FFT spektra se rozhodnout, které frekvence formantů brát jako reprezentující samohlásku. Když se podívám na analýzu samohlásky, která je na obrázku 5.15 nahoře vlevo, řekl bych, že hodnoty prvních třech formantů jsou 636, 2 124 a 3150 Hz, dané analýzou se 14 koeficienty. Frekvenci, která byla zaznamenána na 980 Hz, můžeme považovat pouze za ukazatel něčeho v kvalitě hlasu mluvčího. Tento závěr můžeme podpořit tím, že šířka pásma tohoto vedlejšího formantu je poměrně velká, 499 Hz, je dvakrát větší než v ostatních udaných pólech. Šířka pásma formantu je nepřímě úměrná intenzitě formantu. Cokoliv, co má šířku pásma větší než 400 Hz, je obvykle zanedbatelné.

Přeformulováním dříve uvedeného obecného tvrzení (rule of thumb) můžeme říci, že se obvykle pro mužského mluvčího používají pro výpočet dva koeficienty na každých tisíc nebo část tisíce Hz a další dva pro vypočítání vyšších formantů. Ženská mluvčí bude mít formanty o něco vyšší, a proto budete možná potřebovat pro výpočet dva koeficienty na každých 1200 Hz. Pamatujte, že bez ohledu na rozsah zobrazované frekvence se výpočet analýzy zmenší na polovinu vzorkovací frekvence.⁴⁹ Při vzorkovací frekvenci 11 200 Hz použijte pro mužského mluvčího 14 koeficientů a 12 pro ženskou mluvčí. Při vzorkovací frekvenci 22 400 Hz použijte 20 – 24 koeficientů. Při analýze řeči dětí budete potřebovat méně koeficientů, protože mají formanty dále od sebe a v daném rozsahu frekvence jich bude méně. Jestliže při tvorbě analýzy použijete o něco větší počet koeficientů než je nutné, dostanete nějaké falešné formanty, když se však zamyslíte, tak je budete schopni rozpoznat, což počítač nedokáže.

Měli bychom také poznamenat, že v některých systémech analýzy má volba počtu koeficientů v nastavení FFT/LPC vliv, který přesahuje typ zobrazovaného spektra. Ovlivňuje také volbu Zobrazení formantů ve volbách spektrogramu na obrázku 5.1. Trajektorie formantů samohlásek ásámštiny byly zobrazeny na obrázku 5.9. Nyní vidíme proč na tomto obrázku byly falešné formanty. Trajektorie byly vytvořeny při tvorbě LPC analýzy během každých 10 ms užitím standardního počtu koeficientů. Následkem toho byly vygenerovány falešné formanty. Detekce formantů je užitečná nejenom proto, že ukazuje, kde mohou mít FFT/LPC analýzy problémy.

⁴⁹ Vzorkovací frekvence znamená, kolikrát za sekundu byla nahrána amplituda vlny. Čím vyšší je vzorkovací frekvence, tím kratší je interval mezi jednotlivými vzorky a samplovaná vlna se více podobá původní vlně. Je důležité správně vybrat vzorkovací frekvenci, protože určuje vrchní hranici kmitočtového rozsahu při nahrávání. Většinou se používá vzorkovací frekvence 10 000 Hz, která dovolí nahrát frekvence v rozsahu 0 – 5000 Hz (Hayward, 2000, s. 68).

O dalších možnostech FFT/LPC nastavení, které jsou ukázány na obrázku 5.10, bude řeč později, až budeme pojednávat o souhláskách a typech fonace. S výjimkou délky okna (které by mělo být při analýze samohlásek ponecháno ve své standardní podobě) se další volby většinou týkají vzhledu zobrazení a neovlivňují analýzu.

5.3 DIAGRAMY SAMOHLÁSEK⁵⁰

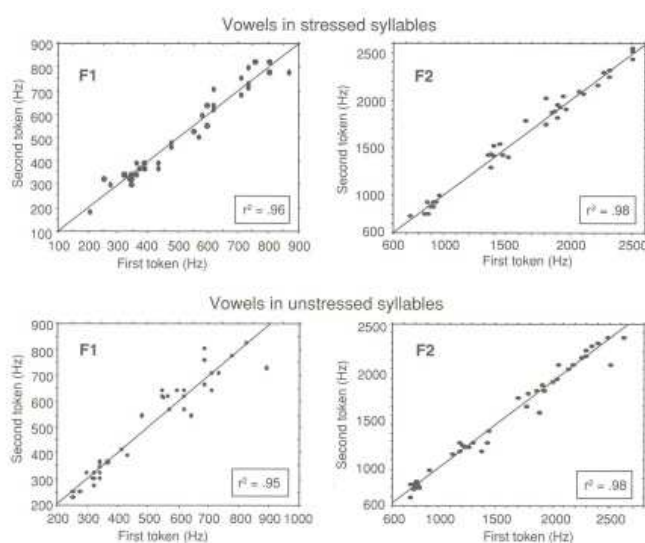
Na začátku této kapitoly jsme tvrdili, že chceme podat hodnotný vědecký popis samohlásek daného jazyka. Vezměme si jako ukázkou poměrně jednoduchý případ samohlásek jazyka banawa, arakwamského jazyka, kterým se mluví v Amazonském deštném pralese. Na tomto jazyce jsem pracoval spolu s kolegou Danem Everettem, který studoval jeho fonologii. Jak bylo nastíněno již v první kapitole, naším prvním úkolem bylo pořídit dobrou sadu slov, které ilustrují jeho samohlásky. Jazyk banawa má pouze čtyři kontrastující samohlásky, a přestože jsme nahrávali samohlásky v přízvučných i v nepřízvučných slabikách, nepotřebovali jsme dlouhý seznam slov. Slova, která jsme použili jsou v tabulce 5.1. Podařilo se nám najít sady minimálně kontrastujících slov (minimally contrasting sets of words) se samohláskami po **t** v přízvučných a po **f** v nepřízvučných slabikách, ale měli jsme potíže při hledání sady minimálně kontrastujících slov, které ilustrují samohlásky po **b**. Nahráli jsme pět mluvčích, kteří dvakrát vyslovili slova z tabulky 5.1. Použitím výše zmíněných metod jsme v nejstálejší části každé ze samohlásek změřili první tři formanty (celkem 5 mluvčích vyslovilo dvakrát 16 slov = 160 samohlásek, 80 z nich v přízvučných slabikách).

Přízvučné slabiky		Nepřízvučné slabiky	
tifa	pít vodu	tafi	jezení
tefe	potrava (m.)	tafe	potrava (f.)
tafa	jíst	tafa	jíst
tufa	načrtnout	tafu	jíst
bita	moskýt	ibi	navzájem
befa	jiný	ibe	proužek
bata	vybrat si	iba	umístit
bufa	dát vařit vodu	ibufa	skočit do vody

Tabulka 5.1 Slova ilustrující samohlásky jazyka banawa v přízvučných slabikách (v první slabice každého slova) po **t** a po **b** a v nepřízvučných slabikách (v druhé slabice) po **f** a po **b**.

Zvažme, jak můžete použít údaje o frekvencích formantů. Tyto zásady se budou týkat jakékoliv sady frekvencí formantů. Nejdříve byste si měli ověřit, zda jsou měření spolehlivá. Když pracujete se svými vlastními daty, můžete vše změřit dvakrát. Při rozhodování o umístění formantů však můžete udělat chybu a když se podruhé podíváte na spektrogram, můžete opět udělat tu samou chybu. Ideální by bylo, kdyby si spektrogramy prohlédli dva lidé a nezávisle na sobě provedli měření.

Data můžete zkontrolovat i jiným způsobem. Můžete se podívat na slova, která jste nahráli dvakrát a tento způsob můžete použít i při kontrole práce ostatních. V případě jazyka banawa můžeme využít toho, že každé slovo bylo zopakováno a vyznačit (plot) do grafu měření první výpovědi proti té druhé. Na obrázku 5.16 jsou zobrazeny výsledky.



Obrázek 5.16 Korelace mezi dvěma měřeními každého F1 a F2, jedno je měření první realizace každého slova a další je měření druhé realizace toho samého slova.

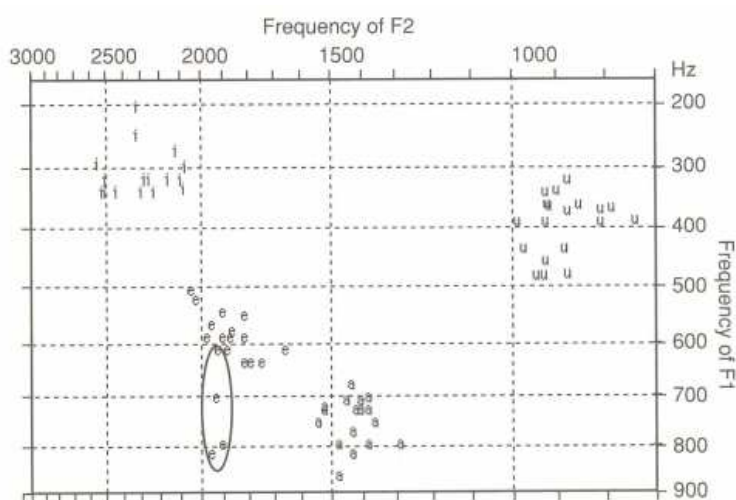
Z grafů na obrázku 5.16 lze vidět, že formanty první realizace (token)⁵¹ každé přízvučné samohlásky daného slova jsou víceméně stejné jako u druhé realizace toho samého slova tím samým mluvčím. Jak mluvčí tak i měření byla shodná. Vyšší hodnoty F1 jsou rozdílné, ale hodnoty F2 ve dvou opakováních toho samého slova jsou velmi podobné. Rozdíly F1 v nepřízvučných samohláskách jsou výrazně větší, což ukazuje, že byly buď udělány chyby při měření nebo, že se nepřízvučné samohlásky více liší ve výšce. Rozdíly hodnot F2 nepřízvučných samohlásek jsou stejné jako u přízvučných samohlásek. Při ověřování těchto

⁵⁰ Je možné též použít výrazy jako „tabulka“ nebo „schéma“. V češtině, kde je poměrně jednoduchý systém samohlásek, používáme výraz „vokalický trojúhelník“. V angličtině, která má složitější systém, se používá výraz „vokalický čtyřúhelník“.

⁵¹ „Token“ lze v podstatě chápat jako hlásku, konkrétní realizaci daného fonému.

dat jsem se díval na všechny případy, kdy se jakýkoliv pár samohlásek lišil o více než 50 Hz. Nenašel jsem nic, co by se mohlo vyřadit jako chyba měření, tudíž jsem akceptoval fakt, že data ukazují pouze na nestálost mluvčího.

I když jsou data správně změřena, tak to neznamená, že nejsou problémy s mluvčími. Když se podíváte na formanty skupiny lidí, měli byste si ověřit, jestli se nějaký mluvčí neliší od ostatních. Na obrázku 5.17 je znázorněn diagram frekvencí prvních dvou formantů přízvučných samohlásek, tak jak je vyslovilo všech pět mluvčích. Elipsa obklopuje čtyři přízvučné **e** samohlásky mluvčího číslo 4. Hodnoty prvního formantu jeho **e** samohlásky se v průměru liší od **e** samohlásek ostatních mluvčích. (Statistická analýza prokázala, že se výrazně odlišují.) Je zřejmé, že mluvčí vyslovuje samohlásku nezvyklým způsobem, s vyšší F1 (jako více otevřenou samohlásku) než ostatní mluvčí. Jeho ostatní samohlásky se neliší od ostatních mluvčích, a proto nemůže být odlišnost této samohlásky připisována nějakému anatomickému faktoru, jako je velmi malé hlasové ústrojí, což by způsobilo vyšší frekvenci všech formantů. Jestliže narazíte na mluvčího, který vyslovuje slovo výrazně odlišně, měli byste při tvorbě diagramu samohlásek daného jazyka tuto část dat vynechat. V poznámce však uveďte, že jsou mluvčí, kteří se odchyľují od obecného vzoru.



Obrázek 5.17 Samohlásky pěti mluvčích jazyka banawa. Elipsa obklopuje čtyři samohlásky jednoho mluvčího, který netypicky vyslovuje /e/ samohlásku.

Stupnice na obrázku 5.17 jsou uspořádané tak, aby z hlediska fonetiky podávaly co nejvíce informací o samohláskách. Formant 1 je na ordinále (vertikální ose) s hodnotami narůstajícími dolu a formant 2 je na abscise (horizontální ose) s rostoucími hodnotami zprava doleva. Tento způsob zaznamenání uspořádá samohlásky podobným způsobem jako

vokalický čtyřúhelník IPA (IPA vowel chart).⁵² Frekvence formantů jsou uspořádány v souladu s Barkovou stupnicí (Bark scale),⁵³ měřítkem sluchové podobnosti tak, že vzdálenost mezi jakýmkoliv dvěma samohláskami reflektuje, jak jsou od sebe hlásky vzdálené. Většina energie samohlásky je obvykle v prvním formantu, a proto je měřítko tohoto formantu více roztažené než u druhého formantu.

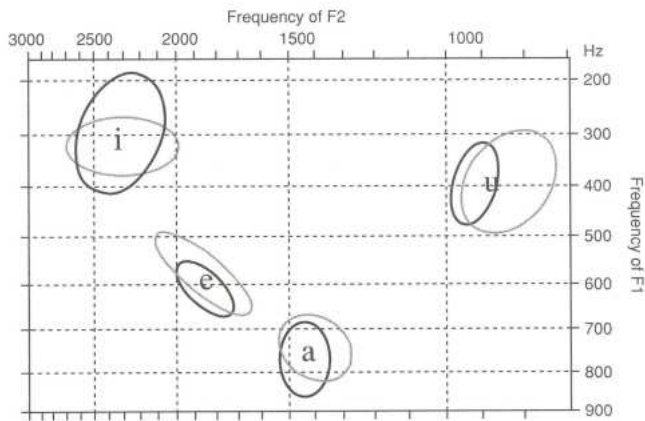
Obrázek 5.17 obsahuje příliš mnoho bodů na to, aby podal dobré informace o kvalitě samohlásek jazyka banawa a nerozlišuje samohlásky po **t** od těch po **b**. Tyto dvě samohlásky musí být odděleny a je potřeba nějakým způsobem zjistit průměr dat. Náležitě grafy můžete vytvořit použitím programu jako je UCLA Plot Formants program, který je dostupný na <http://www.linguistics.ucla.edu/faciliti/sales/software.htm>. Tento program vypočítá pro každou samohlásku průměr (mean) a směrodatnou odchylku (standard deviation) prvních dvou formantů. Okolo skupiny samohlásek vyznačí elipsu s poloměry dvou směrodatných odchylek. Podle běžné statistické hypotézy se dá předpokládat, že 95 % populace, z níž je vybrán tento vzorek mluvčích, bude mít při tvoření samohlásek, jako jsou tyto, hodnoty formantů uvnitř elipsy. Použil jsem tento program pro vyznačení sady elips okolo každé samohlásky po **t** a další sady po **b**, tak jako to je zobrazeno na obrázku 5.18. Protože víme o směrodatných odchylkách, můžeme říci, že přibližně 95% mužských mluvčích jazyka banawa bude tvořit samohlásky, jejichž hodnoty formantů jsou uvnitř elips zobrazených na obrázku 5.18. Jazykem banawa mluví pouze asi 30 mužských mluvčích, takže každá z těchto elips pravděpodobně platí pro všechny kromě jednoho nebo dva z nich; a už jsme našli jednoho mluvčího, který vyslovuje jednu ze samohlásek odlišně.

Na obrázku 5.18 je dobře znázorněna kvalita samohlásek v přízvučných samohláskách jazyka banawa. Dokládá, že (tak jako ve většině jazyků) mají samohlásky po **b** většinou nižší hodnoty F2. Také ukazuje, že **u** tohoto jazyka není tak vysoké jako **i** a je možné ho interpretovat jako **o**. (Skupina lingvistů zabývající se jazykem banawa užívá **u** a jiná **o**.) Pomocí frekvenčního pole formantů (formant chart),⁵⁴ jako je na obrázku 5.18 se můžete rozhodnout, které IPA symboly máte vybrat pro jazyk, na kterém pracujete.

⁵² Pro „IPA vowel chart“ zřejmě neexistuje žádný zažitý název, navrhovala bych použít výraz „vokalický čtyřúhelník IPA“, popřípadě „IPA diagram samohlásek“.

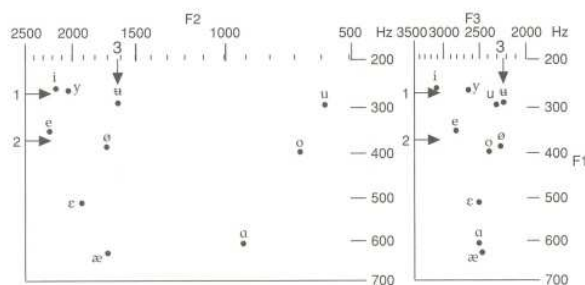
⁵³ Barkova popřípadě Barková stupnice. Pro zobrazení frekvencí formantů používáme Barkovou nebo melovou stupnici, které objektivizují vnímané intervaly. Rozdíly ve výšce totiž nevnímáme jako lineární, nad 800 Hz vnímáme vztah mezi výškou a frekvencí logaritmicky, což nereflektují údaje v Herzích. Cit.: Skarnitzl, Radek: *Psychoakustika* [pdf online]. [cit. 2008-03-16]. Dostupné na <<http://fu.ff.cuni.cz/vyuka/akustika/psychoakustika.pdf>>

⁵⁴ Výraz „frekvenční pole formantů“ užívá Palková ve *Fonetice a fonologii češtiny* (1994), „formant chart“ by také bylo možno nazvat jako „diagram formantů“.



Obrázek 5.18 Grafická znázornění formantů samohlásek jazyka banawa. Elipsy vyznačené tlustou čarou ukazují oblastí zahrnující všechny body mezi dvěma směrodatnými odchylkami průměru samohlásek v přízvučných slabikách po **t**. Slabší elipsy ukazují to samé rozmezí samohlásek v přízvučných slabikách po **b**. Symboly samohlásek jsou umístěny v místě úhrnného průměru (grand mean) každé samohlásky bez ohledu na kontext.

Graf s hodnotami F1 a F2 na jednotlivých osách dobře popisuje samohlásky jazyka banawa a mnoha ostatních jazyků. Tento typ zobrazení však není vhodný pro jazyky, jejichž samohlásky jsou rozlišeny pomocí labializace. V těchto případech je nutné brát v úvahu i F3. Akustické dimenze reprezentované F1, F2 a F3 přímo neodpovídají sluchově-artikulačním dimenzím (auditory-acoustic dimensions) výšky samohlásky, zadnosti samohlásky (vowel backness) a labializaci. Na dvojdimenzionálním grafu reprezentuje F1 většinou výšku samohlásky, F2 však charakterizuje jak její zadnost, tak labializaci. Představuje to problém pro fonetiky popisující samohlásky jazyků jako jsou francouzština, němčina, švédština a dánština, které mají přední labializované samohlásky. Tyto samohlásky lze nejlépe zobrazit tak, že vyneseme jak F1 proti F3, tak i F1 proti F2.

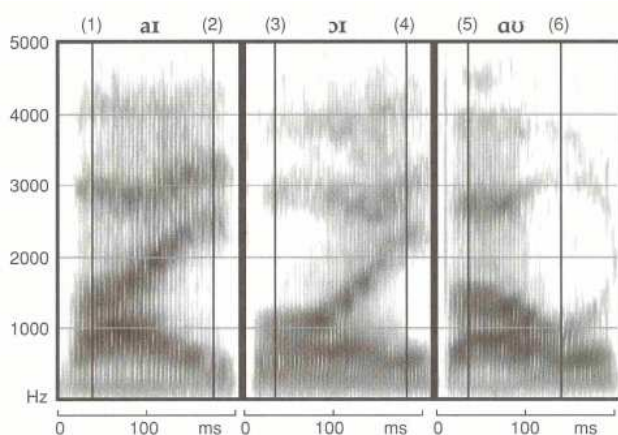


Obrázek 5.19 Grafy formantů švédských dlouhých samohlásek (data z Fanta 1973).⁵⁵ F1-F2 je nalevo a F1-F3 napravo. Místa vyznačená očíslovanými šipkami jsou diskutována v textu.

⁵⁵ V originálním textu není uvedeno, ze které práce G. Fanta byla tato data použita.

Obrázek 5.19 představuje jak graf F1-F2, tak i graf F1-F3 švédských dlouhých samohlásek. Jde o průměr dat 24 studentů, tak jak je publikoval švédský fonetik Gunnar Fant. Vidíte, že pomocí F3 lze rozlišit přední vysoké samohlásky **i** a **y** (šipka 1 na obrázku 5.19), které mají velmi podobné hodnoty F1 a F2. Střední vysoké přední samohlásky **e** a **ø** (šipka 2) jsou rozlišeny pomocí frekvencí F2, jsou však také rozlišeny pomocí F3. Pro vysokou zadní samohlásku **u** (šipka 3) je charakteristická nejnižší frekvence F3 z těchto samohlásek. Tato samohláska má podobnou pozici na grafu F1-F2 jako velarizovaná (retracted) přední samohláska,⁵⁶ jako je anglické **ɪ**. Kdybychom jednoduše znázornili F1 proti F2, nebyli bychom schopni říci, čím se liší od anglického **ɪ**. Vzhledem k tomu, že má velmi nízkou frekvenci F3, víme, že zní velmi odlišně.

Ve spojení se zaznamenáváním samohlásek do grafů samohlásek musíme uvážit, jak na frekvenčním poli formantů reprezentovat diftongy. Na začátku kapitoly jsem tvrdil, že diftongy by se měly měřit blízko začátku a konce samohlásek, v místech neovlivněných tranzientem souhlásky. Někdy je to však problematické. Vezměme v úvahu tři německé diftongy v prvních slabikách slov **vartən**, *weiten*, „rozšířit“; **bortə**, *Beute*, „kořist“; **bautən**, *bauten*, „stavět“. Tyto slabiky jsou zobrazeny na obrázku 5.20. Někdy není snadné určit vhodné body pro měření. Moje návrhy jsou vyznačeny kolmými úsečkami (1)–(6).



Obrázek 5.20 Spektrogramy prvních slabik německých slov **vartən**, *weiten*, „rozšířit“; **bortə**, *Beute*, „kořist“; **bautən**, *bauten*, „stavět“. Očíslované kolmé úsečky označují body měření, o kterých se diskutuje v textu.

První slabika ve slově **vartən**, *weiten*, „rozšířit“ má první souhlásku **v**, která obvykle snižuje všechny formanty (ovšem F3 v tomto případě není příliš ovlivněn). Počáteční bod pro měření

⁵⁶ Velarizovaná přední samohláska se tvoří posunutím hřbetu jazyka k měkkému patru (Řeřicha, 1998).

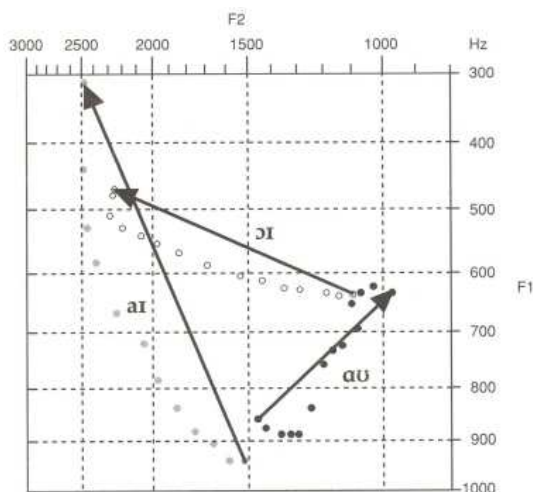
jsem vybral tam, kde je slabý vliv počátečního **v**, kde se stabilizuje F2 a F1 se přibližuje svému maximu. Kolmá úsečka (2) ukazuje místo, kde si myslím, že končí diftong. Je však sporné, zda bych ho neměl umístit dříve, kdy je F2 lehce vyšší. Nejde o jednoduché rozhodnutí, protože zahrnuje souhrn artikulačně akustických vztahů. Při tvoření vysoké přední samohlásky se jazyk pohybuje nahoru dopředu, tudíž se dostane do místa, kde je rezonance přední dutiny⁵⁷ (indikátor toho, jak hodně je samohláska přední) spojena spíše s F3 než s F2. Na diagramu samohlásek se však pro ukázání předozadních dimenzí používá pouze F2, a tak přestože nejde o konečný bod samohlásky, měli bychom možná jako závěrečný bod měření použít maximum F2. Tak jako u všech problémů s měřením, se musíte rozhodnout, stanovit postupy a být jednotný.

Druhá samohláska **ɔɪ**, jako je ve slově **boɪtə**, *Beute*, „kořist“, má poměrně konstantní část na počátku a závěr, který je podobný prvnímu z těchto tří diftongů, F1 však tak neklesá. Body měření jsem vyznačil kolmými úsečkami (3) a (4). Třetí samohláska **ɔʊ**, jako je ve slově **bautən**, *bauten*, „stavět“, představuje další problémy. Není obtížné vyznačit počáteční kolmou úsečku (5), je však problém stanovit závěrečný bod diftongu. Kolmou úsečku (6) jsem umístil v bodě, kde jsou F1 a F2 na minimu, což jsem vybral jako svoje kritérium pro konec diftongu. Po tomto bodu však F1 lehce a F2 značně stoupají. Zvýšení trvá delší dobu, než je běžné pro interval tranzientu souhlásky. Bylo by docela rozumné tyto změny formantů považovat za součást samohlásky a zaznamenat je.

Diftongy lze prezentovat na diagramu formantů. Nejjednodušší je vyznačit počáteční bod a pak narýsovat šipku směřující na konec. Tento postup se používá v učebnicích, je však vynecháno hodně informací. Jak bylo již řečeno, samohláska **ɔɪ** ve druhém slově začíná poměrně nedynamicky a třetí samohláska **ɔʊ**, je složitý diftong, který zahrnuje více než pohyb z místa na místo. Dokonce první samohláska, **ai**, se nepohybuje od počátku do konce stejnoměrně. Jedním ze způsobů, jak zobrazit detaily pohybu diftongu, je vyznačit během samohlásky hodnoty F1 a F2 v intervalech 10 ms. (Na obrázku 5.10 jsme viděli, že systém analýzy může na spektrogramu zobrazit formanty v intervalech 10 ms.) Místo (nebo kromě) počátečních a závěrečných bodů můžou být do programu pro zobrazení formantů přidány trajektorie formantů německých diftongů. Na obrázku 5.21 jsou zobrazeny výsledky reprezentace těchto samohlásek jak šipkami, tak i trajektoriemi formantů. Jak bylo vidět již na spektrogramu, v závěru diftongu **ɔʊ** stoupá F2, toto stoupání je na diagramu vyznačeno třemi

⁵⁷ Není zřejmé, co autor považuje za „přední dutinu“. Pravděpodobně jde o přední část ústní dutiny.

body nalevo od hrotu šipky. Vidíme také, že se samohláska **ai** nepohybuje stejnou rychlostí a že na počátku **ai** jsou relativně malé změny F1.



Obrázek 5.21 Frekvenční pole formantů, které představuje německé diftongy v prvních slabikách slov **var̥tən** , *weiten*, „rozšířit“; **bo̥rtə** , *Beute*, „kořist“; **bḁtən**, *bauten*, „stavět“ vyslovené jedním mluvčím.

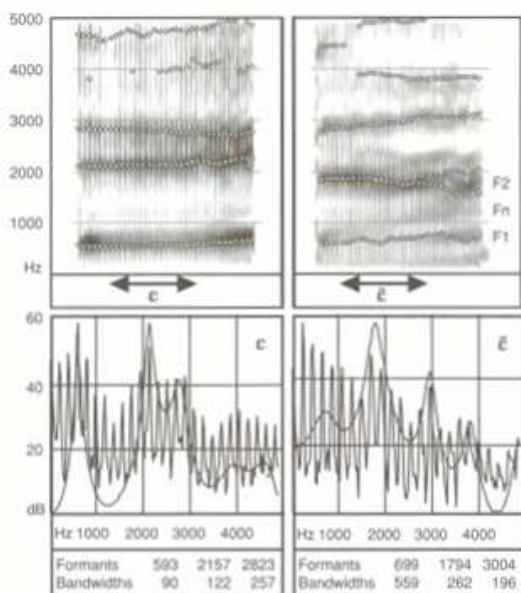
Pro účely této diskuse jsme brali v úvahu pouze jednoho mluvčího, který vyslovil všechny tři německé diftongy. Výsledky zobrazené na obrázku 5.21 jsou tudíž značně omezené. Nevíme, zda se mluvčí shoduje s ostatními, nebo zda tyto slova nemají zvláštní výslovnost. Dokonce ani nevíme, zda se body měření, které jsme vybrali, nacházejí i v ostatních samohláskách. Bez informací o dalších mluvčích a kontextech ostatních samohlásek nemůžeme rozhodnout, které měření nejlépe charakterizuje německé diftongy. Dřív než se rozhodnete, jaké postupy budete používat při zkoumání hlásek určitého jazyka, je dobré si rychle projít dostupná data v co největším množství, abyste věděli, jaké máte stanovit postupy měření. Jen po té, co projdete celou řadu samohlásek, můžete rozhodnout, které postupy měření budou vyhovovat analýze všech dat.

5.4 NAZALIZOVANÉ SAMOHLÁSKY

V úvodní knize jako je tato není možné věnovat více prostoru akustické analýze nazalizovaných samohlásek. Tyto samohlásky jsou charakterizovány částečně tím, co tam je a tím co tam není. Je známo, že první formant nazalizovaných samohlásek má tendenci se vytrácet. V horní části obrázku 5.22 jsou zobrazeny spektrogramy samohlásek ve

francouzských slovech **le**, *laid*, „ošklivý“ a **lẽ**, *lin*, „lněný“. První formant samohlásky vpravo je zřetelně slabší – kdyby nebyl proložený trajektoriemi formantů, měli byste možná problém ho najít.

Mezi spektrogramy orálních a nazálních samohlásek na obrázku 5.22 jsou další rozdíly. V oblasti mezi F1 a F2 je extra energie, kterou jsem označil jako Fn, a která je patrná zejména na konci nazalizované samohlásky a druhý formant nazalizované samohlásky je značně nižší. Ve spodní části obrázku 5.22 jsou zobrazena FFT a LPC spektra. Vrcholy na spektru nalezené LPC analýzou nelze označovat jako vrcholy formantů, protože na nazalizované samohlásky není možné aplikovat LPC algoritmus. Lze je však považovat za dobré změření prominence vrcholů spektra. Jsou užitečnými indikátory toho, co je slyšet. LPC analýza také udává šířku pásma toho, co lze považovat za formanty. Jak vidíte, první dva formanty nazalizovaných samohlásek mají větší šířku pásma – jsou méně ostře definovány a mají nižší amplitudy než odpovídající formanty orálních samohlásek. Zvýšená šířka pásma je částečným znakem nazalizace.



Obrázek 5.22 Průměrné spektrogramy a spektra francouzských orálních a nazálních samohlásek vyjmutých ze slov **le**, *laid*, „ošklivý“ **lẽ**, *lin*, „lněný“, které jsou zprůměrované během intervalu vyznačeného šipkami pod spektrogramy.

Analýza těchto orálních a nazálních samohlásek ukazuje další možnosti dostupné v dobrých systémech analýzy řeči. Můžete vybrat část samohlásky a najít průměrné frekvence formantů. Spektra ve spodní části obrázku 5.22 vznikla zprůměrováním sedmi spekter pořízených každých 10 ms během intervalu vyznačeného šipkami pod spektrogramy. Když je konstantní

stav v prostředí samohlásky, tak dobře zjistíte, že spektrální analýza reflektuje skutečné hodnoty formantů.

5. Závěr

Cílem této práce bylo s pomocí prostudované literatury o teorii překladu přeložit kapitolu *Characterizing Vowels* z původní anglické monografie *Phonetic Data Analysis* od Petera Ladefogeda. Text je opatřen též odborným komentářem, v němž jsem objasnila problematické termíny či potíže vzniklé při překladu. Některé z těchto termínů se v českém fonetickém odborném prostředí neužívají nebo nejsou zažité, tudíž jsem se je pokusila přeložit, popřípadě navrhnout možná řešení, synonymické řady termínů.

Plánovanou součástí práce byla též kapitola věnovaná přehledu české odborné literatury, která se věnuje akustické charakteristice hlásek a popisu metod užívaných v daných publikacích pro charakteristiku hlásek.

Ve všech publikacích se pro charakteristiku samohlásek uvádějí hodnoty F1 a F2, popřípadě hodnoty F3. Tyto hodnoty jsou uváděny v tabulkách a grafech. Ve starších publikacích od B. Hály a M. Romportla jsou uvedeny i metody pro výzkum akustické stránky řeči. Bohužel v dnešní době lze tyto popisy metod brát již jen jako historický doklad dříve užívaných postupů a přístrojového vybavení. V současnosti je pro akustické analýzy užívána vyspělá výpočetní technika.

V *Mluvnici češtiny 1* je uveden pouze graf, který ukazuje rozložení hodnot formantů na formantovém poli. Jednotlivé samohlásky jsou charakterizovány pomocí hodnot F1 a F2. Nejsou zde však uvedeny spektrogramy jednotlivých hlásek, ani zde není uvedeno, jakým způsobem byly hlásky získány a jak byly změřeny hodnoty formantů.

Palková ve své knize *Fonetika a fonologie češtiny* uvádí v tabulkách hodnoty formantů z prací Hály (1941), Romportla (1963) a MČ 1 (1986), jednotlivé samohlásky charakterizuje pouze pomocí spektrogramů. Uvádí pouze, že tyto spektrogramy získala s pomocí počítače a programu J. Liljenkranze, neuvádí však, jakým způsobem byly spektrogramy vytvořeny.

Nejlépe jsou popsány metody, které se v dnešní době používají pro akustickou charakteristiku samohlásek, v práci B. Hedbávné: *Diferenciace kvality vokálů v projevu českých mluvčích*. Autorka v ní uvádí, jaké metody se obecně používají pro akustickou charakteristiku samohlásek. Popisuje jejich výhody a nevýhody a uvádí, které z nich použila ve své práci a jakým způsobem byla data měřena.

Z uvedeného přehledu vyplývá, že v české odborné lingvistické literatuře chybí titul, který se zabývá moderními metodami akustické analýzy. Hedbávná je v podstatě jediným autorem, který uvádí metody pro akustické charakteristiky hlásek. Je to však jen nástin, nikoli zevrubný

popis metod. Navíc jde o diplomovou práci vzniklou ve Fonetickém ústavu Filozofické fakulty Univerzity Karlovy, která nebyla publikována, a tudíž není přístupna široké veřejnosti. Cílem této práce je v návaznosti na výše uvedené prozatím alespoň prostřednictvím překladu zaplnit citelnou mezeru v české odborné literatuře.

Anotace

Jméno a příjmení: Kateřina Braunová

Vysoká škola: Univerzita Palackého v Olomouci

Název katedry a fakulty: Katedra bohemistiky, Filozofická fakulta

Název práce: Characterizing Vowels. Akustická analýza vokálů dle Petera Ladefogeda

Vedoucí práce: PhDr. Petr Pořízka

Počet stran: 53

Počet příloh: 1

Počet znaků: 92 110

Počet titulů použité literatury: 18

Klíčová slova: akustická analýza samohlásek

formanty

spektrogram

spektrum

Cílem této bakalářské práce je překlad kapitoly *Characterizing Vowels* z původní anglické monografie *Phonetic Data Analysis* od Petera Ladefogeda a opatřit text odborným komentářem, v němž jsou vysvětleny problémy terminologické a translatologické povahy. Součástí práce je též kapitola věnující se reflexi akustického zkoumání vokálů v české odborné literatuře. Při překládání jsou řešeny problémy spojené s chybějícími protějšky. Jsou způsobené absencí české odborné literatury.

6. Použité prameny

Čermák, František: *Jazyk a jazykověda. Přehled a slovníky*. První dotisk třetího doplněného vydání. Praha, Univerzita Karlova v Praze 2004.

Dynamický rozsah [online] [cit. 2008-03-16]. Dostupné na

<http://cs.wikipedia.org/wiki/Dynamick%C3%BD_rozsah>

Hála, Bohuslav: *Akustická podstata samohlásek*. Praha, Česká akademie věd a umění 1941.

Hayward, Catherine: *Experimental Phonetics*. Harlow, Pearson Education Limited 2000.

Hedbávná, Barbora: *Diferenciace kvality vokálů v projevu českých mluvčích*. (diplomová práce) Praha 2001.

Knittlová, Dagmar: *Teorie překladu*. Olomouc, Vydavatelství Univerzity Palackého v Olomouci 1995.

Křmová, Marie: Kapitola 6. Akustická a auditivní fonetika. In. *Fonetika*. [online]. 2007 [cit. 2008-04-1]. Dostupné na

<<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch06.html>>

Křmová, Marie: 5.5 Metody poznávání artikulace. In. *Fonetika*. [online]. 2007 [cit. 2008-04-1]. Dostupné na <<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch05s05.html>>

Ladefoged, Peter: *Phonetic Data Analysis*. Malden, MA: Blackwell Publishing 2003.

Mluvnice češtiny 1. Praha, Academia 1986.

Palková, Zdena: *Fonetika a fonologie češtiny*. Praha, Karolinum 1994.

Romportl, Milan: *Základy fonetiky*. Praha, SPN 1973.

Řeřicha, Václav: *Anglicko-český fonetický glosář fonetických a fonologických pojmů*. Olomouc, Univerzita Palackého v Olomouci 1998.

Roach, Peter: *A Little Encyklopaedia of Phonetics* [pdf online]. 2002. [cit. 2008-03-16]

Dostupné na <<http://www.personal.reading.ac.uk/~llsroach/encyc.pdf>>

Skarnitzl, Radek: *Analýza řečového signálu* [pdf online] [cit. 2008-03-16]. Dostupné na <<http://fu.ff.cuni.cz/vyuka/akustika/analyza.pdf>>

Skarnitzl, Radek: *Filtrová teorie produkce řeči* [pdf online] [cit. 2008-03-16]. Dostupné na <http://fu.ff.cuni.cz/vyuka/akustika/source_filter.pdf>

Skarnitzl, Radek: *Psychoakustika* [pdf online] [cit. 2008-03-16]. Dostupné na

<<http://fu.ff.cuni.cz/vyuka/akustika/psychoakustika.pdf>>

Vlček, Vítězslav Vít: *Dvojměrná frekvenční a směrová filtrace – pomocí diskrétní Fourierovy transformace* [pdf online]. 2002. [cit. 2008-03-16]. Dostupné na <<http://home.zcu.cz/~vsoft/dft/10konfvlcek.pdf>>