

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ

ÚSTAV MECHANIKY TĚLES, MECHATRONIKY A  
BIOMECHANIKY

FACULTY OF MECHANICAL ENGINEERING

INSTITUTE OF SOLID MECHANICS, MECHATRONICS AND  
BIOMECHANICS

# ANALÝZA FORMANTŮ ČESKÝCH SAMOHLÁSEK GENEROVANÝCH NAHLAS A ŠEPTEM

ANALYSIS OF CZECH VOWELS TO BE GENERATED ALOUD AND IN A WHISPER

DIPLOMOVÁ PRÁCE

DIPLOMA THESIS

AUTOR PRÁCE

AUTHOR

MICHAL MATUG

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. VOJTĚCH MIŠUN, CSc.

BRNO 2008

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav mechaniky těles, mechatroniky a biomechaniky

Akademický rok: 2007/08

## ZADÁNÍ DIPLOMOVÉ PRÁCE

student(ka): Matug Michal

který/která studuje v **magisterském studijním programu**

obor: **Aplikovaná mechanika (3901T003)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

### **Analýza formantů českých samohlásek generovaných nahlas a šeptem**

v anglickém jazyce:

#### **Analysis of czech vowels to be generated aloud and in a whisper**

Stručná charakteristika problematiky úkolu:

Hlásky se skládají ze souhlásek a samohlásek. Samohlásky jsou generovány spektrálními a modálními vlastnostmi vokálních traktů, různě nastavené podle vyslovované samohlásky. Spektrální vlastnosti kavit vokálního traktu rovněž závisí na okrajových podmínkách kavit na straně hlasivek - otevřeny (šept) nebo uzavřeny (nahlas). Tématem je analýza vlastních frekvencí akustických kavit vokálního traktu při vyslovení jednotlivých samohlásek a při různém typu mluvení - šepot nebo nahlas.

Cíle diplomové práce:

Cílem je analýza formantů českých samohlásek při generování samohlásek nahlas nebo šeptem a při různé základní frekvenci zdrojového hlasu.

Seznam odborné literatury:

I.R.Titze: Principles of voice. Prentice Hall, 1994

Mišun V.: Vibrace a hluk. Skripta VUT Brno, 2005

Mišun V.: Tajemství lidského hlasu. Vydavatelství VUTIUM Brno, připraveno k vydání únor 2008.

Vedoucí diplomové práce: doc. Ing. Vojtěch Mišun, CSc.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2007/08.

V Brně, dne 22.11.2007



prof. Ing. Jindřich Petruška, CSc.  
Ředitel ústavu

doc. RNDr. Miroslav Doupovec, CSc.  
Děkan fakulty

# LICENČNÍ SMLOUVA

## POSKYTOVANÁ K VÝKONU PRÁVA UŽÍT ŠKOLNÍ DÍLO

uzavřená mezi smluvními stranami:

### 1. Pan/paní

Jméno a příjmení: Michal Matug  
Bytem: Gebauerova 901/9, 70200, Ostrava - Přívoz  
Narozen/a (datum a místo): 20.5.1984, Ostrava

(dále jen "autor")

a

### 2. Vysoké učení technické v Brně

Fakulta strojního inženýrství  
se sídlem Technická 2896/2, 61669 FSI VUT v Brně  
jejímž jménem jedná na základě písemného pověření děkanem fakulty:  
doc. Ing. Vojtěch Mišun, CSc.

(dále jen "nabyvatel")

## Článek 1

### Specifikace školního díla

1. Předmětem této smlouvy je vysokoškolská kvalifikační práce (VŠKP):

- disertační práce
- diplomová práce
- bakalářská práce

jiná práce, jejíž druh je specifikován jako .....

(dále jen VŠKP nebo dílo)

Název VŠKP: Analýza formantů českých samohlásek generovaných nahlas  
a šeptem

Vedoucí/školitel VŠKP: doc. Ing. Vojtěch Mišun, CSc.

Ústav: Ústav mechaniky těles, mechatroniky a biomechaniky

Datum obhajoby VŠKP: .....

VŠKP odevzdal autor nabyvateli v:

- tištěné formě - počet exemplářů .....
- elektronické formě - počet exemplářů .....

2. Autor prohlašuje, že vytvořil samostatnou vlastní tvůrčí činností dílo shora popsané a specifikované. Autor dále prohlašuje, že při zpracovávání díla se sám nedostal do rozporu s autorským zákonem a předpisy souvisejícími a že je dílo dílem původním.

3. Dílo je chráněno jako dílo dle autorského zákona v platném znění.

4. Autor potvrzuje, že listinná a elektronická verze díla je identická.

## Článek 2

### Udělení licenčního oprávnění

1. Autor touto smlouvou poskytuje nabyvateli oprávnění (licenci) k výkonu práva uvedené dílo nevýdělečně užít, archivovat a zpřístupnit ke studijním, výukovým a výzkumným účelům včetně pořizování výpisů, opisů a rozmnoženin.
2. Licence je poskytována celosvětově, pro celou dobu trvání autorských a majetkových práv k dílu.
3. Autor souhlasí se zveřejněním díla v databázi přístupné v mezinárodní síti
  - ihned po uzavření této smlouvy
  - 1 rok po uzavření této smlouvy
  - 3 roky po uzavření této smlouvy
  - 5 let po uzavření této smlouvy
  - 10 let po uzavření této smlouvy(z důvodu utajení v něm obsažených informací)
4. Nevýdělečné zveřejňování díla nabyvatelem v souladu s ustanovením § 47b zákona č. 111/1998 Sb., v platném znění, nevyžaduje licenci a nabyvatel je k němu povinen a oprávněn ze zákona.

## Článek 3

### Závěrečná ustanovení

1. Smlouva je sepsána ve třech vyhotoveních s platností originálu, přičemž po jednom vyhotovení obdrží autor a nabyvatel, další vyhotovení je vloženo do VŠKP.
2. Vztahy mezi smluvními stranami vzniklé a neupravené touto smlouvou se řídí autorským zákonem, občanským zákoníkem, vysokoškolským zákonem, zákonem o archivnictví, v platném znění a popř. dalšími právními předpisy.
3. Licenční smlouva byla uzavřena na základě svobodné a pravé vůle smluvních stran, s plným porozuměním jejímu textu i důsledkům, nikoliv v tísní a za nápadně nevýhodných podmínek.
4. Licenční smlouva nabývá platnosti a účinnosti dnem jejího podpisu oběma smluvními stranami.

V Brně dne: .....

.....  
Nabyvatel

.....  
Autor

## Abstrakt

Mezi důležité charakteristiky akustických prostorů vokálního traktu člověka patří spektrální a modální vlastnosti. Jsou přítomny při generování samohlásek a jiných akustických projevů lidské řeči. Rezonanční jevy akustických kavit vokálního traktu můžeme pozorovat ve spektrech lidské řeči, hlavně však při generování samohlásek.

U vokálního traktu se ale ve spektrech samohlásek vyskytuje řada frekvenčních vrcholů, které nemusí nutně být rezonančního původu. Proto je někdy obtížné správně přiřadit frekvenční vrcholy rezonančním vrcholům akustických kavit. To spočívá ve působu akustického buzení vokálních traktů.

Vyslovování samohlásek nahlas a šeptem má odlišné buzení vokálního traktu. Při vytváření samohlásek nahlas je buzen soustavou harmonických složek vztažených k základní frekvenci hlasivek. Při mluvení šeptem je vokální trakt buzen spojitým spektrem generovaným turbulentním prouděním vydechovaného proudu vzduchu přes hlasivkovou štěrbinu.

Formantem nazýváme frekvenci, při které dochází k rezonanci akustického prostoru. Cílem této práce je analýza formantů českých samohlásek generovaných nahlas a šeptem. Experimentální měření těchto formantů bylo provedeno na lidském vokálním traktu pro všechny samohlásky. Dále pak na uměle vytvořených vokálních traktech pro samohlásky A, I. Poté byly modální vlastnosti vokálních kavit pro samohlásky A, I ověřeny metodou konečných prvků za pomoci výpočtového programu ANSYS. V práci byly zkoumány průběhy akustických tlaků pro jednotlivé formanty, vliv velikosti vokálního traktu a vliv správného otevření úst na formanty. Byla také provedena výpočtová simulace harmonického buzení traktu na straně hlasivek.

## **Abstract**

The modal and spectral characteristic belongs among important human acoustic spaces of vocal tract. They occur at generating vowels and other acoustic aspects of human speech. We can observe the resonant phenomena of acoustic cavity of vocal tract in the human speech spectrum, primary however at vowels generation.

However near vocal tract occurs series of frequency tops in the spectrum of vowels, which necessarily may not be resonant origin. That is why sometimes quite difficult assign is right frequency tops to resonant tops of acoustic cavity. It consist in operate of acoustic excitation of vocal tracts.

The pronounced of vowels loudly and in a whisper has different excitation of vocal tract. At generating vowels loudly is excited by scheme of harmonic components outspread to fundamental frequency of glottis. At talking in a whisper is vocal tract excited by continuous spectrum generated by turbulent fluxion of exhaled flatus over glottis.

We give a name "formant" to a frequency, at which happens to resonance of acoustic space. Aim of this work is analysis of Czech vowels formants generated loudly and in a whisper. Experimental metering of these formants was performed on human vocalic tract for all vowels. Further then on artificially created vocalic tracts for vowels A, I. Then were modal characteristics of vocal cavity for vowels A, I, tested by method of final elements with the help of computing program ANSYS. In this work were surveyed courses of acoustic pressures for individual formants, influence sizes vocal tract and influence of correct mouth opening on formants. Also has been effected computational simulation of harmonic excitation on tract by side of glottis.

## **Klíčová slova**

vokální trakt

formant

samohláska

hlas

artikulace

základní frekvence

modální analýza

harmonická analýza

## **Keywords**

vocal tract

formant

vowel

voice

articulation

fundamental frequency

modal analysis

harmonic analysis

## **Bibliografická citace**

MATUG, M. *Analýza formantů českých samohlásek generovaných nahlas a šeptem*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2008. 67 s. Vedoucí diplomové práce doc. Ing. Vojtěch Mišun, CSc.



## **Čestné prohlášení**

Prohlašuji, že jsem diplomovou práci vypracoval samostatně, na základě uvedené použité literatury pod vedením vedoucího diplomové práce.

V Brně 24. května 2008

.....

Michal Matug

## **Poděkování**

Na tomto místě bych rád poděkoval všem, kteří se jakýmkoli způsobem podíleli na vzniku této diplomové práce. A to hlavně vedoucímu diplomové práce doc. Ing. Vojtěchu Mišunovi, CSc. za poskytnuté materiály a především cenné rady a připomínky a Ing. Martinu Vaškovi za spolupráci při experimentálním měření a vyhodnocování výsledků.

Také bych chtěl poděkovat rodičům, kteří se mnou měli trpělivost a porozumění pro dny strávené při tvorbě této práce a podporovali mě i po celou dobu studia.

## Obsah

1	<b>Úvod</b> .....	13
2	<b>Formulace problému</b> .....	14
3	<b>Historie a vývoj syntézy řeči</b> .....	15
	3.1 Přejchod od mechanických k elektrickým syntetizátorům .....	15
	3.2 Vývoj elektrických syntetizátorů .....	17
4	<b>Anatomie vokálních orgánů</b> .....	21
	4.1 Vokální orgány a jejich funkce.....	21
	4.1.1 Ústrojí dýchací (respirační).....	21
	4.1.2 Ústrojí hlasové (fonační).....	22
	4.1.3 Ústrojí modifikační .....	23
	4.2 Aktivní artikulátory .....	24
5	<b>Zdrojový hlas a spektra</b> .....	25
	5.1 Charakteristiky zdrojového hlasu .....	25
	5.2 Vokální trakt .....	26
	5.3 Formanty .....	27
	5.4 Teorie zdroj-filtr pro generování hlasu .....	27
6	<b>Artikulace</b> .....	30
	6.1 Artikulace samohlásek .....	30
	6.2 Artikulace souhlásek .....	32
7	<b>Teorie popisu tvorby hlasu</b> .....	33
	7.1 Myo-elasto aerodynamická teorie .....	33
	7.2 Teorie fonace: bubliny stlačeného vzduchu .....	35
8	<b>Tvorba modelu</b> .....	38
	8.1 Model geometrie .....	38
	8.2 Model materiálu .....	40
	8.3 Model vazeb .....	41
	8.4 Model zatížení .....	42
9	<b>Výsledky výpočtového modelování</b> .....	43
	9.1 Průběhy akustických tlaků .....	43
	9.2 Vliv velikosti vokálního traktu .....	47
	9.3 Vliv polohy otevření úst .....	51
	9.4 Harmonické buzení vokálních traktů .....	52

10	<b>Experimentální analýza formantů</b> .....	<b>54</b>
	10.1 Experimentální určení polohy formantů při mluvení nahlas .....	54
	10.2 Experimentální určení polohy formantů při mluvení šeptem .....	54
	10.3 Analýza formantů .....	54
11	<b>Experimentální měření skutečných vokálních traktů</b> .....	<b>56</b>
	11.1 Samohlásky vyslovované nahlas .....	56
	11.2 Samohlásky vyslovované šeptem .....	58
12	<b>Experimentální měření umělých vokálních traktů</b> .....	<b>60</b>
	12.1 Měření s hlasivkovými segmenty .....	61
13	<b>Závěr</b> .....	<b>63</b>
	Použitý hardware a software .....	65
	Seznam příloh .....	66
	Použitá literatura .....	67

## 1 Úvod

Řeč je jedním z nejdůležitějších dorozumívacích prostředků člověka. Lidé se již téměř 230 let snaží o pochopení její tvorby a umělou produkci. V současnosti existují dvě teorie vzniku hlasu. První teorie definovaná Van den Bergem již v roce 1958 byla nazvána myo-elasto aerodynamickou teorií. Její platnost je doposud všeobecně uznávána odbornou veřejností. Na tuto teorii pak navazovala a stále navazuje převážná část světových odborníků, zabývajících se tvorbou lidského hlasu a především tvorbou zdrojového hlasu, který generují hlasivky. Ukázalo se však, že tato teorie není zdaleka dokonalá. To vedlo k vytvoření nové teorie, nazvané Teorie fonace: bubliny stlačeného vzduchu. Tato teorie byla vymyšlena a je dále vyvíjena na ÚMTMB FSI VUT v Brně.

Přestože je tvorba hlasu biofyzikálním dějem, nebyla dosud sepsána ucelená odborná česká literatura, která by se tímto tématem hlouběji zabývala.

Cílem této práce bylo se podrobněji seznámit s principem tvorby a šíření lidského hlasu a jeho frekvenční analýzou a zejména frekvenčními vrcholy při tvorbě samohlásek – formanty. To je prováděno při mluvení nahlas, kde hlasivky generují periodicky nespojitě spektrum i při mluvení šeptem, kde hlasivky generují šum dávající spojitě frekvenční spektrum. Na experimentální část navazuje část výpočtová. Ta popisuje frekvenční hodnoty jednotlivých formantů pro samohlásky A a I a tvary vlastních kmitů vzduchové kavity ve vokální dutině při různých nastaveních parametrů.

Tato diplomová práce je přípravou pro pokračování v doktorském studiu se zaměřením na vývoj náhradních hlasivek pro generování zdrojového hlasu.

## 2 Formulace problému a cílů jeho řešení

V této práci se zabývám měřením a analýzou formantů všech českých samohlásek – A,E,I,O,U skutečného lidského hlasu měřicí aparaturou PULSE a stejnojmenném softwaru, dále pak modální a harmonickou analýzou vlastností modelů vokální dutiny za pomoci programového systému ANSYS. Tato analýza byla provedena na modelech pro vyslovování samohlásek A,I vyslovované nahlas a šeptem.

Při mluvení nahlas hlasivky generují základní periodický tón, který se ve frekvenčním spektru projevuje jako čarové spektrum. Není proto možné určit přesnou hodnotu formantu. Ten totiž může ležet někde mezi frekvenčními čarami. Tento problém lze zdánlivě vyřešit analýzou při mluvení šeptem, protože hlasivky generují spojité spektrum u kterého je možné téměř přesně odečíst jednotlivé formanty. Při mluvení šeptem je ale mezi hlasivkami štěrbina, která výrazně změní okrajové podmínky, a tím i vlastní frekvence a tvary.

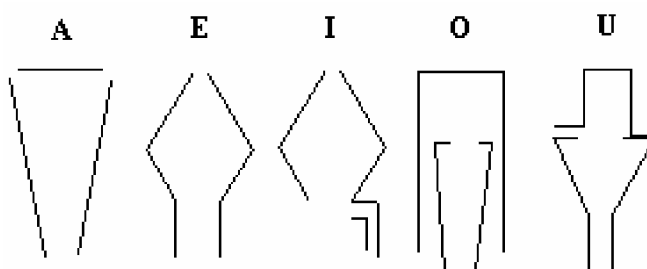
Také počítačové modelování by mohlo být řešením pro určení formantů při mluvení nahlas, kdy namodelováním vokální kavity a předepsáním správných okrajových podmínek můžeme získat téměř přesné hodnoty.

### 3 Historie a vývoj syntézy řeči

Vytvoření umělé řeči bylo snem lidstva už po staletí. Historie syntetizované řeči vznikala od prvního mechanického systému k systémům, které tvoří současné vysoce kvalitní syntetizátory.

#### 3.1 Přejchod od mechanických k elektrickým syntetizátorům

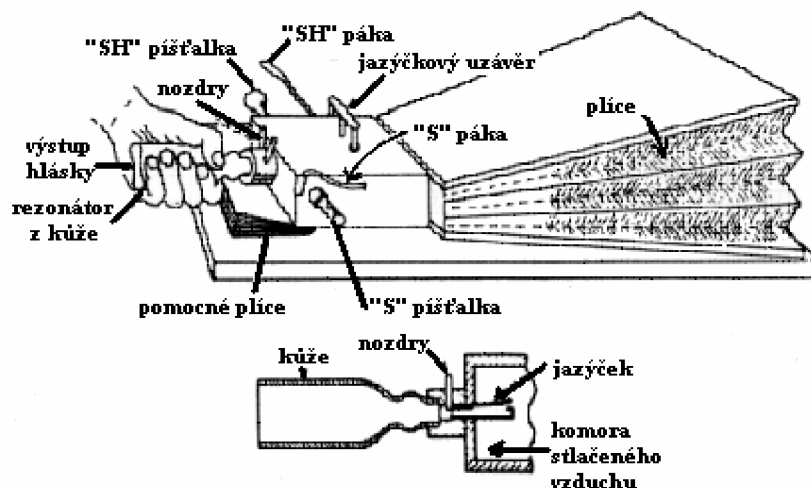
První snaha o umělou produkci lidské řeči byla učiněna už před více než 200 lety. V St. Petersburgu v roce 1779 ruský profesor Christian Kratzenstein vysvětlil fyziologické rozdíly mezi pěti dlouhými samohláskami (A,E,I,O,U) a sestrojil aparát k jejich umělému produkování. Vytvořil akustické rezonátory podobné lidskému vokálnímu traktu a aktivoval je s kmitajícími jazýčky podobně jako je to u hudebních nástrojů. Základní konstrukce rezonátorů je ukázána na obr. 3.1. Zvuk je produkován foukáním do spodní trubice bez jazýčku způsobujícího flétnový zvuk.



**Obr. 3.1 Kratzensteinův rezonátor [8]**

O několik let později ve Vídni v roce 1791 představil Wolfgang von Kempelen svůj akusticko-mechanický mluvicí stroj, který byl schopen produkovat jednotlivé zvuky i nějaké zvukové kombinace. Ve skutečnosti začal Kempelen tvořit svou práci již před Kratzensteinem v roce 1769. Po dvacetiletém výzkumu publikoval knihu, ve které popsal své studie lidské tvorby hlasu a své experimenty s mluvicím strojem. Jeho stroj se skládal z tlakové komory (plíce), kmitajícího jazýčku (hlasivky) a kožené trubice (vokální trakt). Pomocí manipulace tvaru kožené trubice mohl produkovat různé samohlásky. Souhlásky byly simulovány čtyřmi oddělenými částmi řízenými prsty. Jeho studie vedla k teorii, že vokální trakt, dutina mezi hlasivkami a rty jsou hlavní artikulátory. Před Kempelenem byl hrtan obecně považován za

centrum produkce řeči. V polovině 19. století vybudoval Charles Wheatstone slavnou verzi Kempelenova mluvícího stroje, zobrazeného na obr. 3.2.



**Obr. 3.2** Wheatstonova rekonstrukce von Kempelenova mluvícího stroje

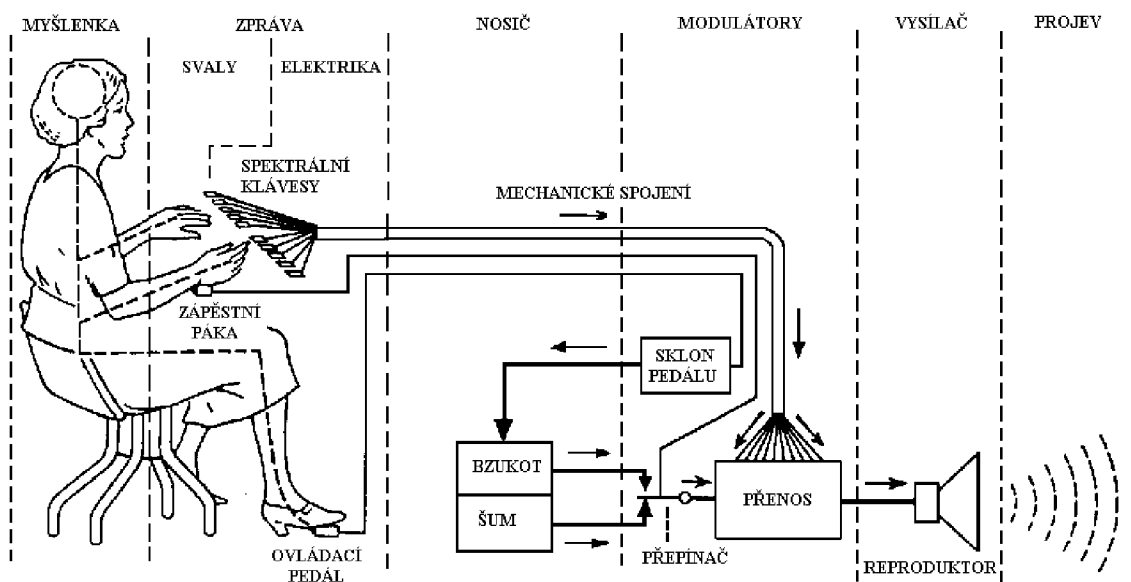
Byl trochu komplikovanější a schopný vytvářet samohlásky a většinu souhlásek a dokonce i zvukové kombinace a slova. Samohlásky byly vytvořeny s kmitajícím jazýčkem a všemi otvory zavřenými. Rezonance byly způsobeny deformujícím se koženým rezonátorem stejně jako u Kempelenova stroje, souhlásky včetně nosovek byly vytvořeny vířivými proudy ve vhodném otvoru bez jazýčku.

Spojení mezi specifickým zvukem samohlásky a geometrií vokálního traktu našel Willis v roce 1838. Objevil přitom, že kvalita samohlásky závisí jen na délce trubky, a ne na jeho průměru. Na konci 19. století Alexander Graham Bell inspirovaný Wheatstonovým mluvícím strojem se svým otcem vybudoval podobný druh stroje. Výzkum a experimenty mechanických a poloelektrických analogových hlasových systémů prováděli i další slavní vědci jako například Herman von Helmholtz.



### 3.2 Vývoj elektrických syntetizátorů

První plně elektrické zařízení syntézy bylo představeno G.W. Stewartem v roce 1922. Syntetizátor měl bzučák jako zdroj a dva sériové kmitavé obvody k modelování akustických rezonancí vokálního traktu. Tento stroj byl schopen generovat jednotlivé statické zvuky samohlásek s dvěma nejnižšími formanty, ale nebyl schopen vytvářet souhlásky nebo spojené projevy. V roce 1932 japonští badatelé Obata a Teshima objevili třetí formant samohlásek. První tři formanty jsou obecně považovány za dobře srozumitelnou řeč. První zařízení považované za hlasový syntetizátor bylo zařízení VODER (Voice Operating Demonstrator) zobrazeno na obr. 3.3. Toto zařízení bylo předvedeno v roce 1939 na světové výstavě v New Yorku. VODER byl inspirován VOCODERem (Voice Coder) vyvinutým v polovině třicátých let. Původní VOCODER byl zařízení na analyzování řeči pomalu se měnícími akustickými parametry, které syntetizátor řídil tak, aby rekonstruoval přiblížení původního řečového signálu. VODER se skládá z páky pro vybrání bzučení nebo zdroje šumu a pedálu pro ovládání základní frekvence. Zdrojový signál postupoval skrz deset pásmově propustných filtrů, jejichž výstupní úrovně byly řízeny prsty. To vyžadovalo značnou schopnost k přehrání věty. Kvalita projevu a srozumitelnost nebyla dobrá, ale byl ukázán potenciál pro produkci umělé řeči.



Obr 3.3 Hlasový syntetizér VODER

Po ukázce VODERu se vědecký svět stále více zajímal o syntézu řeči. To nakonec ukázalo, že srozumitelná řeč může být produkována i uměle.

V roce 1951 Franklin Cooper a jeho spolupracovníci vyvinuli vzor přehrávacího syntetizátoru v Haskinsových laboratořích. To znovu změnilo popsané spektrogramy zvuků v originální a upravené formě.

První formantový syntetizátor PAT (Parametric Artificial Talker) představil Walter Lawrence v roce 1953. PAT se skládal ze tří elektronických formantových rezonátorů zapojených paralelně. Vstupním signálem byl buď bzukot nebo šum. Pohyblivé podložní sklíčko se používalo pro přeměnu namalovaných vzorů do šesti časových funkcí k tomu aby ovládaly tři formantové frekvence vyjadřující amplitudu základní frekvence a amplitudu šumu. Současně Gunnar Fant uvedl první sériový formantový syntetizátor OVE (Orator Verbis Electricis), který se skládal z formantových rezonátorů zapojených do série. O deset let později v roce 1962 Fant a Martony představili zlepšený OVE II syntetizátor, který se skládal ze samostatných součástí, aby modeloval převodní funkci vokálního traktu pro samohlásky, nosovky a určité souhlásky. Poté následovaly OVE III a GLOVE vytvořené v Kungliga Tekniska Hogskolan (KTH) ve Švédsku určené původně pro komerční hlasově-informační systém.

John Holmes uvedl svůj paralelní formantový syntetizátor v roce 1972 po nastudování předešlých syntetizátorů. O rok později uvedl paralelní formantový syntetizátor vyvinutý s JSRU (Joint Speech Research Unit).

První artikulační syntetizátor představil v roce 1958 George Rosen v Massachusetts Institute of Technology. Tento DAVO (Dynamic Analog of the Vocal tract) byl řízen záznamovou páskou z kontrolního signálu vytvořeného ručně. V polovině 60. let byly provedeny první experimenty s LPC (Linear Predictive Coding).

První plný systém převádějící anglický text na řeč byl vyvinut v Electrotechnical Laboratory v Japonsku v roce 1968 japonským vědcem Noriko Umedou a jeho spolupracovníky. Byl založen na artikulačním modelu a zahrnoval syntaktickou modulovou analýzu s důmyslnými objevy.

V roce 1979 demonstrovali Allen, Hunnicutt a Klatt systém MITalk, převádějící text na řeč, vyvinutý v laboratoři M.I.T. O dva roky později Dennis Klatt uvedl slavný Klattalk system. Technologie použitá v systémech MITalk a Klattalk tvoří základy mnoha novějších systémů jako například DECTalk a Prose-2000.

První čtení pomocí optického snímače uvedl Kurzweil v roce 1976. Kurzweilovy čtecí stroje byly schopny číst docela dobře mnohoznakový psaný text. Nicméně byl tento systém pro obyčejné zákazníky příliš drahý (cena přesahovala 30 000 dolarů), ale byly použity v knihovnách a servisních střediscích pro zrakově postižené lidi.

Na konci 70. a začátkem 80. let bylo komerčně představeno mnoho syntetizátorů řeči. První integrovaný obvod pro syntézu řeči byl pravděpodobně Votrax chip, který se skládal ze sériového formantového syntetizátoru a jednoduchého dolnopropustného vyhlazovacího obvodu. V roce 1978 Richard Gagnon představil systém Votrax. V roce 1980 uvedl Texas Instruments lineární předpověďní kódování (Linear Prediction Coding - LPC) založené na Speak-n-Spell syntetizátoru vycházejícím z lineární předpověďní čipové syntézy (TMS 5100). Tento systém byl používán pro elektronickou pomoc čtení u dětí a dostal značné pozornosti. V roce 1982 Street Electronics uvedli Echo low-cost diphone syntetizátor, který byl založen na novější verzi stejného čipu jako Speak-n-Spell (TMS 5220). Zároveň Speech Plus Inc. představili systém převádějící text na řeč Prose-2000. O rok později byly představeny komerční verze slavných syntetizátorů DECTalk a Infovox SA-101.

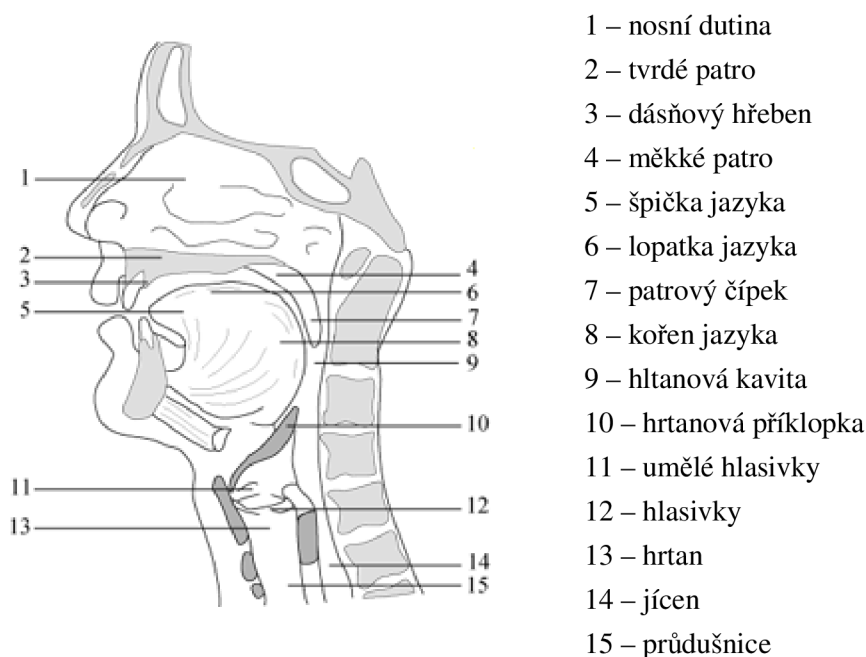
Moderní technologie syntézy řeči vyžadují dost komplikované a důmyslné metodiky a algoritmy.

Jednou z nedávno vzniklých metod syntézy řeči je metoda skrytých markovovských modelů (Hidden Markov Models) dále uváděno HMM. Tato metoda se používala asi dvacet let. Tato metoda je v současné době nejrozšířenější metodou rozpoznávání řeči. Používá se k rozpoznávání jak izolovaných slov tak i souvislé promluvy. Toto rozpoznávání může být závislé i nezávislé na mluvčím. Vzory pro rozpoznávání mohou být jak izolovaná slova tak i menší řečové úseky, jako jsou například hlásky. Modely vzorů vyjadřují statistický popis rozložení v prostoru parametrů. K popisu jejich parametrů je nutné dostatečné množství trénovacích dat. Metoda HMM vychází z představy, že řečový signál je tvořen posloupností určitých stacionárních úseků. Za tyto úseky je možné považovat například části souhlásek a samohlásek. Tyto stacionární úseky jsou navzájem propojeny přechodovými oblastmi. Tuto představu je možné simulovat stavy markovova modelu, které jsou popsány spojitou směsí gausovských rozložení hustot pravděpodobnosti. Doby trvání stavu a přechody mezi stavy jsou charakterizovány pravděpodobnostmi setrvání ve stavu a pravděpodobnostmi přechodů mezi stavy.

V posledních deseti letech byly v syntéze řeči aplikovány neuronové sítě. Jejich poslední výsledky slibovaly docela dobrý potenciál pro použití. Stejně jako HMM tak i neuronové sítě jsou dodnes úspěšně používané metody k rozpoznávání hlasu.

## 4 Anatomie vokálních orgánů

Lidský hlas je tvořen v dýchací soustavě zobrazené na obr. 4.1. V souvislosti s tvorbou hlasu mluvíme o dýchacím ústrojí jako o hlasovém akustickém systému.



**Obr. 4.1 Dýchací soustrava [9]**

### 4.1 Vokální orgány a jejich funkce

Vokální orgány tvoří tři skupiny: ústrojí dýchací, ústrojí hlasové, ústrojí modifikační.

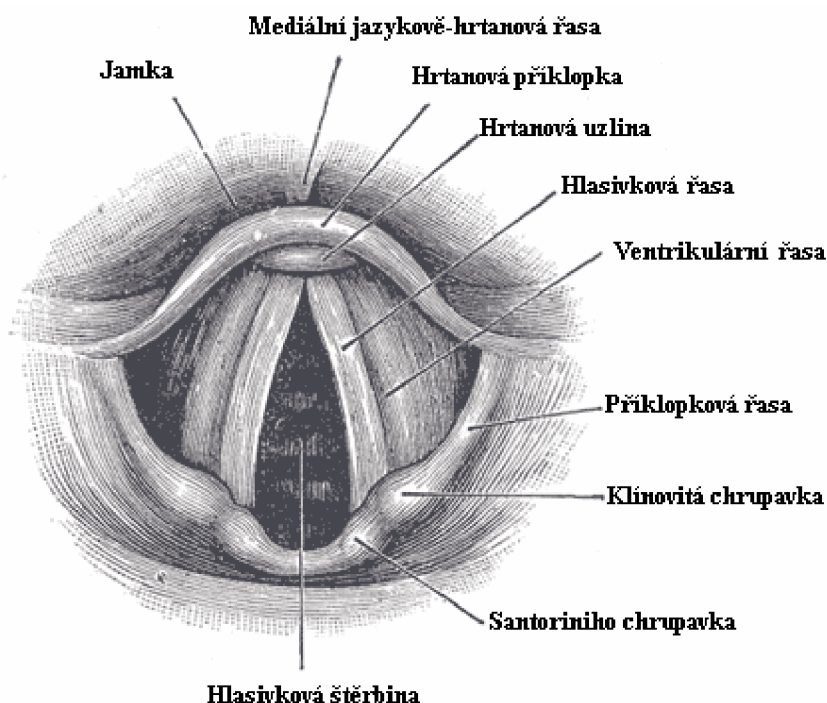
#### 4.1.1 Ústrojí dýchací (respirační)

Dýchací ústrojí slouží primárně k základní životní funkci, kterou je fyziologické dýchání. Při řeči se fyziologické dýchání mění. Poměr vdechu a výdechu je při běžném dýchání asi 2 : 3 (16-20x za minutu), kdy se najednou v klidu vdechne a vydechne jen asi 0,5 litru vzduchu. Při řeči se vdech nápadně zkracuje a výdech se prodlužuje 1 : 7 až 1 : 12 a zvětšuje se objem vdechnutého vzduchu na zhruba 1,5 litru. Množství nadechnutého vzduchu i rytmus dýchání se při řeči řídí zčásti i vědomě, zatímco fyziologické dýchání je reflexivní, i když jej dokážeme ovládat. Při fyziologickém dýchání se doporučuje nadechovat nosem, při řeči se vdech provádí nosem i ústy, protože je nutno v malém čase naplnit plíce maximem vzduchu. Zkušený člověk zvládá dýchání při mluvení bez nežádoucích zvukových efektů.

#### 4.1.2 Ústrojí hlasové (fonační)

Hlasové ústrojí zobrazené na obr. 4.2 je uloženo v hrtanu. Je typicky lidským orgánem a jeho funkcí je vytvářet zdrojový hlas, jehož dalšími úpravami vzniká vlastní řeč.

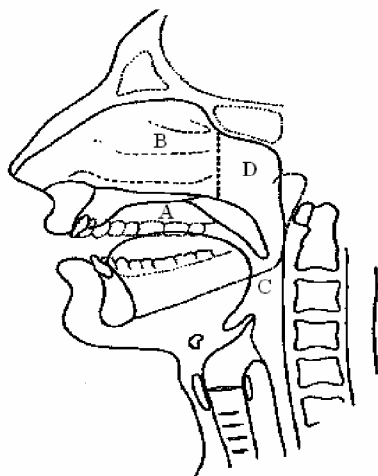
Základem hlasového ústrojí jsou dvě hlasivkové řasy pokryté sliznicí, tzv. hlasivky (chordae vocales). Ty jsou napjaty mezi chrupavkou štítnou (vpředu), kde se navzájem stále dotýkají, a chrupavkami hlasivkovými. Hlasivkové chrupavky jsou umístěny vzadu na rozšířené části prstěncové chrupavky a jsou částečně pohyblivé. Mohou se různým způsobem otáčet, sbližovat a měnit sklon. Jejich kloubní spojení s prstěncovou chrupavkou je mimořádně složité. Také způsob jejich nervového a svalového ovládní je komplikovaný a dosud nebyl dostatečně odborně popsán. Hlasivky jsou složeny z vazivové části, podél níž jsou hlasivkové svaly, podílející se na napínání hlasivek spolu s posuny chrupavek tvořících hrtan a hlasivkových chrupavek. Mezi hlasivkami se vytváří hlasivková štěrbina (glottis). Má tvar trojúhelníku s vrcholem vpředu v klenbě chrupavky štítné a její dvě třetiny jsou ohraničeny blanitou částí hlasivek. Zadní třetina je tvořena okrajem hlasivkových chrupavek, pokrytým podobně jako hlasové řasy jemnou, a tudíž snadno zranitelnou sliznicí.



Obr. 4.2 Hlasové ústrojí [10]

### 4.1.3 Ústrojí modifikační

Modifikační ústrojí je uloženo nad hrtanem. Skládá se z dutiny hrdelní, dutiny nosní a dutiny ústní. Tyto rezonanční dutiny jsou zobrazeny na obr. 4.3



**Obr. 4.3 Rezonanční dutiny**

- A - dutina ústní - orální
- B - dutina nosní - nazální
- C - dutina hrdelní - laryngální
- D - měkké patro

#### Dutina hrdelní (laryngální)

Rozkládá se bezprostředně nad hlasivkami a končí v místech, kde se jazyk při artikulaci nejvíce přiblíží patru. Při řeči se hrdelní dutina účastní jako rezonační prostor. Její objem je proměnlivý pohybem kořene jazyka, pozicí jazyka v ústech a činností hrdelních svalů. Tyto objemové změny hrdelní dutiny se uplatňují zejména při tvoření samohlásek.

#### Dutina nosní (nazální)

Je také rezonančním prostorem, který se však využívá jen při výslovnosti nosových hlásek (nazál). V češtině to jsou zejména souhlásky M a N. Při většině hlásek je průchod do nosní dutiny uzavřen tím, že se měkké patro (velum) zvedá a přitiskuje k zadní stěně ústní dutiny.

Velum je orgán, jehož primární funkcí je zavírat vstup do nosní dutiny při polykání potravy. Při dýchání je spuštěno a vzduch mezi plícemi a chřípím nosu volně prochází. Při artikulaci je v měkkém patru vždy určité napětí, a také patrohrtanový závěr není u všech hlásek stejně silný (nejslabší je u samohlásky A).

V některých jazycích se vytvořily i nazální samohlásky (nosovky), a to například v polštině nebo francouzštině, u nichž je nosní rezonance uplatněna v celém průběhu trvání hlásky, nebo jen částečně. České samohlásky mohou být pouze nazalizovány, tzn. při méně pečlivé výslovnosti mohou v sousedství nazál získat nosní přídech.

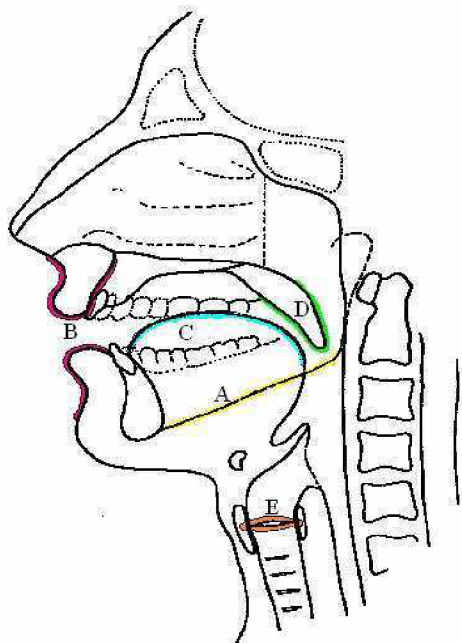
### Dutina ústní (orální)

Vpředu je ohraničena rty, vzadu přechází do dutiny hrdelní. Ústní dutina se při řeči uplatňuje jako rezonanční prostor, který je přítomen u realizace každé hlásky. Objem tohoto prostoru je proměnlivý a pozměňuje se jak pohybem jazyka, tak pohybem rtů a čelistí. Spolu s proměnami hrdelní dutiny je tato změna využívána pro vytváření samohlásek. Spolupůsobí však i při tvoření souhlásek. Přeprada v ústní dutině, kterou se souhláska tvoří, člení prostor úst a důsledkem je pak charakteristická výška šumu.

Význam ústní dutiny pro výslovnost je především v tom, že v ní probíhá diferenciacce většiny zvuků řeči. Jsou v ní, nebo na jejím okraji uloženy aktivní i pasivní vokální orgány. Aktivní jsou pohyblivé, jako pasivní orgány označujeme místa, na nichž se vytvářejí překážky nutné pro tvoření souhlásek.

#### 4.2 Aktivní artikulátory

Aktivní vokální artikulátory uvedené na obr. 4.4 mají funkci změny tvaru vokálního traktu. Je to především dolní čelist (mandibula) jejímž pohybem se mění velikost čelistního úhlu. Mimo ústní dutinu jsou aktivním orgánem hlasivky (chordae vocales), uvnitř ústní dutiny velum (měkké patro, zadní patro), zejména jeho nejzadnější část čípek (uvula), dále nejdůležitější artikulační orgán jazyk (lingua) a rty (labia). Měkké patro a rty jsou při tvoření souhlásek orgány pasivní.



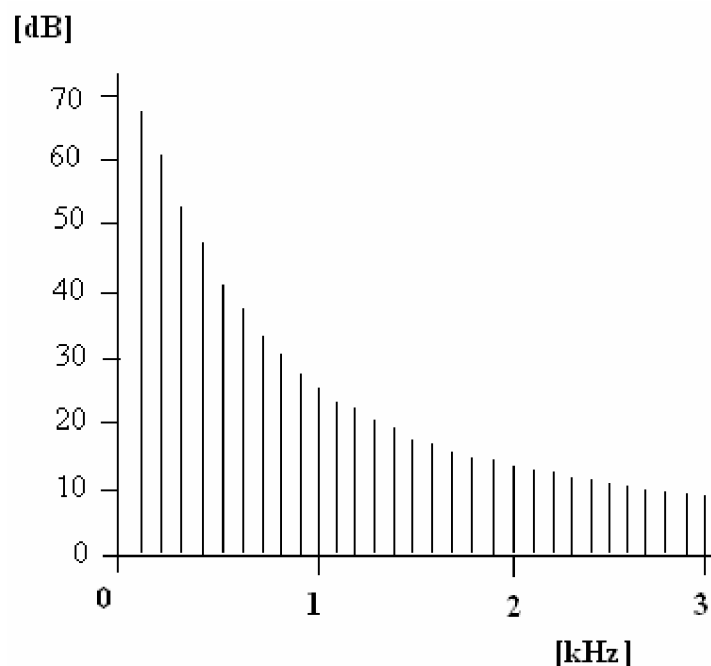
**Obr. 4.4 Aktivní vokální artikulátory**

- A - dolní čelist (mandibula)
- B - rty (labia)
- C - jazyk (lingua)
- D - měkké patro (velum)
- E - hlasivky (chordae vocales)



## 5 Zdrojový hlas

Zdrojový hlas je zvuk, který se generuje na výstupu z hlasivek. Skládá ze základního tónu o frekvenci  $F_0$ , která představuje základní frekvenci zdrojového hlasu a určitého počtu vyšších harmonických složek, které se postupně snižují o 6 dB na oktávu, viz obr 5.1.



Obr. 5.1 Zdrojový hlas

### 5.1 Charakteristiky zdrojového hlasu

- **výška** – fyzikálně odpovídá základní frekvenci kmitání hlasivek  $F_0$
- **znělost** (hlasivky vibrují)
- **neznělost** (hlasivky nevibrují)
- **hlasitost** – fyzikálně odpovídá intenzitě,
- **kvalita** – fyzikálně odpovídá poměru harmonických tónů ve hlase k neharmonickým složkám
- **barva** – je dána hmotou hlasivek, jejich délkou a tvarem rezonančních dutin,
- **flexibilita** – odpovídá komplexnímu vjemu předchozích kvalit dohromady.

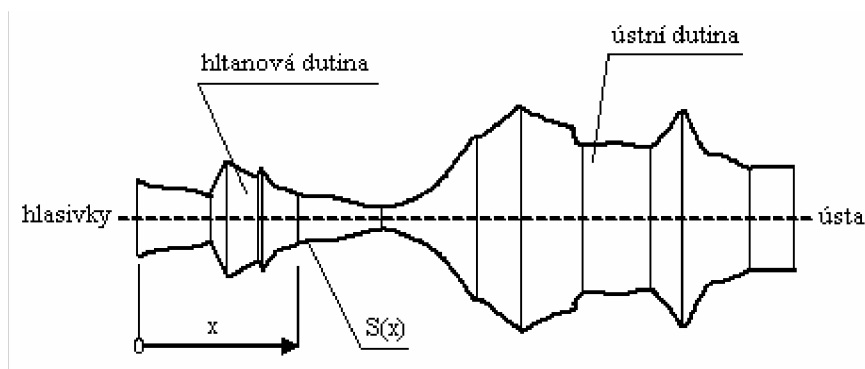
Základní frekvence zdrojového hlasu se u dětí, žen a mužů výrazně liší, jak je ukázáno v tabulce 5.1. To je dáno rozměry vokálního traktu. S jeho zvětšováním se základní frekvence snižují.

$F_0$ [Hz]			
	$F_0$ – střední	$F_0$ - minimum	$F_0$ – maximum
děti	<b>330</b>	200	500
ženy	<b>220</b>	150	350
muži	<b>110</b>	80	200

**Tab. 5.1** Základní frekvence  $F_0$  u dětí, žen a mužů

## 5.2 Vokální trakt

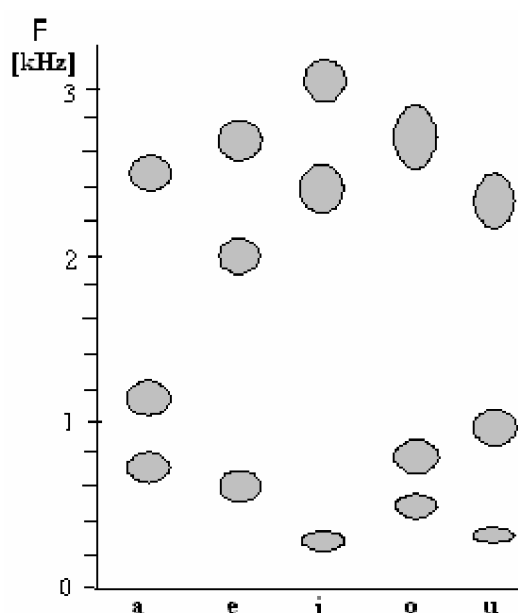
Vokální trakt má složitý tvar, který může být popsán sérií příčných průřezů funkčně závislých na vzdálenosti od hlasivek. Tyto průřezy mohou být modifikovány různým způsobem. Nejdůležitějším faktorem, který ovlivňuje tvar vokálního traktu a kvalitu samohlásek je poloha jazyka v ústní a hltanové dutině. Tento soubor svalů může zaokrouhlovat, zužovat a zmenšovat hltan a tím výrazně měnit jeho velikost. Vokální trakt je obvykle uvažován za nezávislý vůči hlasívkám.



**Obr. 5.2** Rotační model vokálního traktu pro samohlásku A

### 5.3 Formanty

Vlastní frekvence akustických módů vzduchového prostoru vokálního traktu se nazývají formanty a označují se symboly  $F_1$ ,  $F_2$ ,  $F_3$ , atd. Sled rezonančních frekvencí akustických kavit vokálních traktů odpovídá podélným, příčným i vzájemně vázaným akustickým módům. Všechny formanty jsou určeny tvarem rezonančních dutin, dále tvarem a velikostí otvoru mezi rty, polohou a tvarem jazyka a stavem otevření hlasivek. Z hlediska akustické definice dané samohlásky jsou pro srozumitelnost významné pouze první tři formanty. Srovnání formantů jednotlivých českých samohlásek je uvedeno na obr. 5.3.



**Obr. 5.3 Srovnání formantů  $F_1$ ,  $F_2$ ,  $F_3$  u českých samohlásek**

### 5.4 Teorie zdroj – filtr pro generování hlasu

Vokální trakt je akustický filtr, který zesiluje frekvenční složky zdrojového hlasu hlasivek, které jsou blízké a rovny rezonančním frekvencím vokálního traktu. Budícím efektem vokálního traktu jsou hlasivky a především vzduchové pulsy, které hlasivky generují. Tento princip se nazývá teorií zdroj - filtr pro generování samohlásek a byl prvně definován Fantem (1960). Schema této teorie je uvedena na obr. 5.4.

Tato teorie vysvětluje vznik spekter samohlásek kombinací spektra zdrojového hlasu generovaného hlasivkami a filtrací tohoto spektra přenosovou funkcí vokálního traktu. Po

průchodu vokálním traktem je zdrojový hlas upraven rezonančními frekvencemi tohoto vokálního traktu.

Pro amplitudové spektrum výstupního signálu platí:

$$V(jf) = H(jf) \cdot S(jf)$$

kde je:

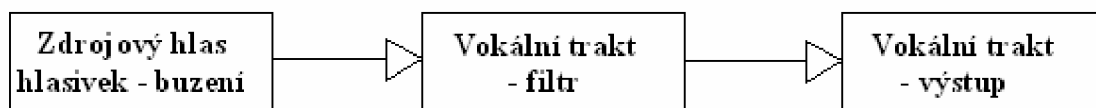
přenosová funkce vokálního traktu  $H(jf)$

spektrum zdrojového hlasu  $S(jf) = \int_{-\infty}^{+\infty} s(t) \cdot e^{-j2\pi ft} \cdot dt$

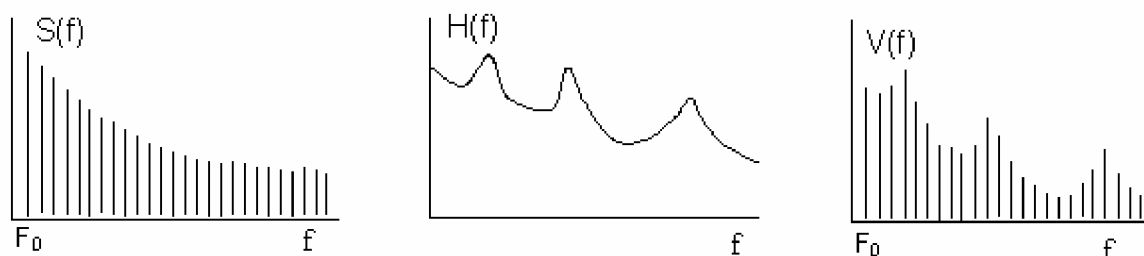
časová realizace záznamu  $s(t)$

Přenosová funkce vokálního traktu odpovídá amplitudovému spektru při jeho buzení konstantní amplitudou během celého frekvenčního rozsahu, tzn. odezva vokálního traktu na buzení s konstantní a jednotkovou amplitudou.

a) Modely:

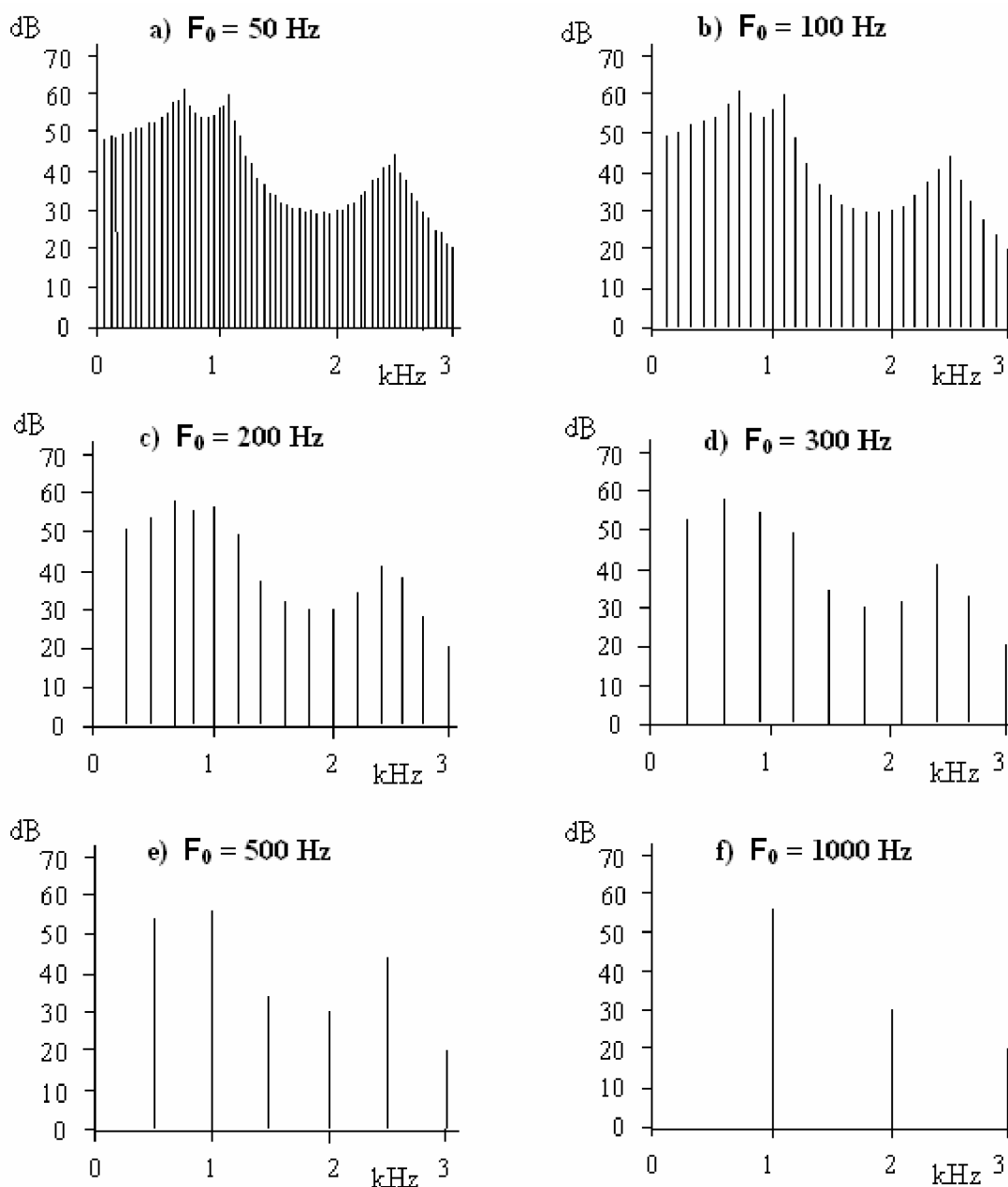


b) Spektra signálů:



Obr. 5.4 Model zdroj – filtr pro generování samohlásek

Zvyšováním základní frekvencí  $F_0$  se také zvětšují odstupy harmonických složek. Obecně se harmonické složky zdrojového hlasu mohou vzdalovat od rezonančních frekvencí vokálního traktu (formantů dané samohlásky). Generovaná samohláska je poté nedostatečně akusticky definovaná (viz obr. 5.5). To způsobuje, že samohlásky zpívané vysokými tóny zní obdobně a je proto jedno, která z nich se vlastně zpívá. Tento jev se projevuje u vysokých hlasů jako například tenor nebo soprán.



Obr. 5.5 Spektra samohlásky A pro různé základní frekvence zdrojového hlasu

## 6 Artikulace

### 6.1 Artikulace samohlásek

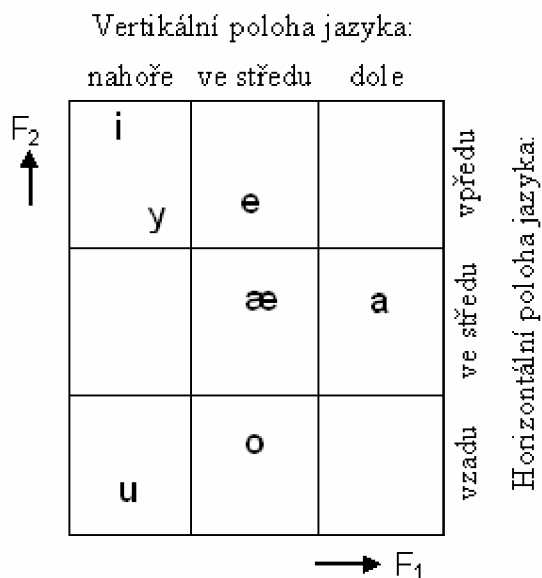
Samohlásky a souhlásky se zásadně liší způsobem generování. Samohlásky vznikají rezonančními jevy akustických kavit vokálního traktu. Pro vygenerování jednotlivých samohlásek není nutný proudící vzduch přes vokální trakt. Samohlásky tedy můžeme budít jen akustickými vlnami, například vnějším zdrojem. Za samohlásky považujeme zvuky tvořené rezonančními frekvencemi (formanty) při různém nastavení částí vokálního traktu včetně nastavení rtů a jazyka. Samohlásky se liší v mnoha fonetických vlastnostech.

#### **Spektrální vlastnosti samohlásek:**

Jednotlivé samohlásky se liší nastavením tvarů vokálního traktu. Jeho akustické módy a odpovídající rezonanční frekvence se nazývají formanty. Každému tvaru vokálního traktu odpovídá jiný soubor vlastních frekvencí – formantů. Základní a přesnou charakteristikou samohlásek je tvar vokálního traktu, tj. tvar hltanové a ústní dutiny. Výraznější změny frekvenčního rozložení formantů můžeme dosahovat hlavně změnou příčných rozměrů jednotlivých částí vokálního traktu. Soubor nejméně tří formantů definuje akusticky každou samohlásku. Každé samohlásce odpovídá jiný soubor tří formantů. Tím je v daném jazyce menší počet samohlásek (např. v češtině) a jednotlivé samohlásky jsou akusticky lépe vzájemně odlišitelné, přesněji definované. Naopak čím je v daném jazyce více samohlásek nebo různých jejich variant (např. v angličtině), tím se akusticky vzájemně pro posluchače hůře identifikují.

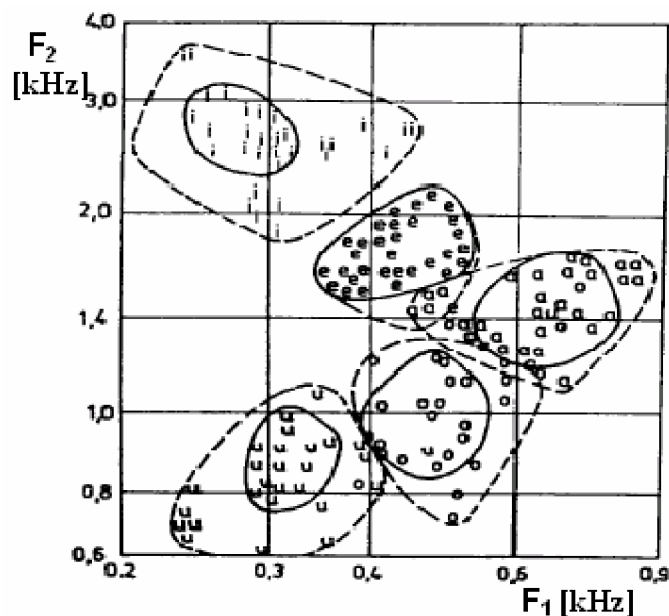
Neutrální samohláska  $\text{æ}$  (tj.  $\text{ə}$ ), vyskytující se u anglických samohlásek má při rovnoměrně rozložených rezonančních frekvencích zhruba stejné maximální amplitudy intenzit.

Schematicky se poloha jazyka při vyslovování jednotlivých samohlásek zobrazuje pomocí tzv. Hellwagova trojúhelníku uvedeného na obr. 6.1.



Obr. 6.1 Poloha prvních dvou formantů podle polohy jazyka

Z důvodu odlišnosti tvaru a velikosti vokálních traktů u jednotlivých lidí se frekvence jednotlivých formantů se mohou pohybovat v určitých tolerancích jak je vidět na obr. 6.2.



Obr. 6.2 Rozptyl polohy prvních dvou formantů českých samohlásek

## 6.2 Artikulace souhlásek

Artikulace souhlásek je realizována odlišným způsobem než je tomu při artikulaci samohlásek. To je dáno tím, že zatímco samohlásky jsou akusticky tvořeny souborem rezonančních vlastností vokálního traktu, souhlásky potřebují proudící vzduch přes různá zúžení podél vokálního traktu. Z tohoto důvodu musí být vokální trakt pro tvorbu samohlásek volný po celé jeho délce a otevřený na straně ústního otvoru. U souhlásek může být zúžení vokálního traktu nebo jeho blokování umístěno v kterémkoliv místě. Souhlásky se liší od samohlásek aktivní částí vokálního traktu. U samohlásek je využíván celý vokální trakt, zatímco u souhlásek se uplatňuje proměnlivá část vokálního traktu. Charakterem generovaných zvuků jsou rezonanční frekvence u samohlásek a neperiodické šumy a periodické zvuky u souhlásek.



## 7 Teorie popisu tvorby hlasu

Teorie definovaná Van den Bergem v roce 1958 byla nazvána myo-elasto aerodynamickou teorií a její platnost je doposud všeobecně uznávána odbornou veřejností. Na tuto teorii pak navazovala a dosud stále navazuje převážná část světových odborníků, zabývající se tvorbou lidského hlasu a především tvorbou zdrojového hlasu, který generují hlasivky.

Ukázalo se však že tato teorie není dokonalá a to vedlo k vytvoření nové teorie nazvané teorie fonace: bubliny stlačeného vzduchu.

### 7.1 Myo-elasto aerodynamická teorie

Základní myšlenky myo-elasto aerodynamické teorie jsou založeny na tom, že zdrojový hlas hlasivek je generován při průchodu vzduchu přes hlasivky. Základním parametrem je objemová rychlost vzduchových pulsů, které vznikají přerušováním proudu vzduchu z trachey vibrujícími hlasivkami. Jako budící efekt pro vokální trakt byla považována objemová rychlost, použitá v normalizovaném tvaru. Později byla objemová rychlost nahrazena derivací objemové rychlosti.

Při tvorbě zdrojového hlasu má zásadní úlohu proudící vzduch a jeho parametry. Těmi jsou tlak, hmotnost a rychlost proudu vzduchu. Pohyb hlasivek je zajištěn zejména díky Bernoulliho efektu. Je to podtlak proudícího vzduchu při průchodu proudu přes zúžená místa vibrujících hlasivek. Bernoulliho podtlak umožňuje uzavírání hlasivek. Aeroelastická interakce mezi proudícím vzduchem a elastickou tkání vybuzuje oscilace hlasivek, takto vzniká tón o základní frekvenci  $F_0$ . Vlastní frekvence hlasivek je ovlivněna intenzitou proudícího vzduchu přes hlasivky.

Tato teorie má však i zásadní nedostatky. Lidský hlas se skládá ze samohlásek a souhlásek. Samohlásky jsou buzeny jen akustickými vlnami vygenerovanými hlasivkami nebo jiným náhradním zdrojem a nevyžadují proudící vzduch, zatímco souhlásky proudící vzduch vyžadují. Samohlásky, které jsou produktem zdrojového hlasu a spektrálních vlastností vokálního traktu (formantů) tudíž nemohou být buzeny rovněž i proudícím vzduchem, jak to prezentuje myo-elasto aerodynamická teorie.

Člověk vědomě ovlivňuje funkci hlasivek (frekvenci, intenzitu), ovšem toho nemůže dosáhnout změnou parametrů typu rychlost proudění vzduchu a jeho hmotností. U výše zjednodušeně uvedené verze funkce hlasivek hraje důležitou funkci a roli samotný proudící vzduch a jeho proměnné parametry (tlak, rychlost, hmotnost). Pokud sledujeme proměnné  $V(t)$  – objemovou rychlost,  $V'(t)$  – derivaci objemové rychlosti a  $P_{SPG}(t)$  – supraglotický tlak, můžeme určit, že pouze proměnná veličina, mající jednotky [Pa] může být korektním budícím efektem vokálního traktu. Touto proměnnou může proto být pouze intraglotický tlak mezi hlasivkami, vznikající při expanzi celkového subglotického tlaku  $P_{SGC}(t)$  na supraglotický tlak  $P_{SPG}(t)$ . Tato skutečnost vyplývá z teorie přenosů.

Objemová rychlost a její derivace neobsahují jednotky tlaku, takže nelze pomocí těchto parametrů ovlivnit a budít vokální trakt a jeho formanty. Objemová rychlost a její derivace jsou výslednými efekty řešení dané konkrétní situace konkrétních hlasivek a nelze je považovat za příčiny buzení hlasivek

Z toho vyplývá, že velkým nedostatkem sledované teorie je nutnost používat normalizované tvary budících účinků. Amplitudy objemových rychlostí musí být nutně voleny jednotkovými, což je obtížně použitelné u reálných soustav  $P_{SGC}(t)$ .

Normalizované hodnoty budících účinků nemohou při fonační fázi hlasivek, když jsou sepnuté, respektovat zvyšující se hodnotu subglotického tlaku. Při nedostatečné hodnotě subglotického tlaku nemusí vůbec docházet k rozpojení hlasivek a tudíž i k jejich fonaci. Proto uvedená teorie neřeší změnu střední hodnoty subglotického tlaku, která je příčinou změny intenzity zdrojového hlasu. Neřeší dokonale ani rozkmit hlasivek, tj. průběh mezery glotis v čase a její změnu při změně střední hodnoty subglotického tlaku.

Některé aplikace vyžadují úpravu parametrů modelu podle charakteru řešeného problému. Tento přístup umožňuje dosahovat v určité míře shodu vlastností použitého modelu s reálnou soustavou. Modely však neumožňují korektní definování odezvy při různých středních hodnotách subglotického tlaku. Aktivitou hlasivek se transformuje přerušovaný proud tlakového vzduchu z plic na akustický tlak, který je budícím efektem vokálního traktu a dále na zvýšenou rychlost proudu, která je nutná při tvorbě souhlásek. Proto je nutné, aby tlak vzduchu v trachey byl co nejvyšší (vyšší intenzita hlasu) a nebyl výrazně snižován přeměnou na rychlost proudění vzduchu hlasivkami.

Bernoulliho podtlak by se musel projevovat během cyklu otevírání a uzavírání mezery glotis, protože nejvyšší rychlost proudění vzduchu je v okamžiku rozevírání hlasivek.

Pokud je Bernoulliho efekt přítomen při pohybu hlasivek, pak by musel mít vyšší hodnotu při rozevírání hlasivek, kdy je vyšší subglotický tlak a proto i vyšší rychlost průtoku mezerou glotis a nikoliv při jejich uzavírání. Z toho vyplývá, že by případný Bernoulliho efekt byl spíše brzdou než hnacím jevem pohybu hlasivek. Výrazný vliv proudícího vzduchu nemůže mít spojitost se změnou frekvence hlasivek, která je jednoznačně definovaná parametry hlasivek, tzn. pružnými a hmotnostními charakteristikami, jejich strukturou a tvarem.

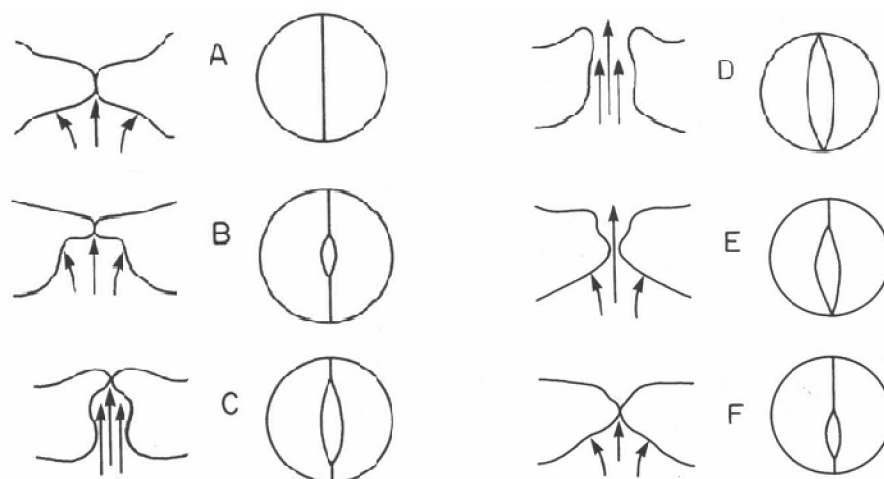
## 7.2 Teorie fonace: bubliny stlačeného vzduchu

Na základě řady výše zmíněných nejasností a nedostatků dosud definovaných principů různými autory byl definován a je dále vyvíjen na ÚMTMB FSI VUT Brno nový princip funkce hlasivek označovaným jako princip „bubliny tlakového vzduchu“ zkráceně princip „bubliny“.

Tlakové pulsy vznikají přerušováním vyššího subglotického tlaku v trachey vibrujícími hlasivkami a změna subglotického tlaku během periody vibrací je dána změnou mezery glotis. Tím dochází k vzájemné vazbě mezi změnou subglotického tlaku a změnou velikosti glotis. Proto je možné doplnit a rozšířit definici principu funkce hlasivek na: „princip bublin tlakového vzduchu generovaných na základě charakteristiky subglotického tlaku“. Bubliny tlakového vzduchu jsou vypouštěny vibrujícími hlasivkami do supraglotického prostoru, kde expandují a generují akustické vlny, jež tvoří zdrojový hlas.

Hlavními silami, působícími na hlasivky při fonaci jsou subglotický tlak pod hlasivkami a v celé trachey. Tento přetlak působí na poměrně velkou vnitřní plochu subglotické oblasti hlasivek, takže poměrně značná síla rozevírá hlasivky. Dále zde působí pružné síly hlasivkových svalů, které působí proti rozevírání hlasivek a setrvačné síly hlasivek.

Charakteristikami zdrojového hlasu, které jsou při mluvení nahlas člověkem vědomě ovlivnitelné jsou intenzita zdrojového hlasu, která je dána pouze střední hodnotou subglotického tlaku a výška základního tónu hlasu. Těmito dvěma parametry je zdrojový hlas dostatečně a zcela definován.



**Obr. 7.1 Fáze pohybu hlasivek při jedné periodě fonace**

### **Důležité fáze fonace:**

Pohyb hlasivek uvedený na obr. 7.1 startuje z tzv. fonační fáze, kdy jsou hlasivky u sebe – A, v první fázi je tlaková bublina vtačována do prostoru mezi hlasivkami, což umožňuje rozevírání jejich spodní části – B, bublina dále postupuje směrem nahoru k vokálnímu traktu až začne fáze otevírání horní části hlasivek – C. Následně se hlasivky rozevřou a směrem nahoru dochází k vytlačování tlakové bubliny do supraglottického prostoru a dále do vokálního traktu – D, bublina expanduje a generuje se zdrojový hlas. Dále dochází k uzavření hlasivek v jejich spodní části – E, vzápětí nastává postupné spojování hlasivek odspodu směrem nahoru – F. Nakonec se hlasivky dostávají zpět do původního fonačního postavení – A.

V tomto sledu je vytvořena vzduchová bublina a je převedena z subglottického prostoru do supraglottického.

Frekvence hlasivek je jednoznačně definována hmotně-pružnými vlastnostmi struktury hlasivek a jejich vrstev a nikoliv rychlostí proudění vzduchu. Základním předpokladem fonace je samobuzené kmitání hlasivkové struktury. Frekvence hlasivek je proto blízká nebo totožná některé vlastní frekvenci struktury hlasivek. Ne každý strukturní mód hlasivek je však schopný zajistit fonaci definovanou výše uvedeným principem. Tuto funkci bude nejspíš schopen zajistit některý vyšší tvar kmitu fonujících hlasivek. Podle použitého strukturního módu hlasivek při dané fonaci se zřejmě jedná o různé hlasové registry (modální – při spojování hlasivek a falseto – při jejich nespojování). Hnacím jevem hlasivek v subglottickém prostoru při jejich fonaci je tlakový vzduch, který je vždy vyšší než v prostoru supraglottickém. Důležitá je změna subglo-

tického tlaku vzduchu během periody fonace. I když subglotický tlak v trachey se snaží plíce udržovat konstantní, v důsledku vibrací hlasivek a měnící se mezerou glotis, dochází k odpouštění vzduchu do supraglotického prostoru. Tím se hodnota subglotického tlaku mění v závislosti na velikosti mezery glotis. Tato závislost se nazývá „charakteristikou subglotického tlaku“ a je zásadní při fonaci. U tohoto principu jde proto o zásadně odlišný princip funkce hlasivek než který je prezentován ve světové literatuře.

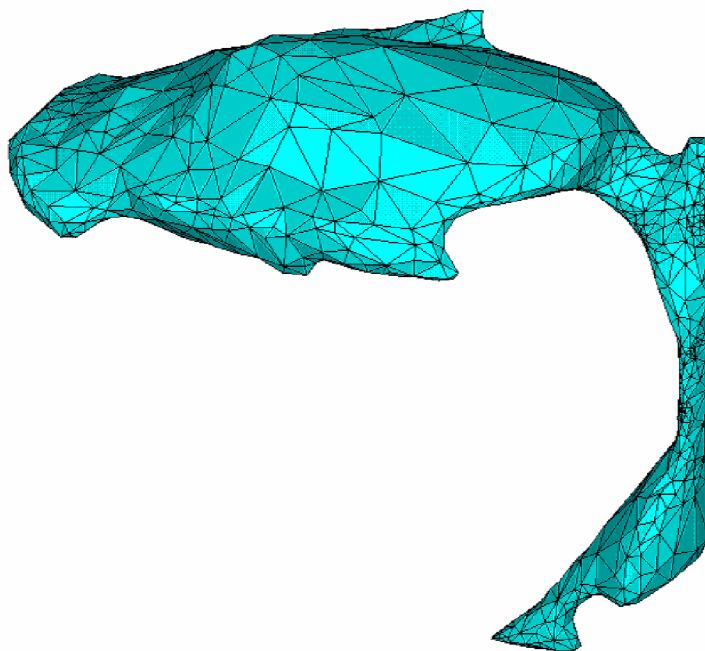
Experimentální výsledky dokazují, že pohyb vzduchu přes hlasivky lze definovat procházejícími bublinami stlačeného vzduchu ze subglotického do supraglotického prostoru. Vzduch procházející přes hlasivky je ve tvaru množin samostatných malých objemů – bublin stlačeného vzduchu (pulsů). Přítomnost bublin je nutnou podmínkou pro generování dostatečného počtu harmonických složek pro buzení formantů jednotlivých samohlásek. Bubliny v supraglotickém prostoru expandují a tím generují akustické vlny, jejichž celkové frekvenční složení definuje zdrojový hlas hlasivky. Hlasivky při fonaci vykonávají jeden ze strukturních módů, přičemž jeho vlastní frekvence definuje základní frekvenci zdrojového hlasu. Tento speciální tvar módu hlasivek umožňuje vytvarovat z časově téměř sinusového průběhu subglotického tlaku speciální periodický průběh supraglotického tlaku, který vedle základní frekvenční složky zdrojového hlasu vygeneruje také dostatečný počet vyšších harmonických. Počet vyšších harmonických složek k základnímu hlasu je dostatečný, pokud umožňuje vybudování alespoň tří nejnižších formantů používaných samohlásek. Celkový subglotický tlak vzduchu se při fonaci skládá z konstantní složky a střídavé složky v důsledku otevírání a zavírání hlasivek, tzn. expanzí a odpouštěním částí vzduchu (bublin) do supraglotické oblasti.

## 8 Tvorba počítačového modelu

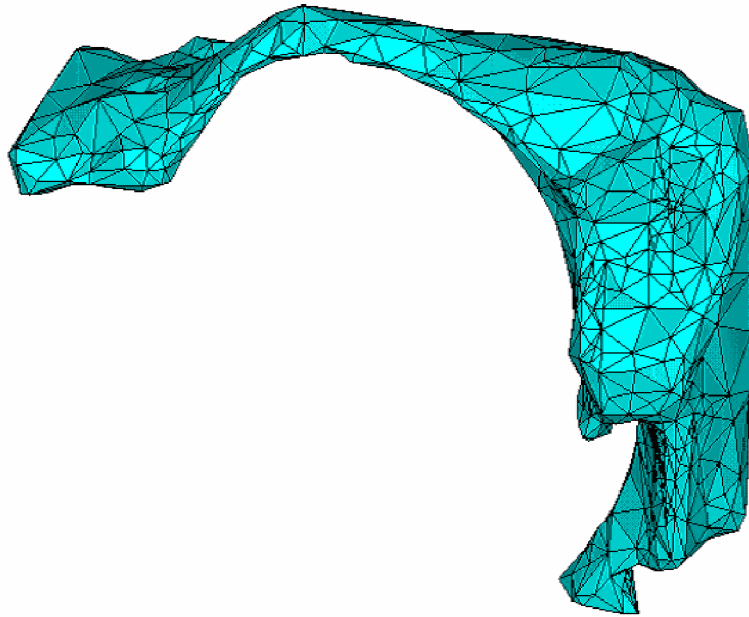
### 8.1 Model geometrie

Prostorové modely vokálních traktů pro samohlásky A (obr. 8.1), I (obr. 8.2) byly převzaty z materiálů Ústavu mechaniky těles, mechatroniky a biomechaniky. Tvorba modelů by totiž byla velmi pracná a není hlavním úkolem této práce. Oba modely jsou vymodelovány bez nosní dutiny. Ta není pro generování samohlásek důležitá a může se proto zanedbat.

Převzaté modely byly vytvořeny z řezů získaných magnetickou rezonancí. V ANSYSu se pro tvorbu sítě použily akustické prvky FLUID30.



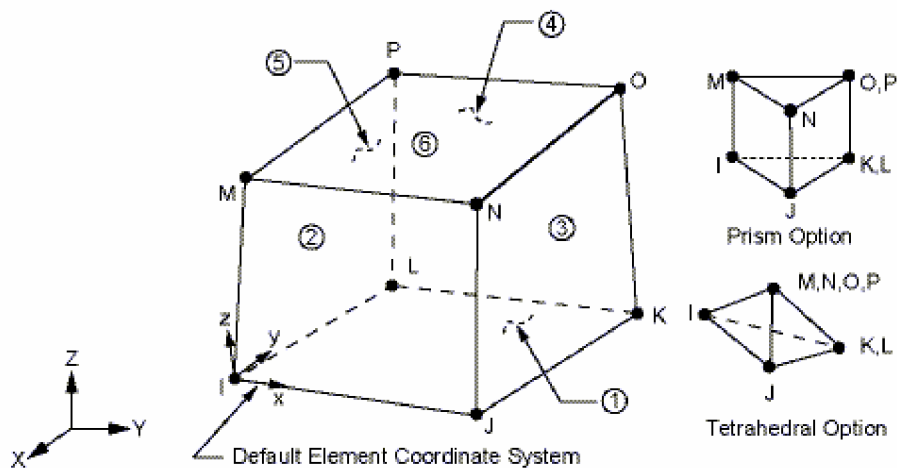
**Obr. 8.1** Prostorový model samohlásky A



Obr. 8.2 Prostorový model samohlásky I

### Charakteristiky prvku FLUID30:

FLUID30 je osmiuzlový prvek, který se používá v prostorových úlohách pro modelování akustického pole. Ve výpočtovém modelu bylo zvoleno čtyřstěnové nastavení prvku. Schematické znázornění prvku je na obr. 8.3.



Obr. 8.3 Schema prvku FLUID30 [13]

Prostorový prvek FLUID30 má dvě nastavení

- prism (trojboký hranol)
- tetrahedral (čtyřstěn)

Podrobnější informace o vstupních parametrech prvku FLUID30:

UX,UY,UZ	posuvy ve směru os
PRES	tlak
PREF	referenční tlak
DENS	hustota prostředí
SONC	rychlost zvuku v daném prostředí
MU	pohltivost zvuku
KEYOPT(2) = 0	prvek je nastaven jako přechodový
KEYOPT(2) = 1	prvek je nastaven jako vnitřní

## 8.2 Model materiálu

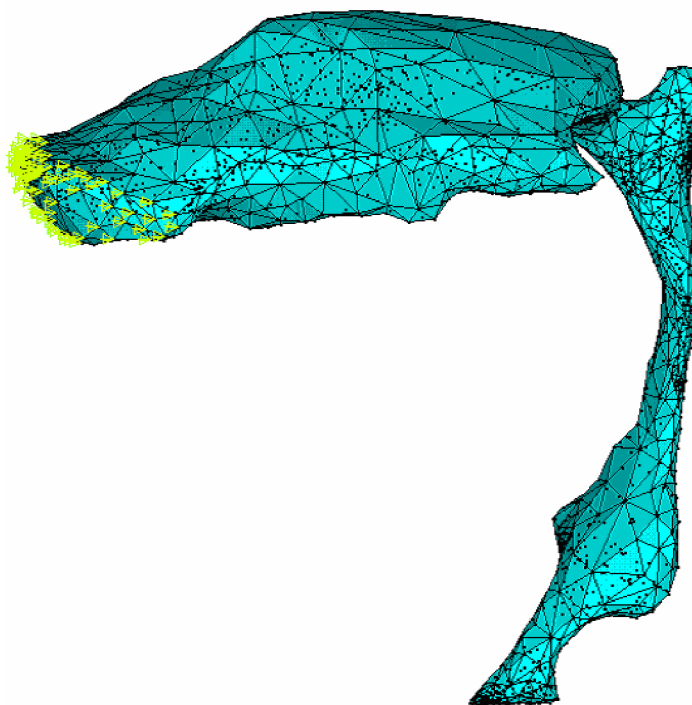
Pro použitý prvek FLUID30 jsou zapotřebí zadat fyzikální parametry proudícího média. Těmi jsou v našem případě hustota vzduchu a rychlost zvuku. Hodnoty těchto veličin jsou uvedeny v následujícím přehledu:

- Hustota vzduchu  $\rho = 1,225 \text{ kg}\cdot\text{m}^{-3}$
- Rychlost zvuku  $c = 340 \text{ m}\cdot\text{s}^{-1}$

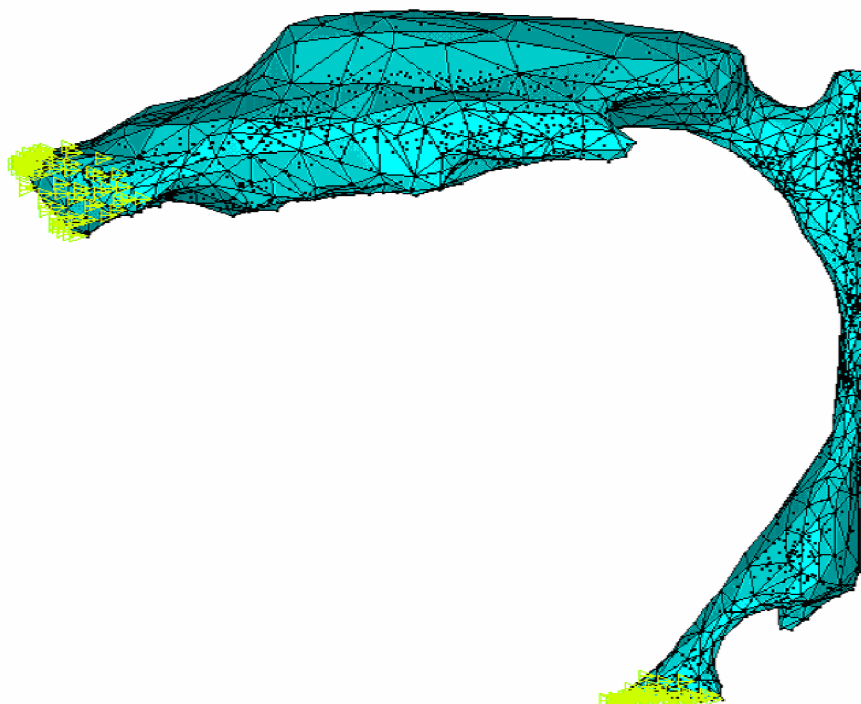


### 8.3 Model vazeb

Pro modelování nastavení vokálních traktů při mluvení nahlas jsme předepsali okrajové podmínky nulového akustického tlaku na uzly v oblasti úst (obr. 8.3). Pro mluvení šeptem je předepsán nulový tlak pro uzly v oblasti úst a také v oblasti hlasivkové štěrbiny (obr. 8.4). Pro všechny modely byly rovněž nastaveny nulové posuvy na všechny externí uzly s výjimkou míst s nulovým akustickým tlakem.



**Obr. 8.3 Okrajové podmínky pro samohlásku A nahlas**



**Obr. 8.4 Okrajové podmínky pro samohlásku A šeptem**

#### 8.4 Model zatížení

Pro harmonickou analýzu byl nastaven akustický tlak v oblasti hlasivkové štěrbiny o velikosti 1 Pa. Tato hodnota byla zvolena z důvodu toho, že lze snadno přepočítat na reálný tlak vynásobením vhodným koeficientem.

## 9 Výsledky výpočtového modelování

Při správném nastavení okrajových podmínek byly namodelovány a spočteny vlastní frekvence vokálního traktu při mluvení nahlas a šeptem pro samohlásky A,I.

V tab. 9.1 jsou uvedeny první tři vypočtené vlastní frekvence vokálních traktů (formanty), pro všechny zkoumané samohlásky.

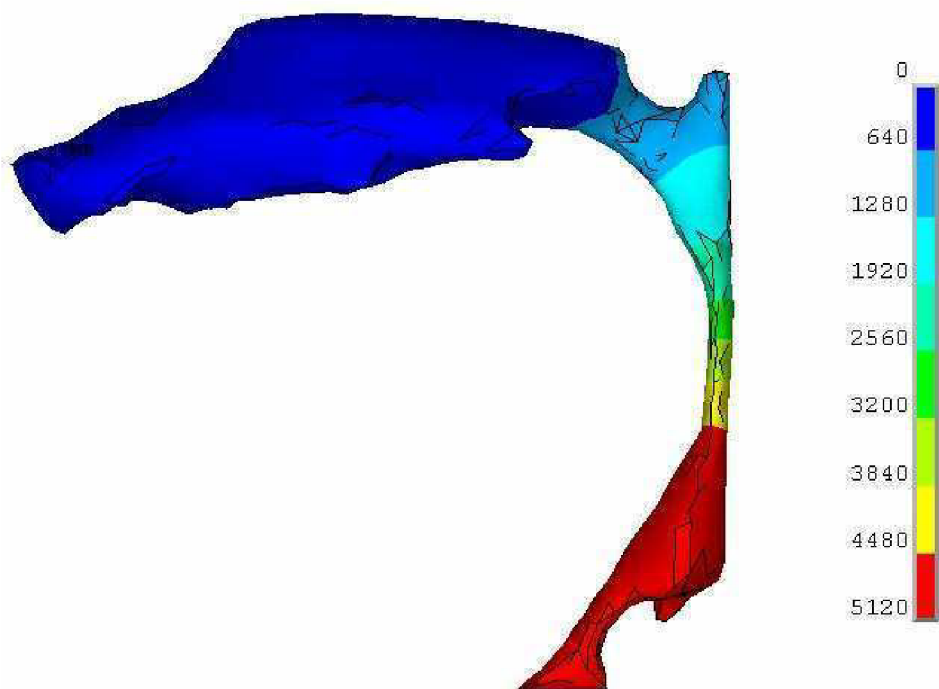
Samohláska	A			I		
	F1 [Hz]	F2 [Hz]	F3 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]
vlastní frekvence						
hodnoty nahlas	586	1053	2733	245	2271	3131
hodnoty šeptem	986	1275	2753	514	2502	3175

**Tab. 9.1** Hodnoty formantů

### 9.1 Průběhy akustických tlaků

Vlastní tvary ukazují, že se při nejnižších vlastních frekvencích vyskytují pouze podélné kmity vzduchového sloupce. Průběhy akustických tlaků pro modely samohlásky A nahlas a šeptem jsou uvedeny na obr. 9.1 až 9.3. Příčné kmity se objevují u obou modelovaných samohlásek až při páté vlastní frekvenci, což je vidět na obr 9.4. Na obr. 9.5 a 9.6 jsou zobrazeny tvary pro samohlásku I. Kombinované kmity jsou přítomny od sedmé vlastní frekvence. Příčné a kombinované kmity však nemají příliš velký vliv na srozumitelnost samohlásek, ale pouze na jejich následné zabarvení.

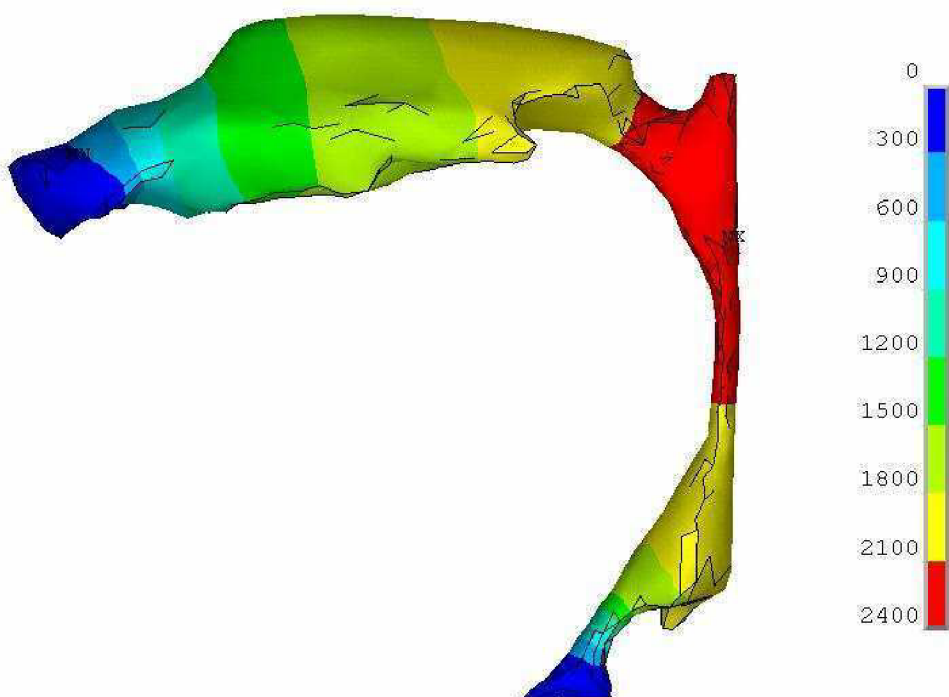
Při modelování s okrajovými podmínkami pro mluvení šeptem došlo k výraznému posunu prvního formantu do oblasti vyšších frekvencí (vůči mluvení nahlas) a jeho přiblížení k druhému formantu, který se sice také zvýšil, ale podstatně méně.



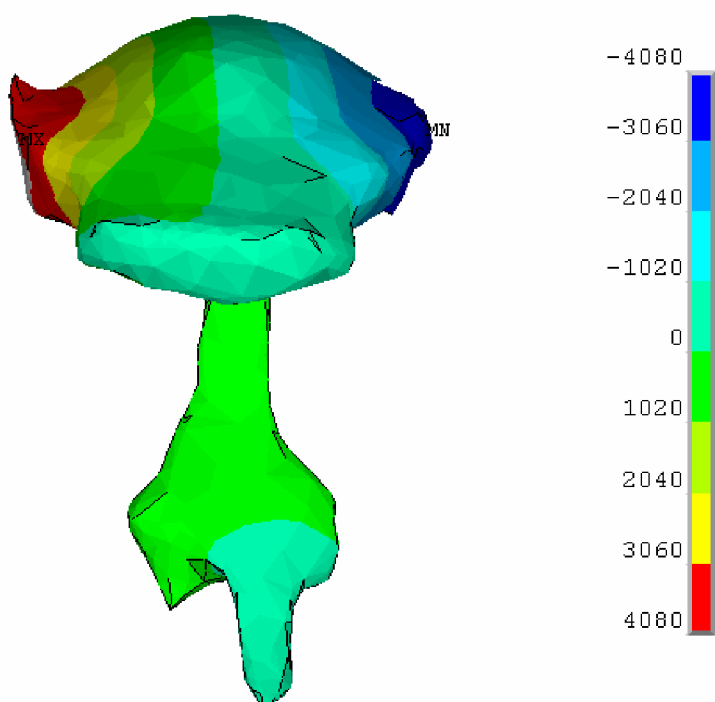
Obr. 9.1 Samohláska A nahlas pro první tvar



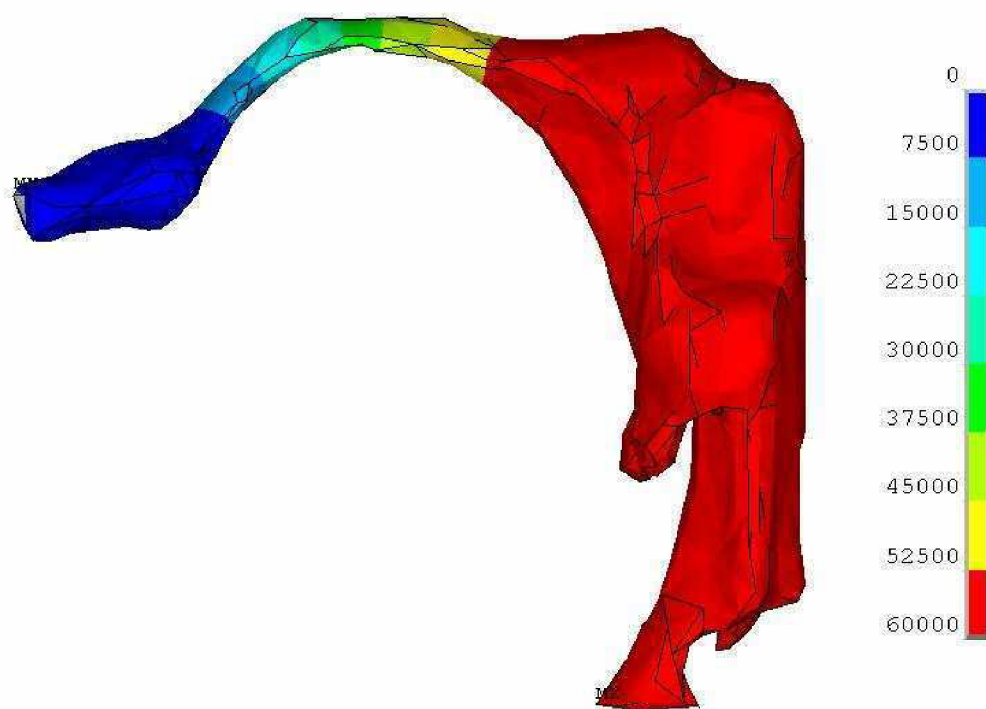
Obr. 9.2 Samohláska A nahlas pro druhý tvar



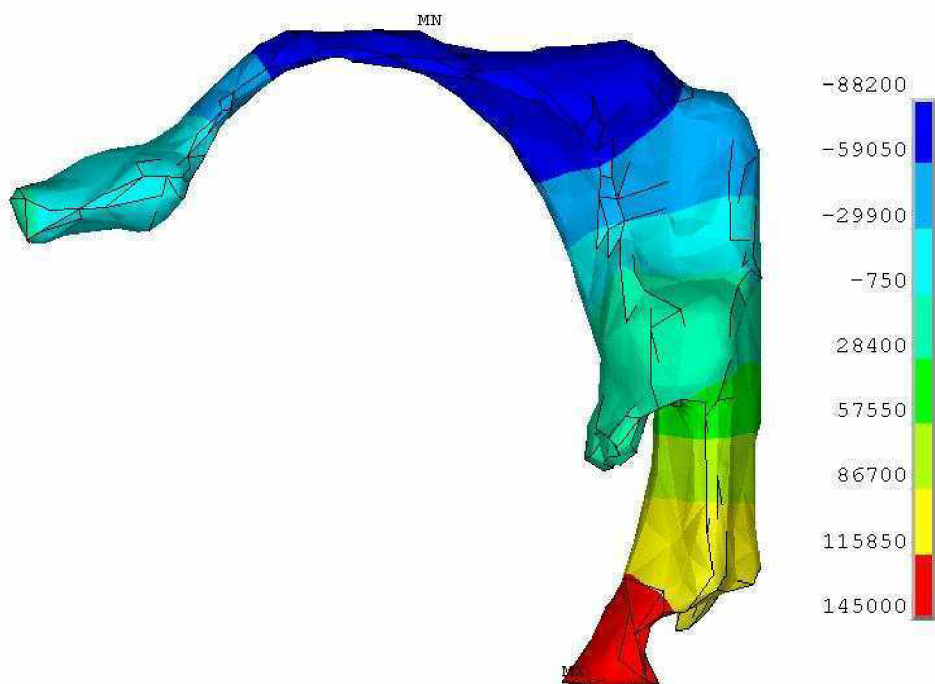
Obr. 9.3 Samohláska A šeptem pro první tvar



Obr. 9.4 Samohláska A šeptem pro pátý tvar



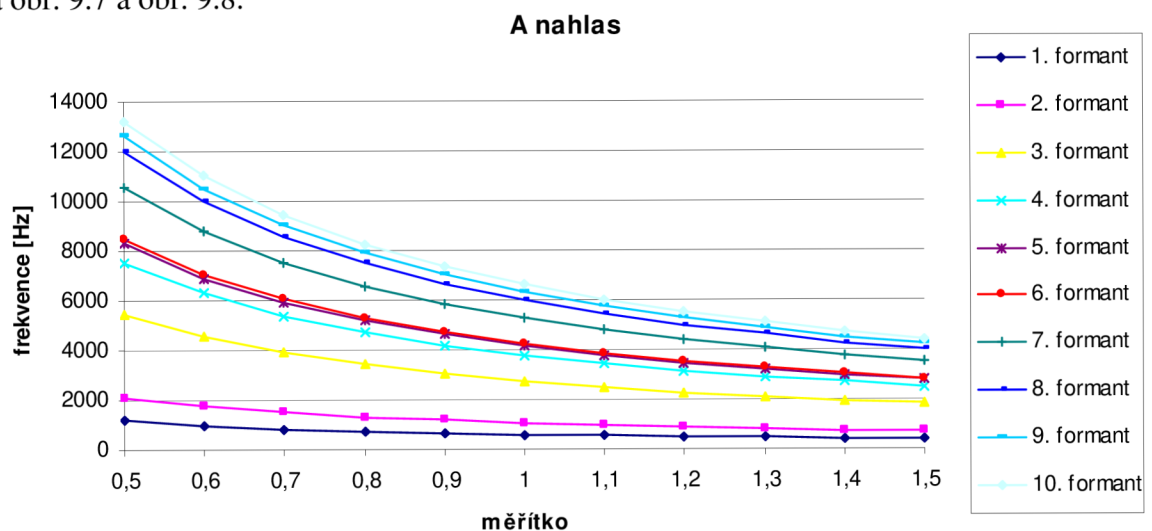
Obr. 9.5 Samohláska I nahlas pro první tvar



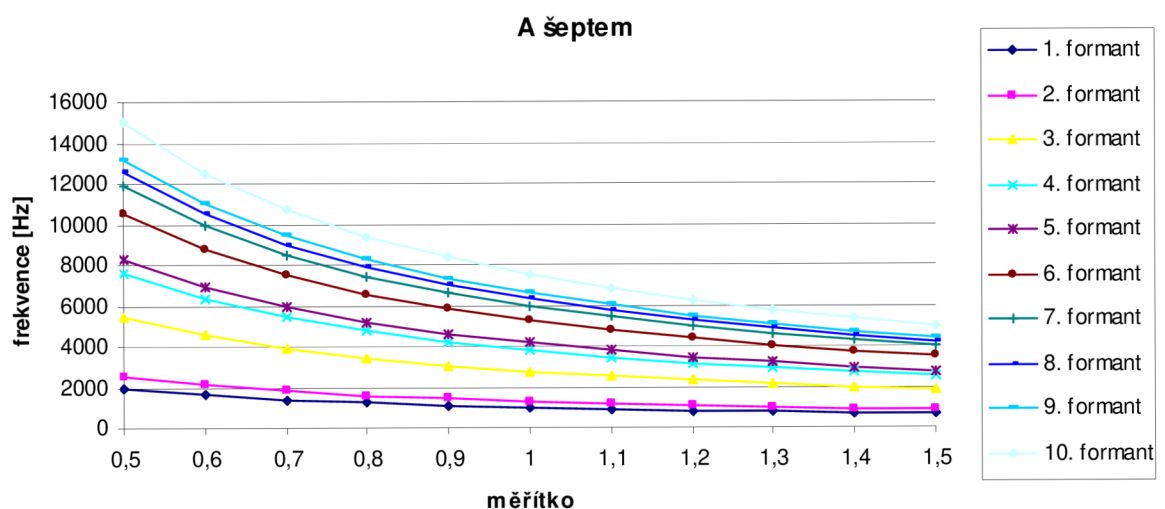
Obr. 9.6 Samohláska I nahlas pro druhý tvar

## 9.2 Vliv velikosti vokálního traktu

V práci byl sledován také vliv velikosti vokálního traktu na formanty. Výpočtovou simulací bylo dokázáno, že se zmenšováním traktu se formantové frekvence posouvaly k vyšším hodnotám a naopak. U samohlásky A se výrazně nemění vzdálenost prvních dvou formantů, od třetího už je posun od prvního formantu daleko zřetelnější jak je dobře ukázáno na obr. 9.7 a obr. 9.8.

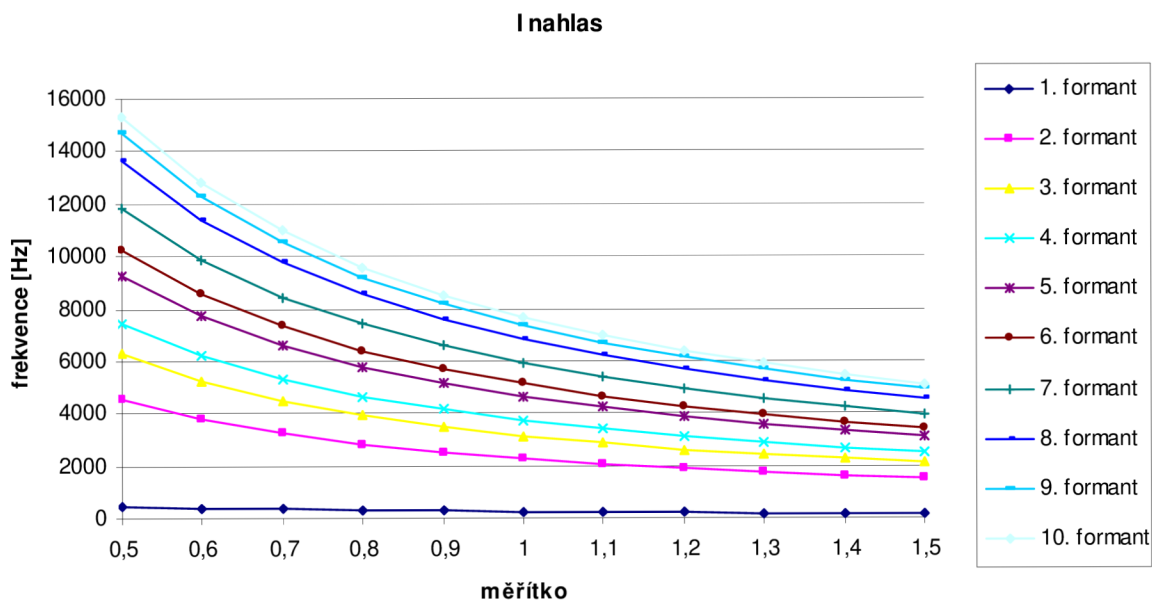


Obr. 9.7 Závislost frekvencí na velikosti vokálního traktu u samohlásky A nahlas

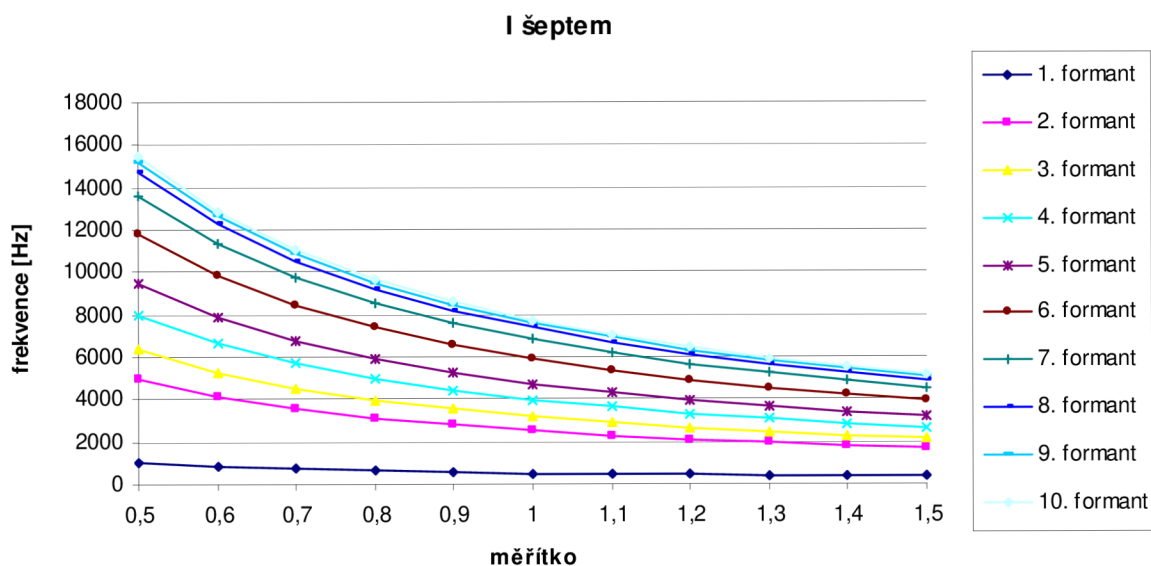


Obr. 9.8 Závislost frekvencí na velikosti vokálního traktu u samohlásky A šeptem

U samohlásky I je nárůst frekvence při zmenšování vokálního traktu nepatrný pouze u prvního formantu. Od druhého už je posun vůči prvnímu formantu daleko zřetelnější jak je patrné z obr. 9.9 a obr. 9.10.



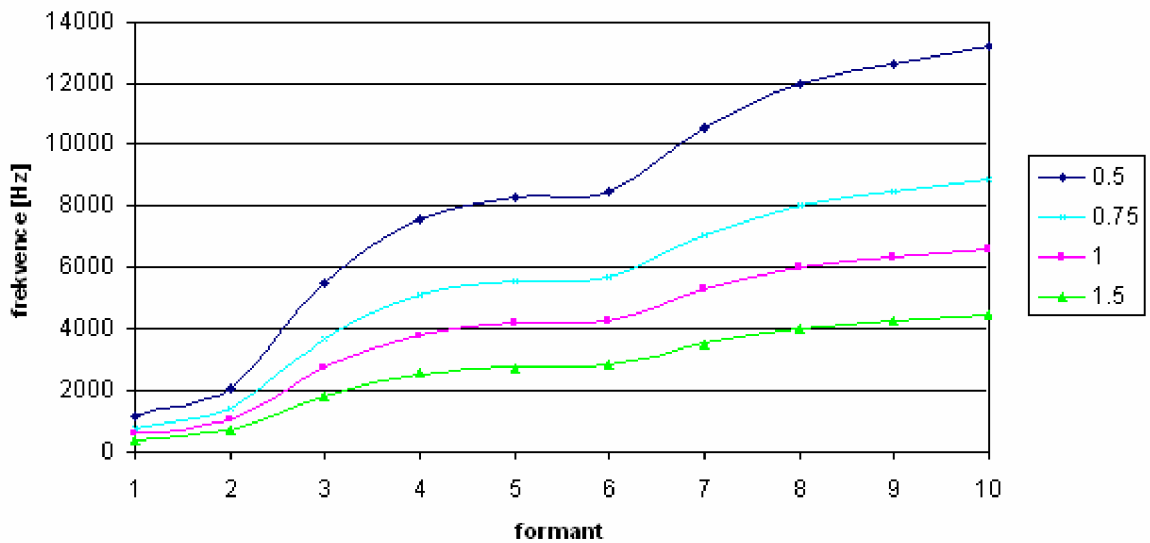
Obr. 9.9 Závislost frekvencí na velikosti vokálního traktu u samohlásky I nahlas



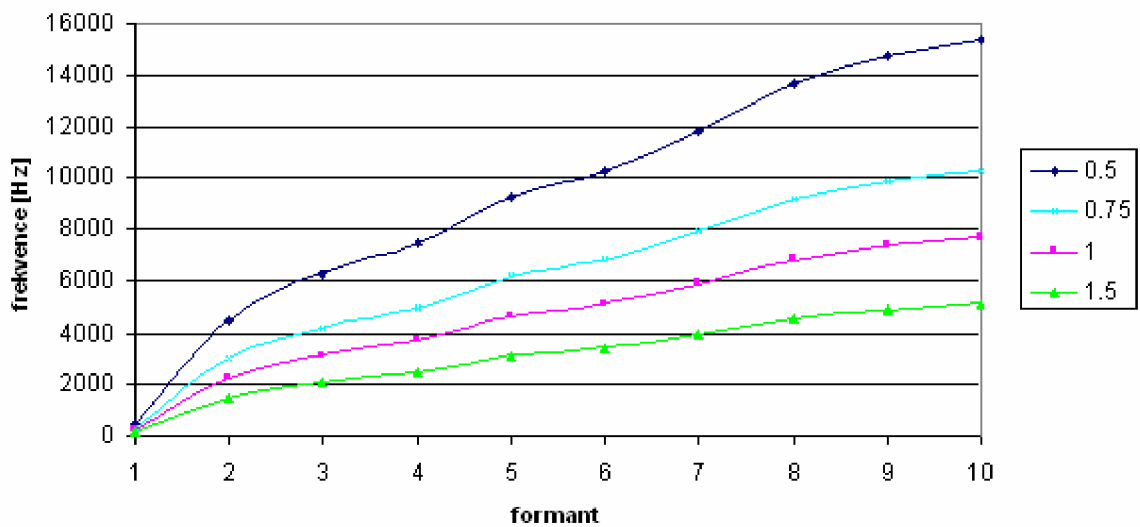
Obr. 9.10 Závislost frekvencí na velikosti vokálního traktu u samohlásky I šeptem



Závislosti frekvencí jednotlivých formantů na velikostech vokálních traktu u samohlásek jsou zobrazeny na obr. 9.11 a obr. 9.12.

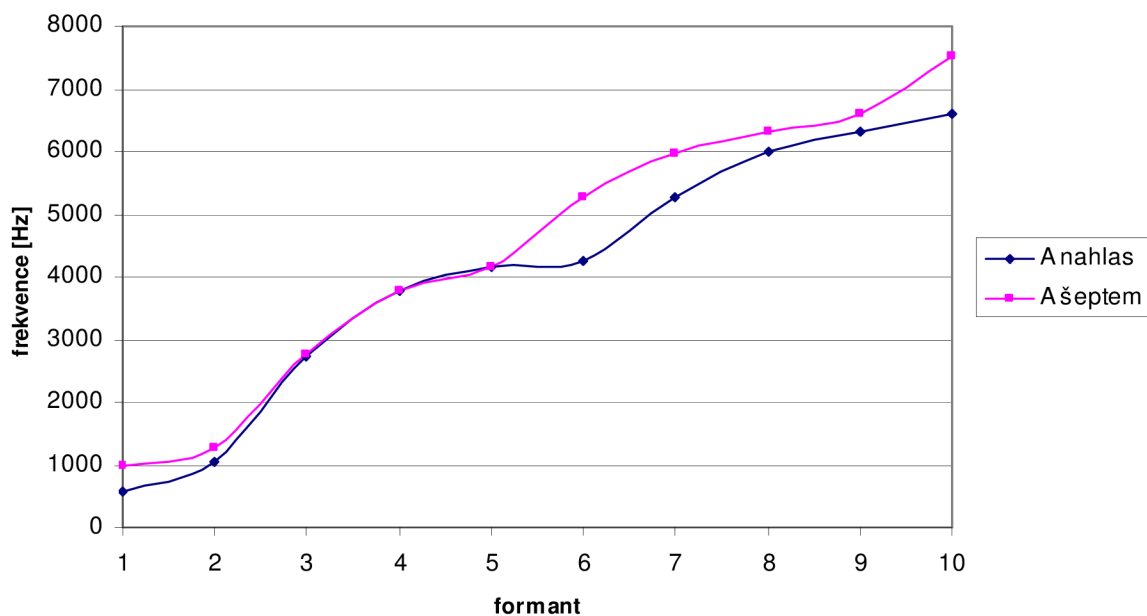


Obr. 9.11 Samohláska A nahlas

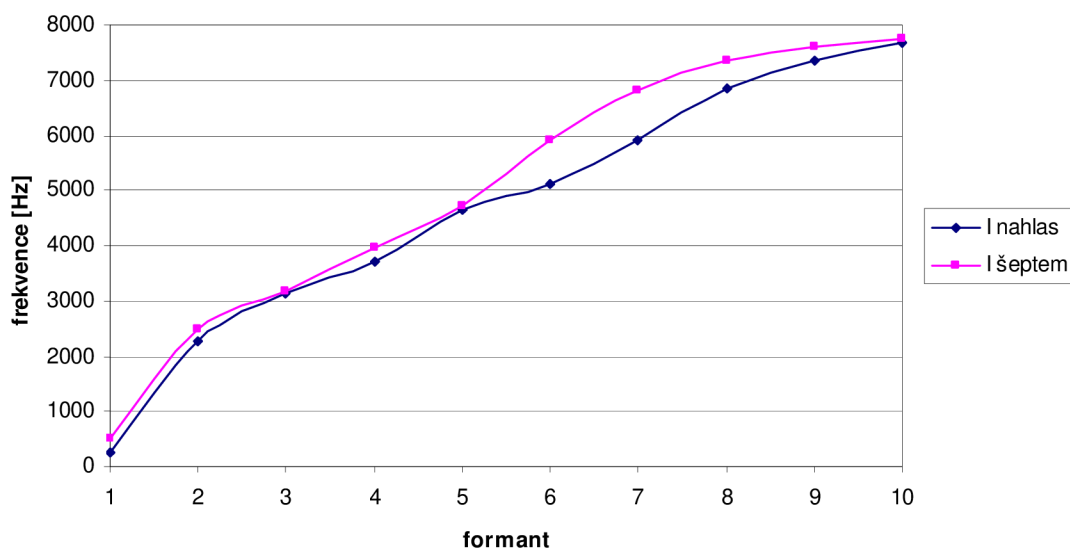


Obr. 9.12 Samohláska I nahlas

Bylo zjištěno, že se při vyslovování samohlásek nahlas vyskytuje jistý (šestý) formant, který se u mluvení šeptem nevyskytuje, a tím způsobuje posun v číslování formantů. Z toho vyplývá, že například 7. formant při mluvení nahlas frekvenčně odpovídá 6. formantu pro mluvení šeptem, atd. To je patrné z obr. 9.13 a obr. 9.14.



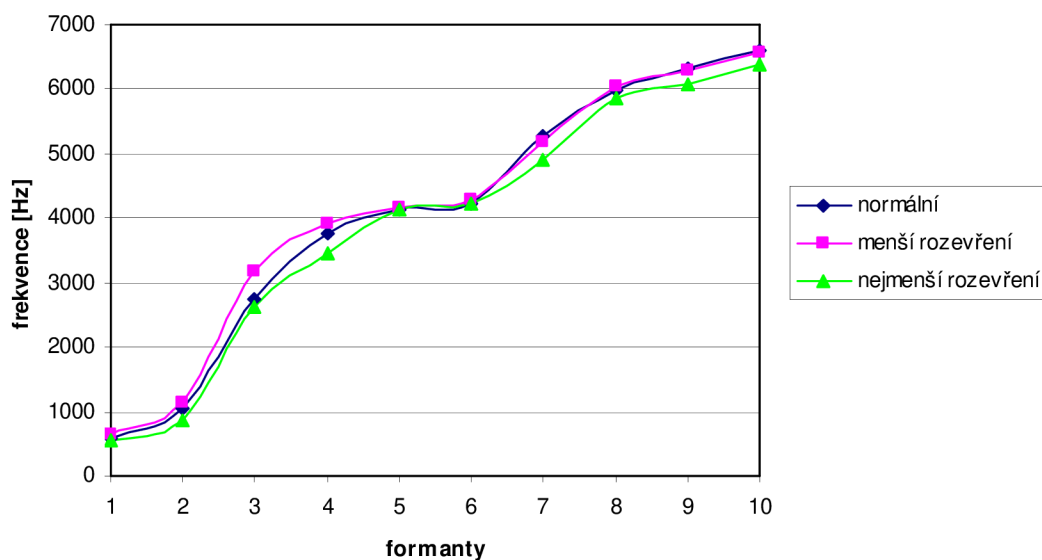
Obr. 9.13 Samohláska A



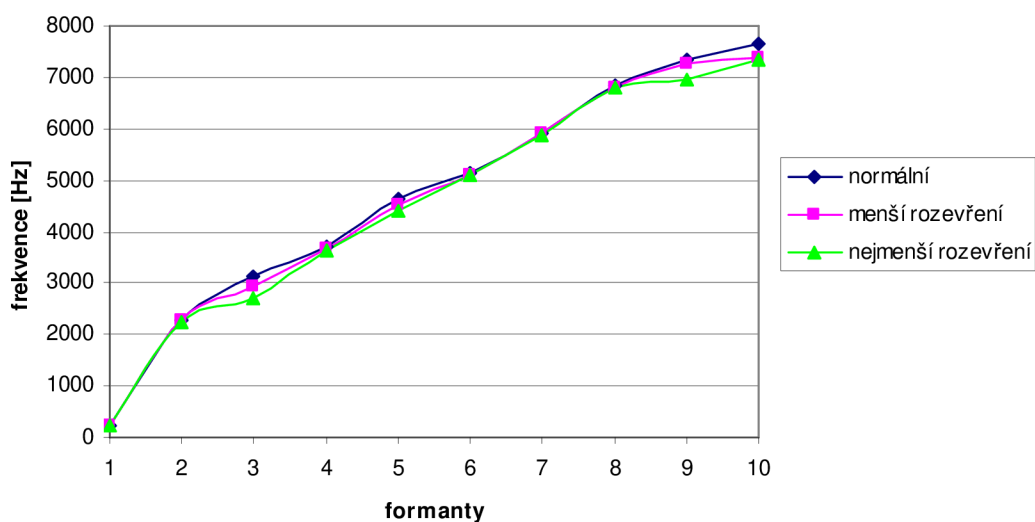
Obr. 9.14 Samohláska I

### 9.3 Vliv polohy otevření úst

Následně bylo počítačovým modelováním ověřováno, jaký vliv na formanty má otevření úst při vyslovování zkoumaných samohlásek. U samohlásky A je patrné ovlivnění v oblasti třetího a čtvrtého formantu, viz obr. 9.15, avšak nelze s určitostí říct zda v daném místě hodnota stoupá nebo klesá. U samohlásky I se zmenšujícím se ústním otvorem nepatrně klesal třetí a pátý formant jak je vidět na obr. 9.16.



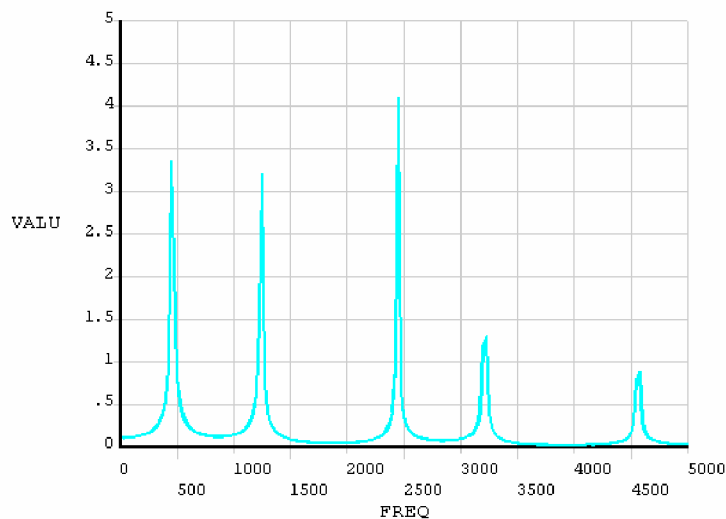
Obr. 9.15 Samohláska A



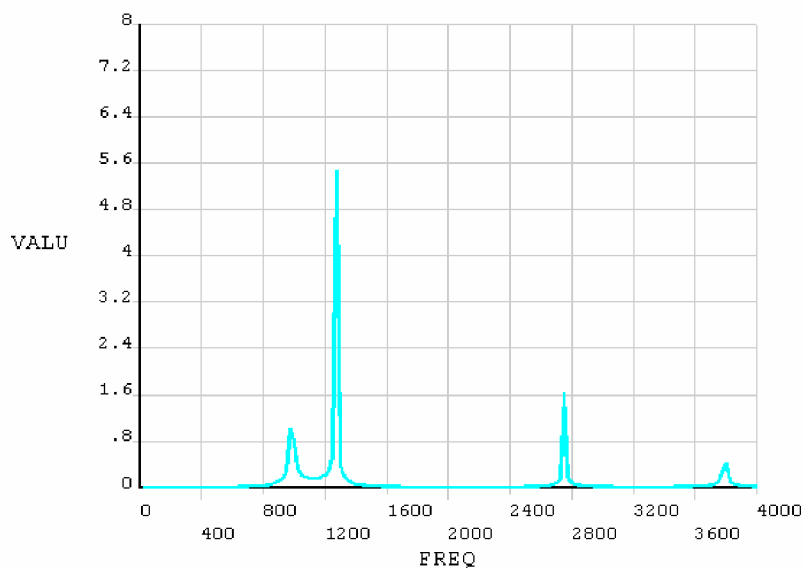
Obr. 9.16 Samohláska I

## 9.4 Harmonické buzení vokálních traktů

Na modelech vokálních traktů pro jednotlivé samohlásky bylo provedeno buzení na straně hlasivek pomocí harmonických tlakových pulzů. Velikost tlaku byla nastavena jednotková (1 Pa), aby se dala hodnota pro skutečný tlak snadno dopočítat vynásobením vypočtených hodnot vhodným koeficientem. Z těchto výpočtů byly zjištěny amplitudo-frekvenční charakteristiky uvedené na obr. 9.17 a 9.18.

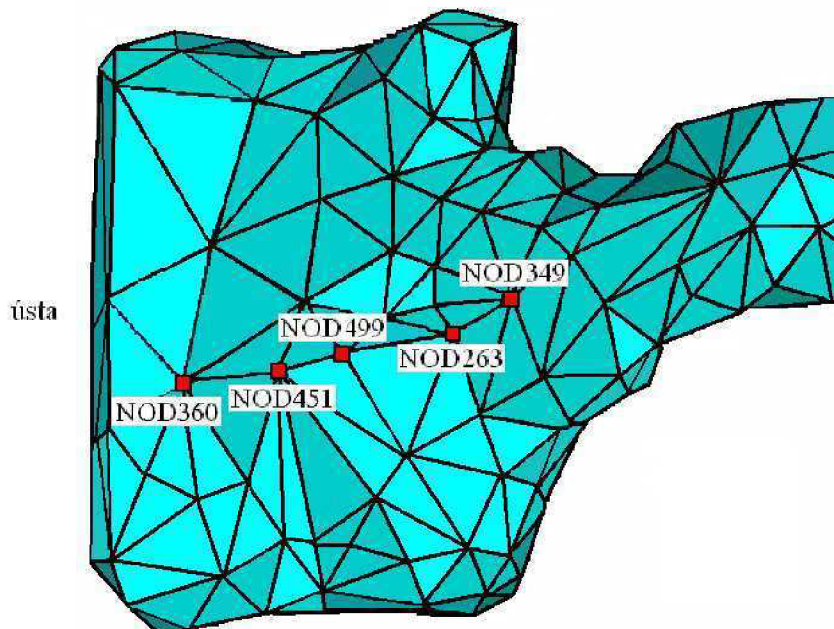


**Obr. 9.17 Samohláska A nahlas**

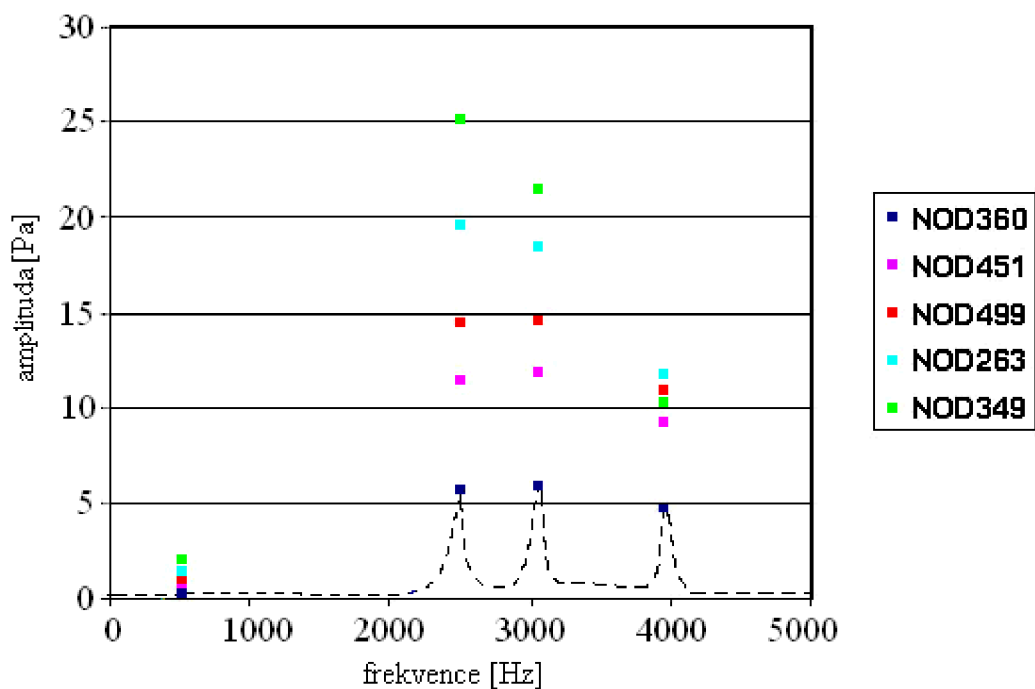


**Obr. 9.18 Samohláska A šeptem**

Amplitudo-frekvenční charakteristiky byly zjištěny pro různé vzdálenosti od úst. Místa vykreslovaných závislostí jsou uvedena na obr. 9.19. Vrcholy výsledných tlaků pro jednotlivé uzly jsou pak vykresleny na obr. 9.20.



Obr. 9.19 Uzly výstupních tlaků samohlásky I



Obr. 9.20 Vrcholy charakteristik samohlásky I šeptem

## 10. Experimentální analýza formantů

### 10.1 Experimentální určení polohy formantů při mluvení nahlas

Při mluvení nahlas generují hlasivky tón, který má základní frekvenci  $F_0$ . Ve spektru se dále vyskytují harmonické složky, které jsou násobkem základní frekvence  $2xF_0$ ,  $3xF_0$  atd. Pouze tyto harmonické složky mohou být formantem zvýrazněny. Pokud leží formant mezi dvěma harmonickými složkami, pak dojde k určitému zvýraznění obou těchto složek. Přesné určení formantu však z tohoto důvodu není možné. Ve skutečnosti se ukázalo, že při experimentálním určení polohy formantů je nejlepší vytvářet hlubší tón. Potom je i při nízkých frekvencích přítomno velké množství frekvencí harmonických k frekvenci základní. Je tím snadnější identifikace jednotlivých formantů samohlásek, a také větší pravděpodobnost, že je některá z harmonických frekvencí přibližně rovna některému z prvních dvou formantů. Když nějaká harmonická frekvence leží v těsné blízkosti formantu, tak je tato frekvence velmi zvýrazněna a formant je možné určit přesněji.

### 10.2 Experimentální určení polohy formantů při mluvení šeptem

Při mluvení šeptem hlasivky generují spojité spektrum. Formanty lze proto určit přesně. Dochází ovšem k posunu formantů směrem k vyšším hodnotám, protože při šepotu je mezi hlasivkami mezera. Tato změna okrajových podmínek způsobuje změnu modálních vlastností vokálního traktu.

### 10.3 Analýza formantů

Dlouhodobým zkoumáním formantů českých samohlásek bylo zjištěno, že na srozumitelnost samohlásek vyslovovaných nahlas, má významný vliv základní frekvence. Na této frekvenci kmitají hlasivky a je zdrojem akustického signálu. Při měření prováděného u vysokého ženského hlasu, dochází k velkým problémům. Základní frekvence mohou dosáhnout hodnoty až 1 kHz. V rozsahu ( $0 \div 3$ ) kHz, kde leží první dva formanty, které největší měrou přispívají ke správnému určení samohlásky, se pak vyskytuje poměrně malý počet frekvencí. Ty mohou být následně ovlivněny rezonančními dutinami vokálního traktu. Určování přesných poloh formantů u takového hlasu je pak velmi obtížné. Někdy může při zpěvu těchto

vysokých sopránových tónů dokonce dojít k tomu, že základní frekvence tónu jsou vyšší než první formant. To je následně příčinou špatné srozumitelnosti ženského zpěvu ve vysokých polohách. Frekvenční spektra samohlásek A,E jsou velice podobná, což vede při zpěvu k zaměňování těchto dvou samohlásek.

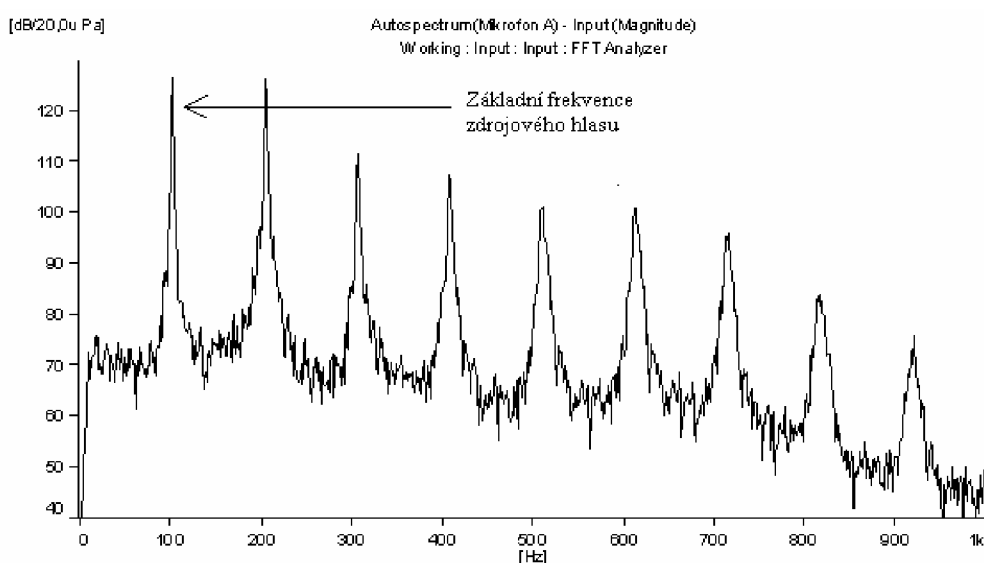
Znalosti formantů samohlásek můžeme efektivně využívat u rezonančních vlastností vokálního traktu. Při správném nastavení základní frekvence hlasivek můžeme dosáhnout přiblížení některé z harmonických frekvencí k prvnímu formantu. To se projeví výrazným zesílením intenzity této harmonické složky a tím i dané samohlásky. V praxi se tento jev často využívá při tzv. ladění formantů.

Jiná situace se vyskytuje při vyslovování samohlásek šepem. Zde je zdrojový zvuk z hlasivek ve formě šumu. Ten generuje spojitě spektrum a vybudí poměrně přesně jednotlivé formanty. Tyto formanty však nejsou shodné s formanty při vyslovování nahlas, protože při mluvení šepem jsou hlasivky stále otevřené a při vyhodnocování je proto nutné počítat s jinými okrajovými podmínkami. Při mluvení šepem se první dva formanty posunují k vyšším frekvencím. První formant se posunuje výrazně více než druhý. U ostatních formantů není posun tak zřetelný.

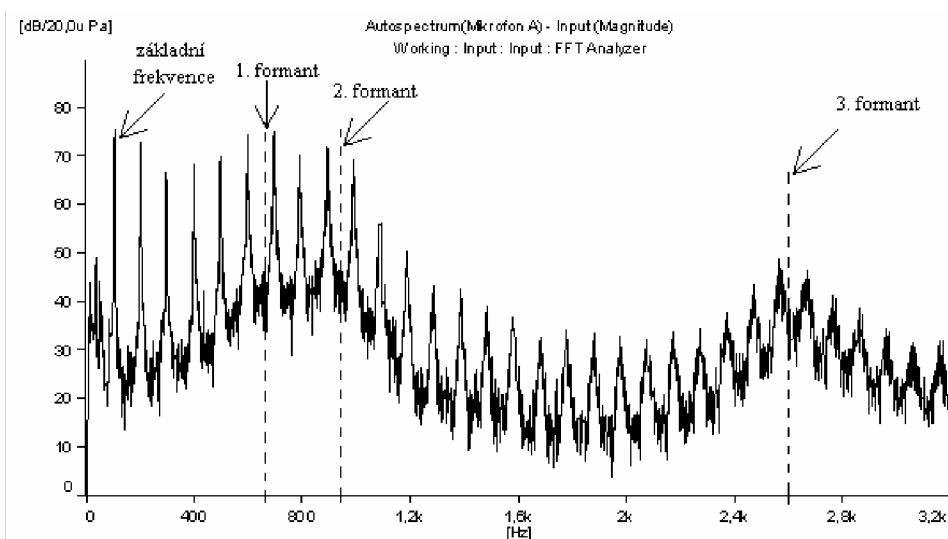
## 11. Experimentální měření skutečných vokálních traktů

Tato měření byla prováděna na skutečném lidském vokálním traktu pomocí zařízení PULSE a zpracována na stejnojmenném softwaru. Výstupem těchto měření bylo autospektrum pro zdrojový hlas uvedené na obr. 11.1. Spektra pro jednotlivé samohlásky nahlas jsou patrné z obr. 11.2 až 11.4, pro samohlásky šeptem pak jsou na obr. 11.5 až 11.7.

### 11.1 Samohlásky vyslovované nahlas

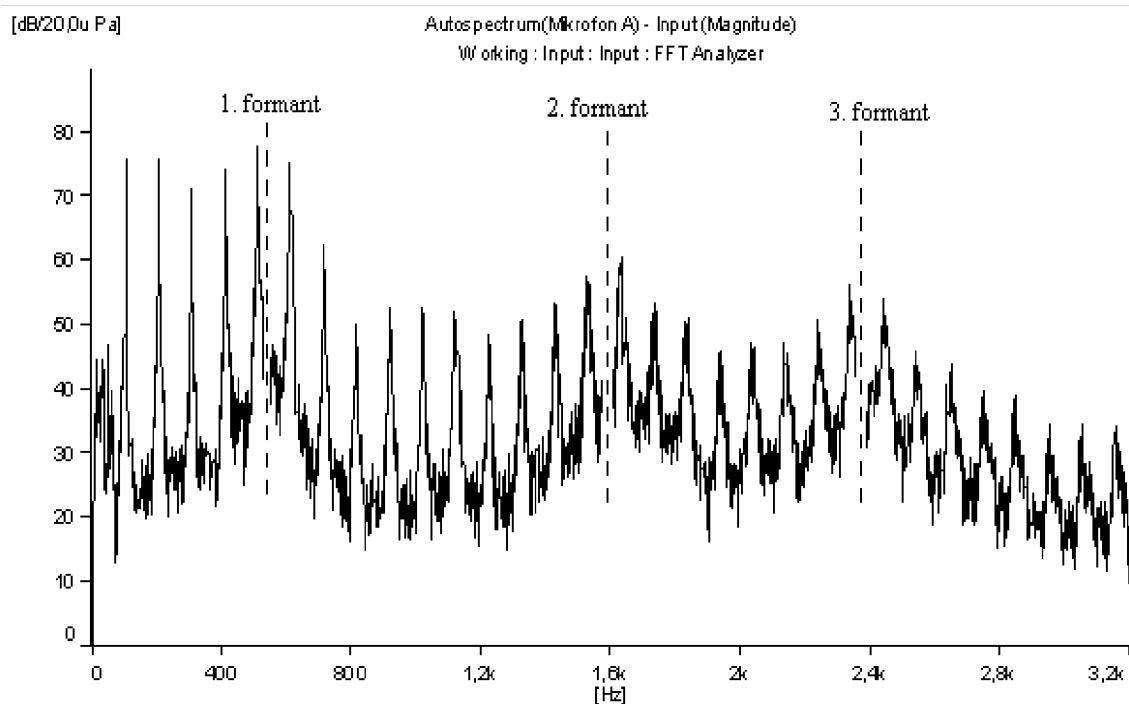


Obr. 11.1 Autospektrum zdrojového hlasu pro generování souhlásek nahlas

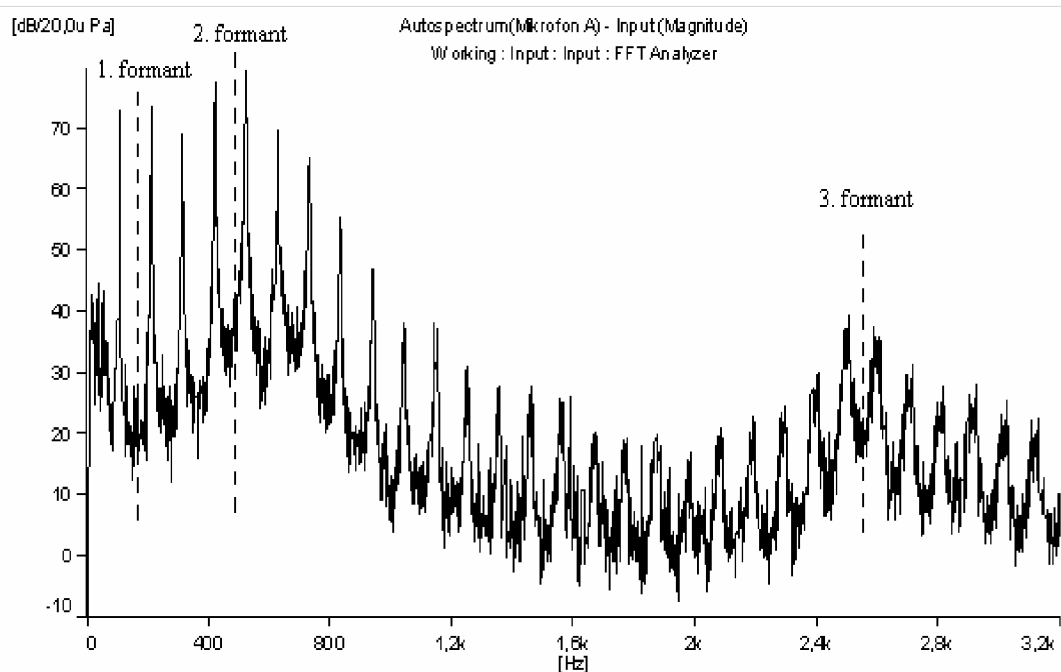


Obr. 11.2 Autospektrum samohlásky A vyslovené nahlas



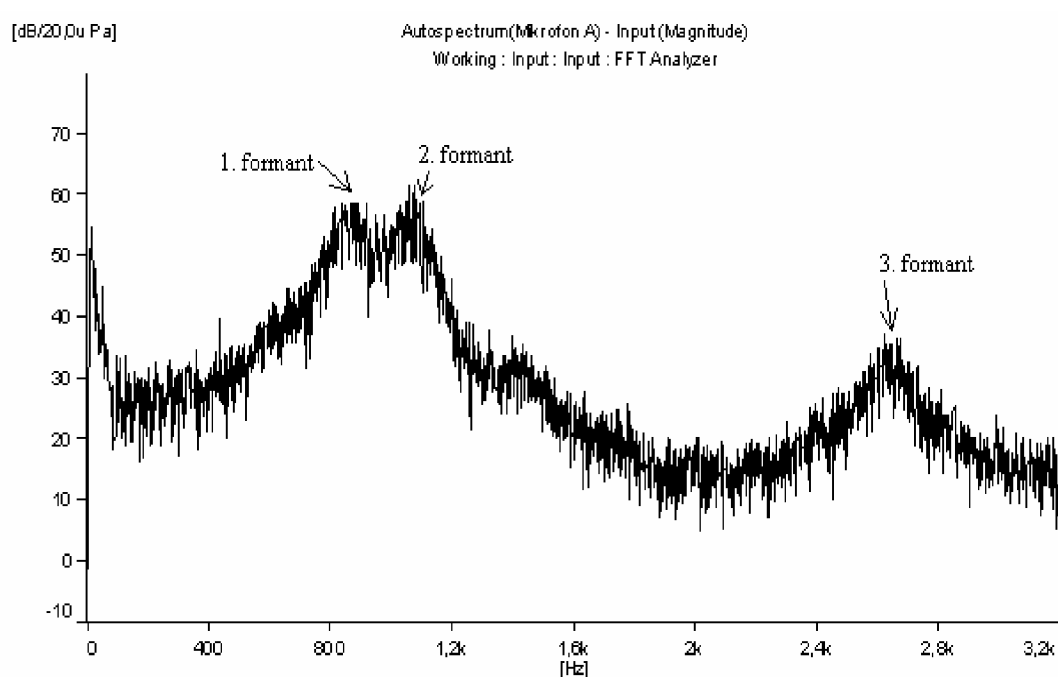


Obr. 11.3 Autospektrum samohlásky E vyslovené nahlas

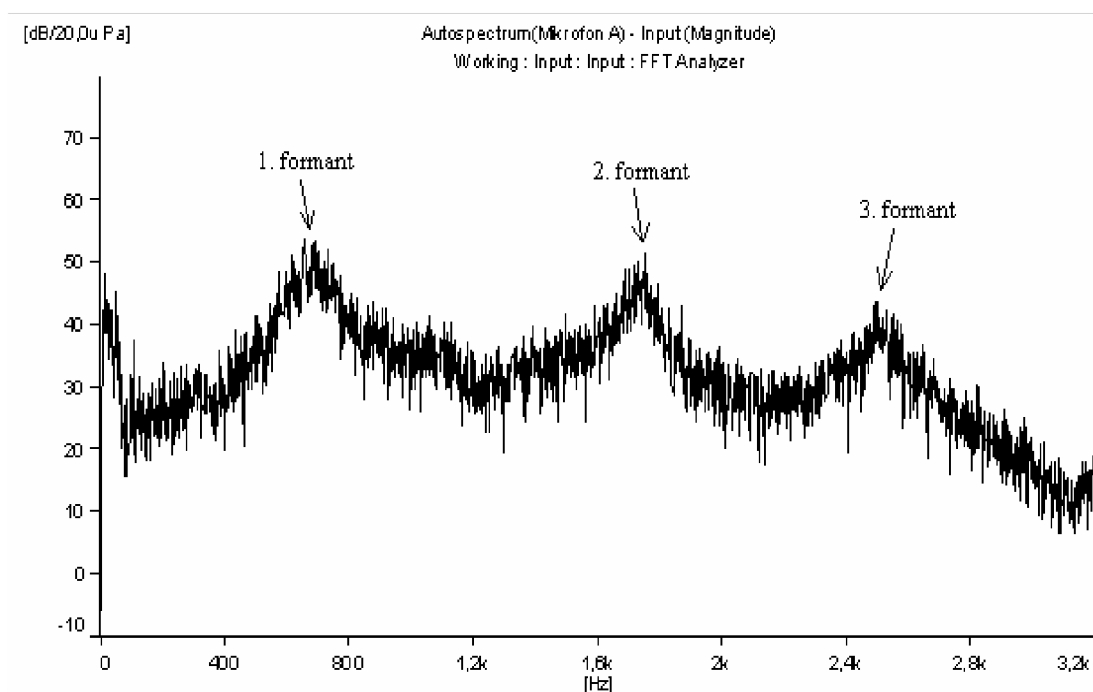


Obr. 11.4 Autospektrum samohlásky O vyslovené nahlas

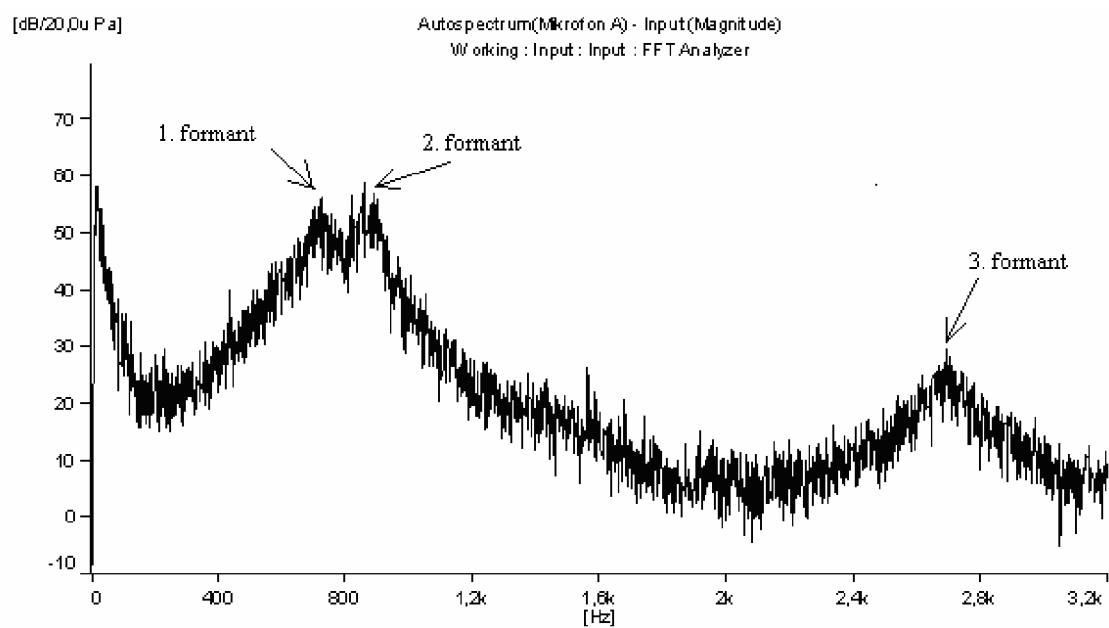
## 11.2 Samohlásky vyslovované šeptem



Obr. 11.5 Autospektrum samohlásky A vyslovené šeptem



Obr. 11.6 Autospektrum samohlásky E vyslovené šeptem



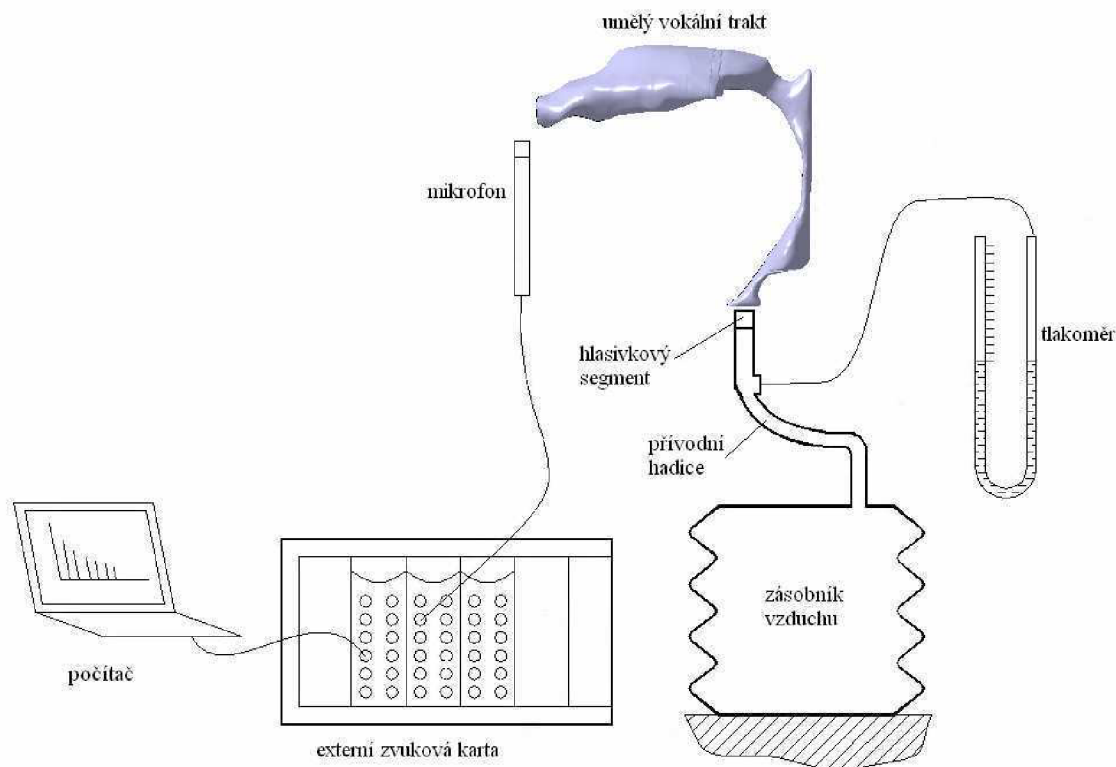
**Obr. 11.7** Autospektrum samohlásky O vyslovené šeptem

## 12. Experimentální měření umělých vokálních traktů

Tato měření byla prováděna na umělém vokálním traktu vyrobeném na Ústavu konstruování v 3D tiskárně SST 1200. Tato tiskárna používá FDM technologii (Fused Deposition Modeling), při které je model vytvářen postupným nanášením tenkých vrstev modelovacího materiálu.

Schema zapojení měřící aparatury je zobrazeno na obr. 12.1. Pro buzení byly použity hlasivkové segmenty s třemi různými průměry otvorů, které se pohybovaly v rozmezí 2,5 až 4,5 mm. Měření jsme prováděli pro různé hodnoty tlaků vytvářených zásobníkem vzduchu v rozsahu  $(1 \div 10)$  cm vodního sloupce. Tyto tlaky byly regulovány na konstantní úrovni pomocí tlakoměru. Výstupní signál byl přenášen přes měřící mikrofon do externí zvukové karty připojené k PC. Tento signál byl následně vyhodnocen v matematickém programu MATLAB v. R2006a.

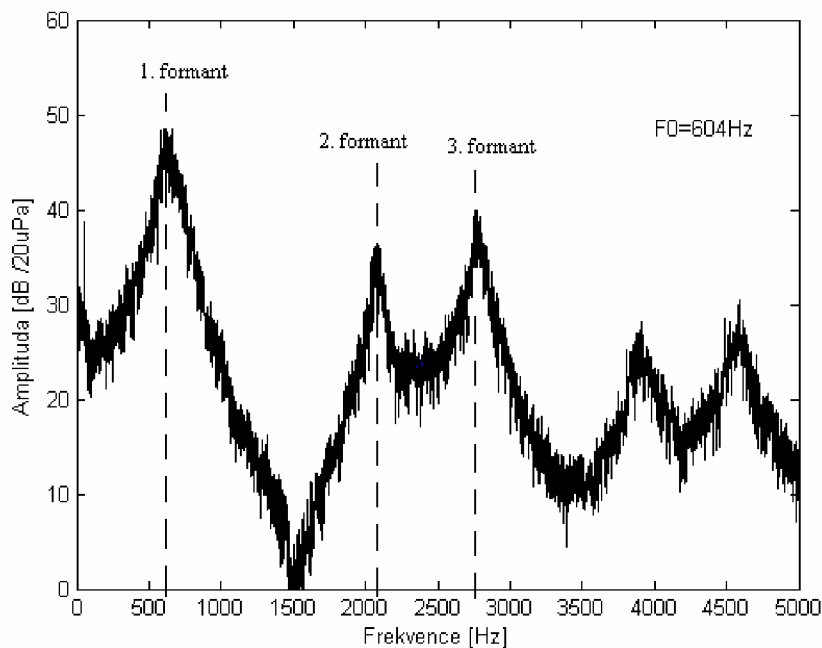
Experimentálně byla ověřena i verze, kdy je model vokálního traktu buzen umělým zdrojovým hlasem J2, podrobněji popsaná v [1]. Pro vytváření šumu byl jako akustický budič použit sluchátkový reproduktor o průměru 16 mm.



Obr. 12.1 Měřící aparatura

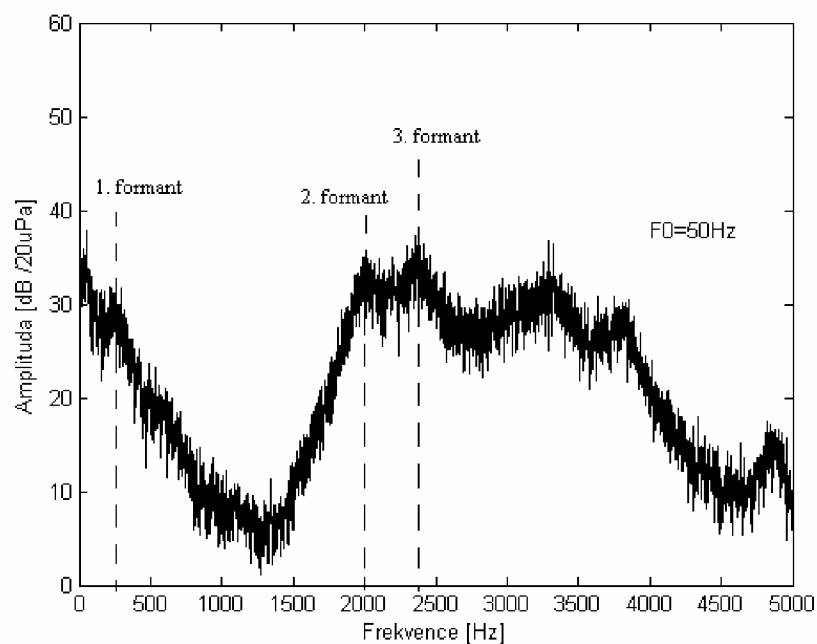
## 12.1 Měření s hlasivkovými segmenty

Na obr. 12.2 je zobrazeno spektrum pro samohlásku A vygenerovanou pomocí hlasivkového segmentu pro vyslovování šepem s otvorem 2,5 mm a tlakem 10 cm vodního sloupce. Při těchto experimentech se u samohlásky A projevilo splnutí prvních dvou formantů. To je způsobeno Zobrazený 2. formant je tedy ve skutečnosti 3. formantem a vysvětluje to rozdíl mezi rozložením formantů pro skutečný a umělý vokální trakt



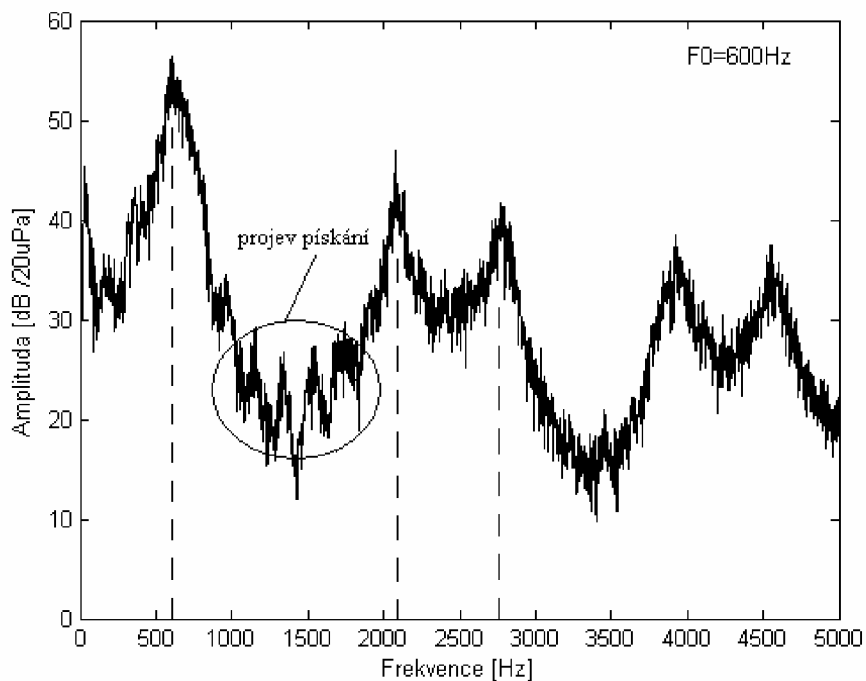
**Obr. 12.2 Samohláska A u umělého vokálního traktu**

Na obr. 12.3 je spektrum pro samohlásku I vygenerovanou pro stejný hlasivkový segment a tlak 8cm vodního sloupce.



**Obr. 12.3 Samohláska I u umělého vokálního traktu**

U některých měření se vyskytly frekvenční vrcholy, které však ve skutečnosti nejsou způsobeny rezonancemi vokálního traktu, ale pískáním přes hranu hlasivkového segmentu. To je patrné z obrázku 12.4



**Obr. 12.4 Samohláska A s projevem pískání**

## 13 Závěr

Cílem této práce bylo seznámení s akustickými vlastnostmi vokálního traktu. Pro srovnání bylo provedeno výpočtové modelování a experimenty na skutečném a umělém vokálním traktu.

Z analýz bylo zjištěno, že na srozumitelnost samohlásek při mluvení nahlas má velký vliv základní frekvence. Na této frekvenci kmitají hlasivky a je zdrojem akustického signálu. Při vysokých základních frekvencích kolem 2000 Hz nemohou být některé nižší formanty vybudeny. Velký odstup harmonických frekvencí navíc znemožňuje správné buzení formantů. Tento jev je dobře zřetelný z obr. 5.5. Tím může dojít k zaměňování jednotlivých samohlásek

Pokud známe formanty samohlásek a vhodně upravíme základní frekvence, můžeme přiblížit harmonické frekvence k prvnímu formantu, čímž se výrazně ovlivní intenzita harmonické složky a tím i dané samohlásky. V praxi se tento jevu využívá při tzv. ladění formantů.

Jinak je to při vyslovování samohlásek šeptem. Zde je zdrojovým zvukem šum, který generuje spojité spektrum. Jednotlivé formanty se proto vybudí přesně. Frekvenčně však tyto formanty neodpovídají formantům při vyslovování nahlas. To je způsobeno stálým otevřením hlasivek. Při výpočtu musíme proto počítat s jinými okrajovými podmínkami.

Výpočtovým modelováním jsme spočítali vlastní frekvence a akustické tlaky vokálního traktu pro samohlásky A a I. Výpočet byl proveden na modelech těchto samohlásek při nastavení okrajových podmínek pro mluvení nahlas a šeptem.

Podélné kmitání vzduchového sloupce se ve vokálním traktu vyskytuje při prvních čtyřech vlastních tvarech, což potvrzuje naše předpoklady. Teprve od pátého tvaru se začíná připojovat příčné kmitání, které však nemá vliv na srozumitelnost samohlásek, ale pouze na zabarvení hlasu. Vybrané vlastní tvary jsou uvedeny v kapitole 9.1.

V práci jsme sledovali také vliv velikosti vokálního traktu na formanty, uvedený v kapitole 9.2. Se zmenšováním vokálního traktu se formantové frekvence posouvaly k vyšším hodnotám a naopak. U samohlásky A se výrazně neměnila vzdálenost prvních dvou formantů. Od třetího je již posun od prvního formantu daleko zřetelnější. Při vyslovování samohlásek nahlas se vyskytuje formant, který se u mluvení šeptem nevyskytuje.

U samohlásky A je při vlivu otevření úst patrné ovlivnění v oblasti třetího a čtvrtého formantu, nelze však s určitostí říci zda v daném místě hodnota stoupá nebo klesá. U samohlásky I se zmenšujícím se ústním otvorem nepatrně klesá třetí a pátý formant.

Amplitudo-frekvenční charakteristiky samohlásek byly zjišťovány pomocí harmonické analýzy modelů vokálního traktu pro jednotkový tlak 1 Pa.

Výsledky počítačového modelování jsou celkem ve shodě s experimentálním měřením na skutečném vokálním traktu. Odlišnosti jsou způsobeny jinými rozměry kavit. Při měření umělých vokálních traktů s hlasivkovým segmentem se u samohlásky A projevuje splývání prvních dvou formantů. Při určitých hodnotách tlaků se vyskytují přídavné frekvenční vrcholy, které jsou způsobeny pískáním přes hranu hlasivkového segmentu.



## Použitý hardware a software

Pro experimentální měření samohlásek na skutečném lidském vokálním traktu byla použita aparatura PULSE. Na zpracování výstupních dat pak stejnojmenný software od firmy Bruel&Kjaer.

Pro měření vytvořených umělých vokálních traktů byl jako akustický budič pro mluvení šeptem použit sluchátkový reproduktor o průměru 16 mm. Experimentálně byla ověřena i verze, kdy je model vokálního traktu buzen umělým zdrojovým hlasem J2. Jako A/D i D/A převodník byla použita externí zvuková karta Focusrite Safire LE připojená k PC. Pro zpracování dat byl použit program MATLAB R2006a. Pro všechny měření byl jako snímač použit měřicí mikrofón Behringer ECM8000. Při experimentech byla nastavena vzorkovací frekvence na 44 100 Hz.

Výpočtové modelování bylo prováděno ve výpočtovém programu ANSYS v.11 na současném hardwarovém vybavení počítačové učebny ÚMTMB.

## **Přílohy**

Na přiloženém CD jsou obsaženy následující přílohy:

Elektronická verze diplomové práce      Diplomová práce.pdf

## Použitá literatura

- [1] Mišun V.: Tajemství lidského hlasu. Vydavatelství VUTIUM Brno, 2008.
- [2] Mišun V.: Vibrace a hluk. Skripta VUT Brno, 2005
- [3] Titze I.R.: Principles of voice production. Prentice Hall, 1994
- [4] Hrůza V.: Analýza formantů českých samohlásek. Diplomová práce, 2003
- [5] Švec J.: Studium mechanicko-akustických vlastností lidského hlasu. Disertační práce, Olomouc, Univerzita Palackého, 1996
- [6] Hostička P., Mišun V.: Experimentální určování polohy formantů při různém vokálním projevu. Proc. Interaction and Feedbacks 2001, Praha, 2001
- [7] Ryalls J., Behrens S.: Introduction to speech science, Allyn & Bacon, 2000
  
- [8] History and Development of Speech Synthesis, dostupné na:  
[http://www.acoustics.hut.fi/publications/files/theses/lemmetty\\_mst/chap2.html](http://www.acoustics.hut.fi/publications/files/theses/lemmetty_mst/chap2.html)
- [9] Krčmová M.: Artikulační fonetika, dostupné na:  
<http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch05.html>
- [10] Vocal fold, dostupné na: <http://content.answers.com/main/content/wp/en-commons/thumb/5/5e/400px-Gray1204.png>
- [11] Speech analysis tutorial, dostupné na:  
<http://www.ling.lu.se/research/spechtutorial/tutorial.html>
- [12] Akustika, vznik a šíření zvuku, frekvenční analýza a syntéza, dostupné na:  
[http://homen.vsb.cz/~ber30/texty/varhany/anatomie/pistaly\\_akustika.htm](http://homen.vsb.cz/~ber30/texty/varhany/anatomie/pistaly_akustika.htm)
  
- [13] Ansys manual