

MODELLING OF PRODUCTION OF CZECH FORMANTS BY VOCAL FOLDS MODEL

K. Přikryl*, V. Mišun*

Summary: *Acoustic speech output in humans is commonly considered to result from a combination of a source of sound energy (e.g. the larynx) modulated by a transfer (filter) function determined by the shape of the supralaryngeal vocal tract. Finite element (FE) models were used to model vocal tracts. Shapes used in this case were obtained using the magnetic resonance imagine (MRI) technique. Different vocal tract are shown corresponding to the czech vowels /a/,/i/,/o/,/u/. Acceleration of the air column at the point of the glottis was the source of the motion. The output spectrum is the product of the glotal flow spectrum and frequency response of the vocal tract.*

1. Úvod

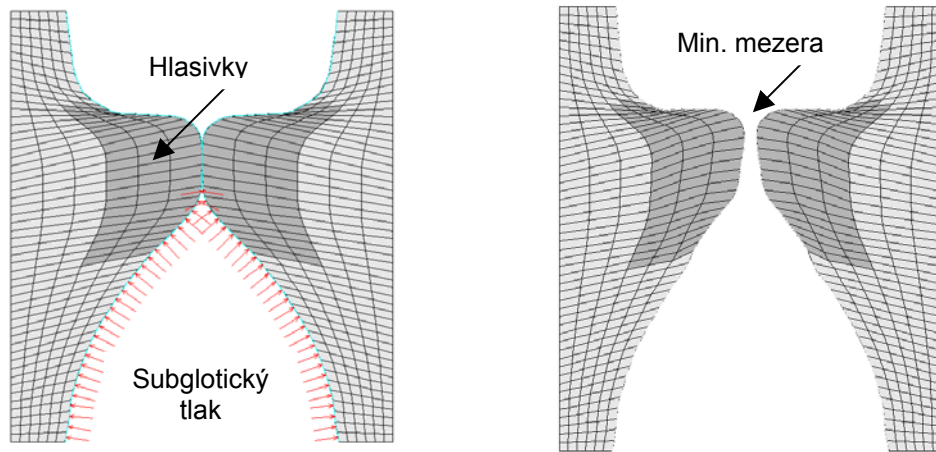
Vytváření samohlásek je podmíněno tvarem kavity vokálního traktu a jemu odpovídajících formantů (rezonančních frekvencí) a při mluvení nahlas také zdrojem, kterým jsou pohybuující se hlasivky. Podhlasivkový tlak vzduchu při mluvení nahlas je přetvářen hlasivkami na periodické bublinové pulsy (Mišun 2001). Pulsy vzduchu jsou dále akusticky filtrovány vokálním traktem. Znamená to, že při tomto procesu se zvýrazňují některé amplitudy frekvenčních složek zdrojového signálu, což má za následek vytvoření slyšitelné samohlásky. V Přikryl (2003) byl popsán proces „zdroj-filtr“ při vytváření samohlásek analyticky popsaným zdrojem, který byl vytvořen na základě objemového průtoku vzduchu hlasivkovou mezerou (Burnett G.C). Uvažuje se, že v časové oblasti je filtrem plocha průřezu vokálního traktu v závislosti na vzdálenosti od hlasivkové mezery. Proud vzduchu ve vokálním traktu je potom transformován na výstupní tlak v oblasti úst. Ve frekvenční oblasti spektrum odezvy je násobek spektra zdroje a frekvenční odezvy vokálního traktu.

2. Modelování vytváření pulsů hlasivkami

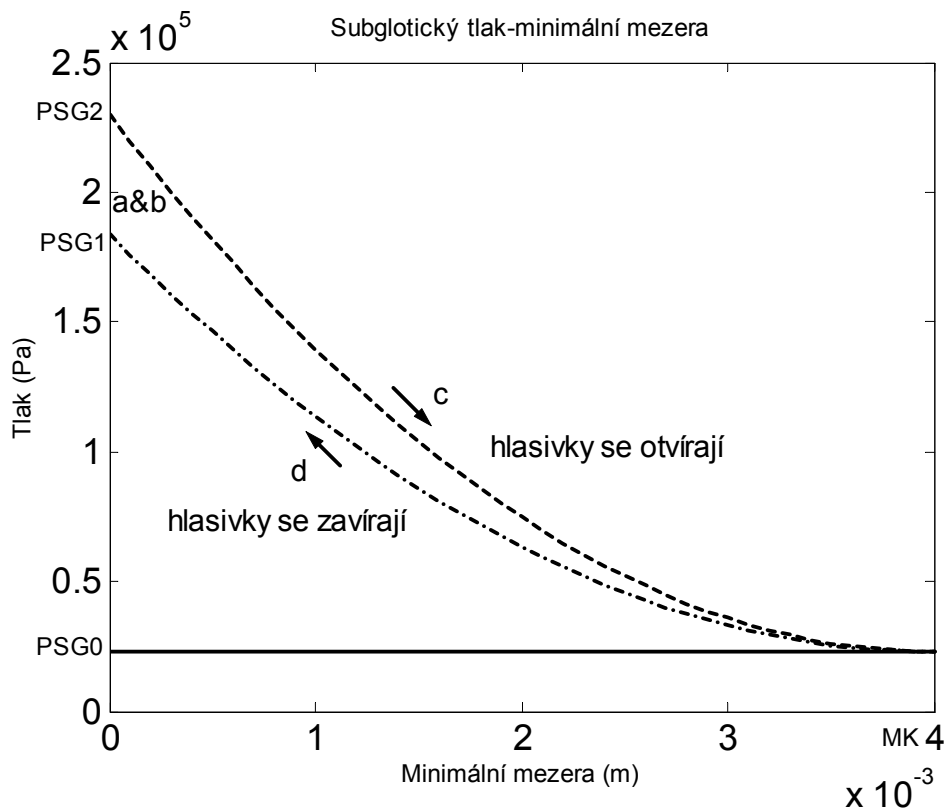
V literatuře Titze (1994) byly popsány různé mechanické modely hlasivek. V práci Mišun (2001) byla provedena definice hlasivek na základě teorie „bubliny“. Hlasivky jsou modelovány pomocí metody konečných prvků (MKP) a jejich rovinný model spolu s částí hrtanu je na obr.1. Obě části jsou modelovány s různými moduly pružnosti podle Přikryl & Mišun (2001). Na obr. 1 jsou odlišeny odstínem šedé barvy. Pro funkci hlasivek (vytváření „bublin“) byl definován subglotický tlak, který je funkcí minimální hlasivkové mezery a jeho grafická závislost je na obr. 2. Jsou zde vyznačeny větve otevírání (větve c), zavírání (větve d) a zavřeno (větve a & b). Body PSG2 a PSG1 jsou podhlasivkové tlaky v okamžiku otevření

* Doc.Ing.Karel Přikryl,CSc., *Doc.Ing.Vojtěch Mišun,CSc.,VUT Brno, FSI, ÚMT, Brno 616 69, tel.: +420 5 4114 2861, fax.: +420 5 41142876, e-mail: prikryl@umtn.fme.vutbr.cz, misun@umtn.fme.vutbr.cz

hlasivek, resp. uzavření hlasivek v již probíhajícímu cyklu vytváření „bublin“, PSG0 je stálý podhlasivkový tlak.



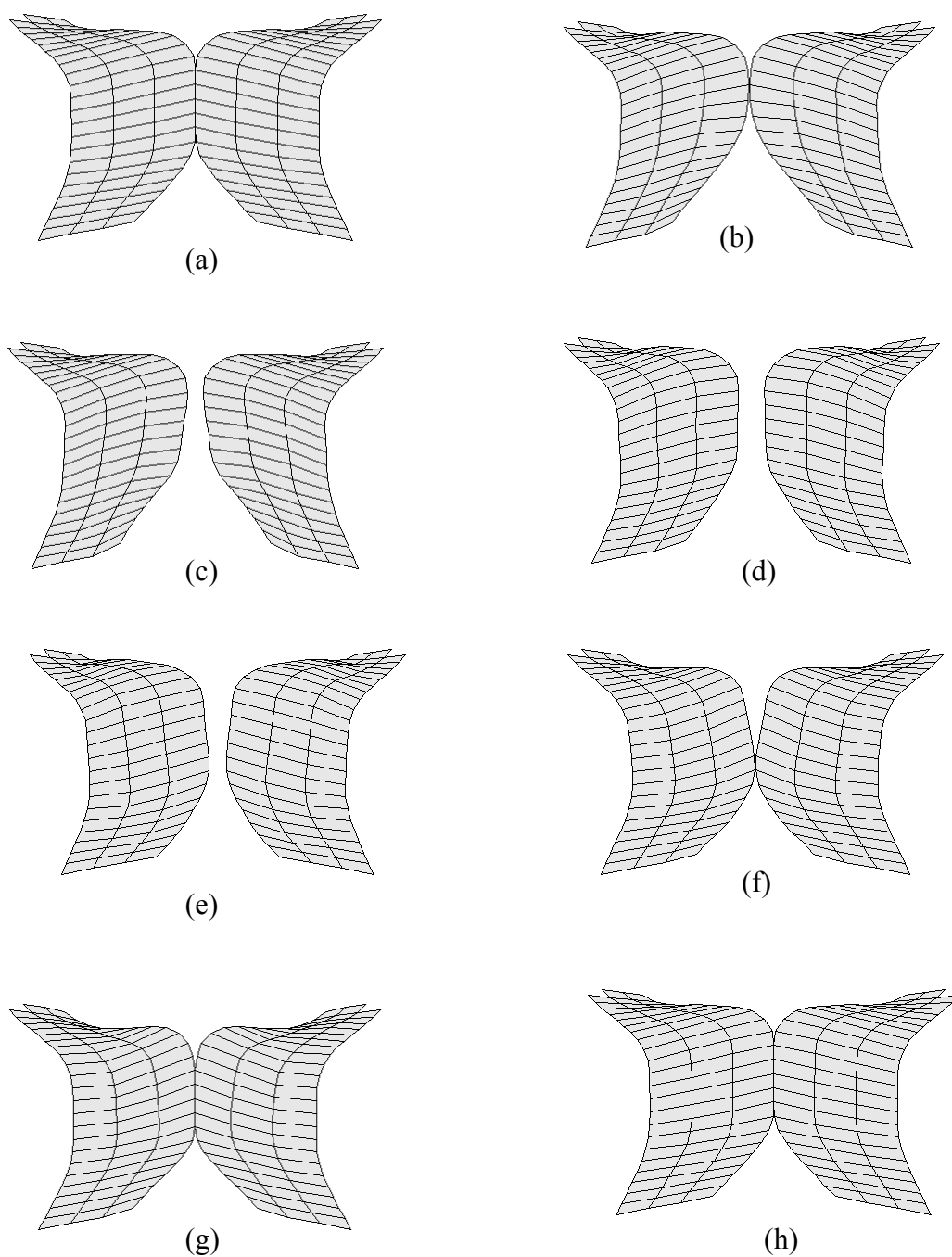
Obr. 1 Rovinný model hlasivek se subglotickým tlakem a minimální hlasivkovou mezerou



Obr. 2 Subglotický tlak PSG v závislosti na minimální hlasivkové mezeře

Tlak PSG2 je určen výpočtem. Je to tlak, při kterém jsou hlasivky právě otevřeny. Je funkcí předpětí na začátku fonace, kdy jsou hlasivky u sebe (zavřeno). V řešeném případě je vypočteno $PSG2=230000$ Pa. Tlaky PSG1 a PSG0 jsou voleny jako procentuální část vypočítaného PSG2. V modelu bylo stanoveno $PSG1=184000$ Pa, a $PSG0=23000$ Pa.

Na obr. 3 jsou zobrazeny fáze pohybu hlasivek modelované předepsaným subglotickým tlakem podle obr.2.



Obr. 3 Fáze pohybu hlasivek v koronální rovině

Na obr. 3a jsou hlasivky v kontaktu. Je to počáteční fáze a není žádný průtok vzduchu. S rostoucím subglotickým tlakem se hlasivky začínají otevírat podle obr. 3b. Dosáhne-li tlak hodnotu PSG_2 , hlasivky podle obr. 3c ztratí kontakt, otevřou se. Nastává fáze otevírání

hlasivek a vytváření „bubliny“ vzduchu. Otevírání hlasivek pokračuje až k hodnotě nejvýše rovné předpokládané hodnotě MK (obr. 3d) se současným vytlačováním vzduchu. Dále se hlasivky začnou zavírat podle obr. 3e až 3f a vytlačovat „bublinu“. Po vytlačení „bubliny“ nastávají závěrečné fáze podle 3g, a hlasivky se uzavřou úplně a pohyb se cyklicky opakuje. Z obr. 3 je zřejmé, že výchylky ve spodní a horní části hlasivek nejsou ve fázi. Spodní část předbíhá část horní. Vzniká tak mukózní (slizniční vlna). Cyklický pohyb hlasivek je vyvolán vratnými pružnými a setrvačnými účinky hlasivek.

V průběhu numerického řešení pomocí programového souboru ANSYS byla prováděna přechodová analýza a zaznamenáván průběh subglotického tlaku PSG-min.mezera v časové oblasti. Výsledek je na obr. 4. Byla řešena maticová rovnice

$$[M]\{\ddot{u}\} + [B]\{\dot{u}\} + [K]\{u\} = \{f(t)\} \quad (1)$$

kde

M je matice hmotností

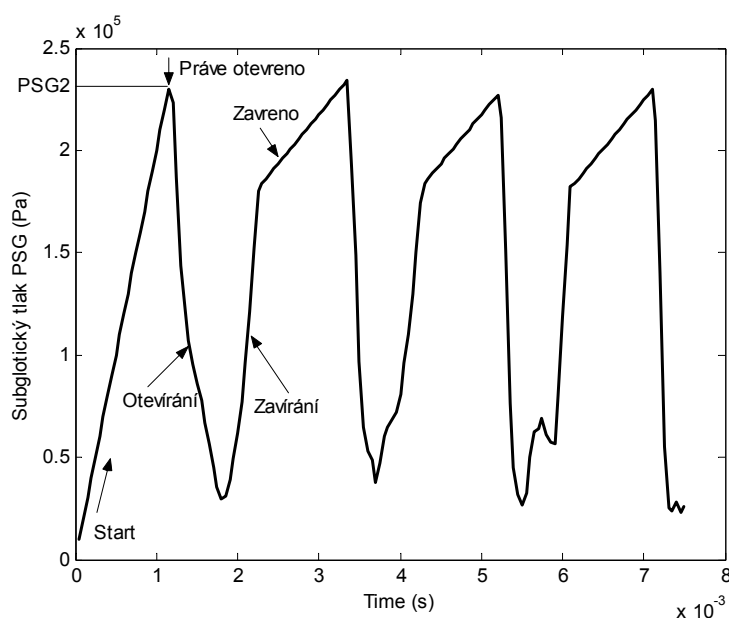
B je matice tlumení

K je matice tuhostí

\ddot{u}, \dot{u}, u jsou vektory zrychlení, rychlostí a výchylek

$f(t)$ je vektor zatížení

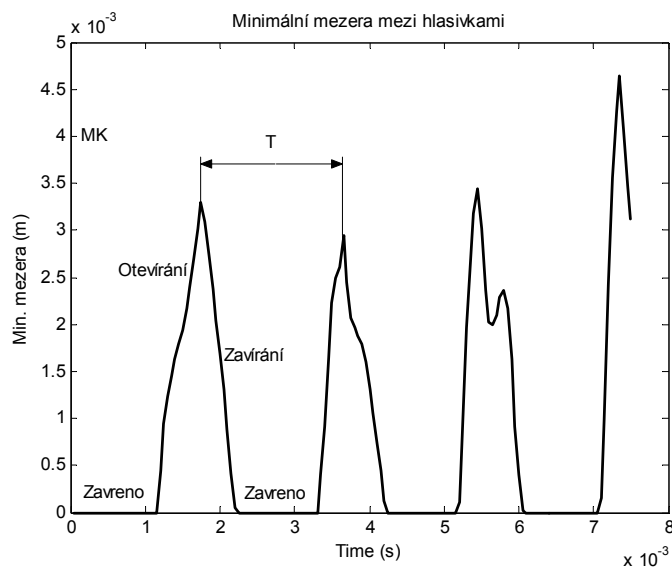
Tlumení bylo zahrnuto jako proporcionalní, úměrné matice hmotnosti



Obr. 4 Subglotický tlak PSG(min. mezera) v závislosti na čase

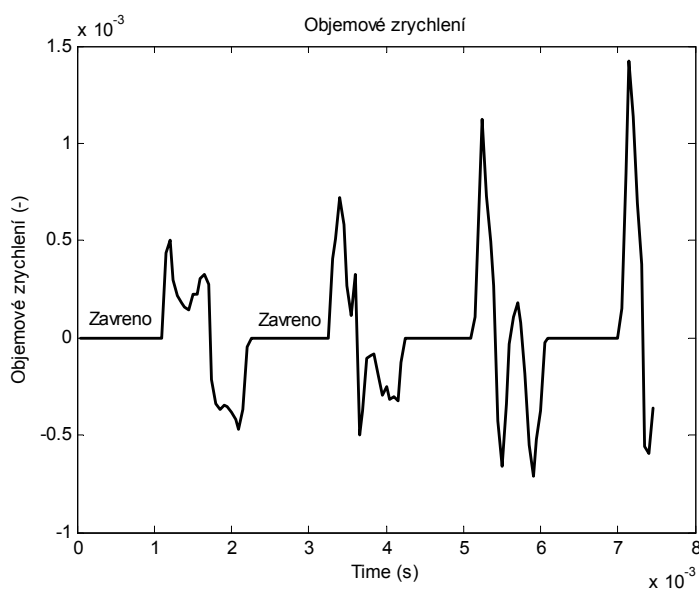
Je zřejmé, že důležité charakteristiky fonace jako je základní frekvence, kvalita hlasu, amplituda vibrací apod. se mění s materiálovými konstantami modelu a předpětím v počáteční fázi. Objemová rychlost Q_v , je závislá na příčné ploše hlasivkové štěrbině A_g .

Protože v modelu není zahrnuta inertance vokálního traktu, neboť není uvažován vzduchový sloupec, je podle B.H.Story (2000) časová závislost objemové rychlosti Q_v a hlasivkové štěrbiny A_g stejná. Plocha hlasivkové štěrbiny se v modelu stanoví z minimální mezery mezi hlasivkami násobená délkou hlasivek. Protože je model rovinný, délka hlasivek je rovna jedničce. Tvar průběhu A_g je tedy roven časové závislosti minimální mezery mezi hlasivkami, která je na obr. 5.



Obr. 5 Minimální mezera mezi hlasivkami resp. průběh vlny A_g v závislosti na čase

Derivací objemové rychlosti dQ_v/dt resp. průběhu (dA_g/dt) obdržíme průběh objemového zrychlení, kterým budeme budít model vokálního traktu. Soustava je lineární a záleží pouze na tvaru a spektru budících účinků. Výsledek je na obr. 6.



Obr. 6 Průběh vlny objemového zrychlení dQ_v/dt , kterým se zatíží model traktu

3. Model vokálního traktu

Pro modelování vokálního traktu byly použity konečnoprvkové modely získané pomocí techniky magnetické rezonance viz. Kršek (2000). Akustický prostor byl modelován pomocí 3D akustických elementů s těmito materiálovými vlastnostmi : hustota $\rho_0 = 1.2 \text{ kgm}^{-3}$, rychlost zvuku $c_0 = 353.2 \text{ ms}^{-2}$. Otevřená ústa jsou modelována pomocí okrajové podmínky pro akustickou impedanci otvorů kavit podle „Levin & Schwingerovy formule (SYSNOISE) :

$$Z = \left(0.24(ka)^2 + j*0.56(ka)\right)\rho c \quad (3)$$

kde Z je komplexní impedance, k je vlnové číslo a je poloměr otevření. Tento vztah platí pokud je (ka) menší než 1. Při řešení přechodové analýzy je povolena pouze reálná část komplexní impedance. Žádné jiné okrajové podmínky při modelování nebyly zavedeny. Byla modelována produkce samohlásek /a/, /i/, /o/, /u/, jejichž konečnoprvkové modely jsou na obr. 7.



Obr. 7 Konečnoprvkové modely vokálních traktů při vyslovení /a/, /i/, /o/, /u/

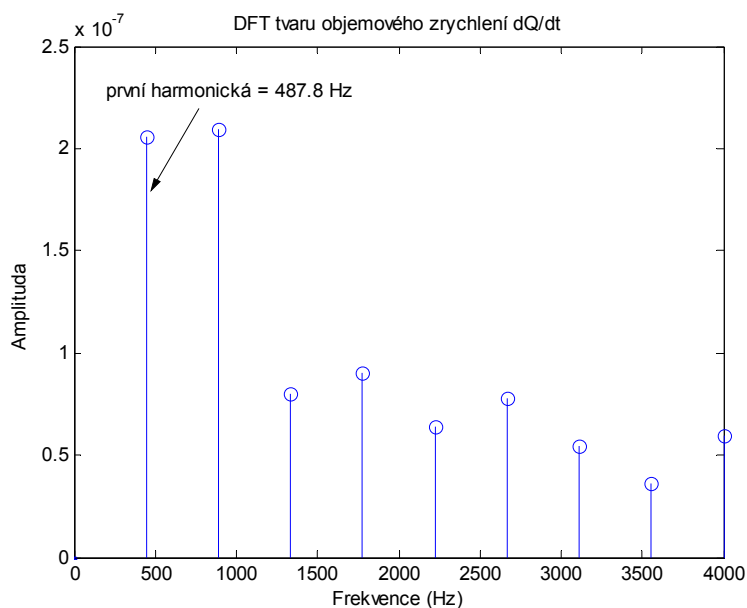
4. Přechodová analýza

Přechodová analýza je provedena na modelech kavit podle obr. 7 podle rovnice (4) s následující spektrální analýzou podle teorie filtrů.

$$[M]\{\ddot{p}\} + [B]\{\dot{p}\} + [K]\{p\} = \{F_A\} \quad (4)$$

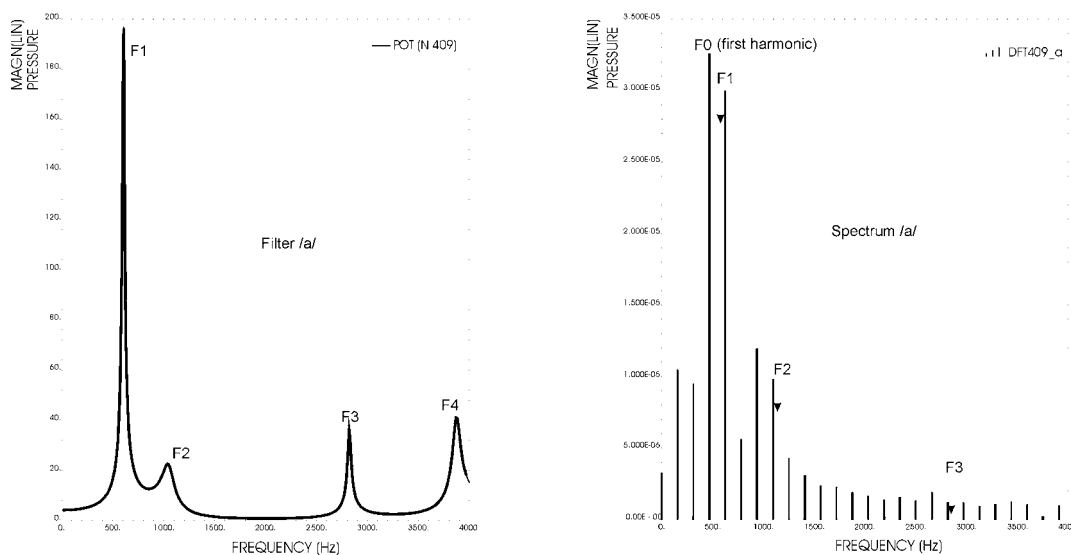
$\{F_A\}$ je vektor sil, které jsou úměrné normálovému zrychlení, které působí na povrchy akustických elementů kavity, $\{p\}$ je akustický tlak a matice tlumení $[B]$ vznikne v důsledku impedance v místě otevřených úst. V časové oblasti se k řešení užívá Newmarkova metoda a výsledkem je průběh akustických tlaků v místě úst. Ve frekvenční oblasti je spektrum odezvy násobek spektra zdroje a frekvenční odezvy vokálního traktu.

Spektrum zdroje obdržíme pomocí Diskrétní Fourierovy Transformace (DFT) průběhu objemového zrychlení dQ_v/dt , jehož tvar je na obr. 6. Výsledek DFT je na obr. 8.



Obr. 8 Spektrum zdroje v místě hlasivek

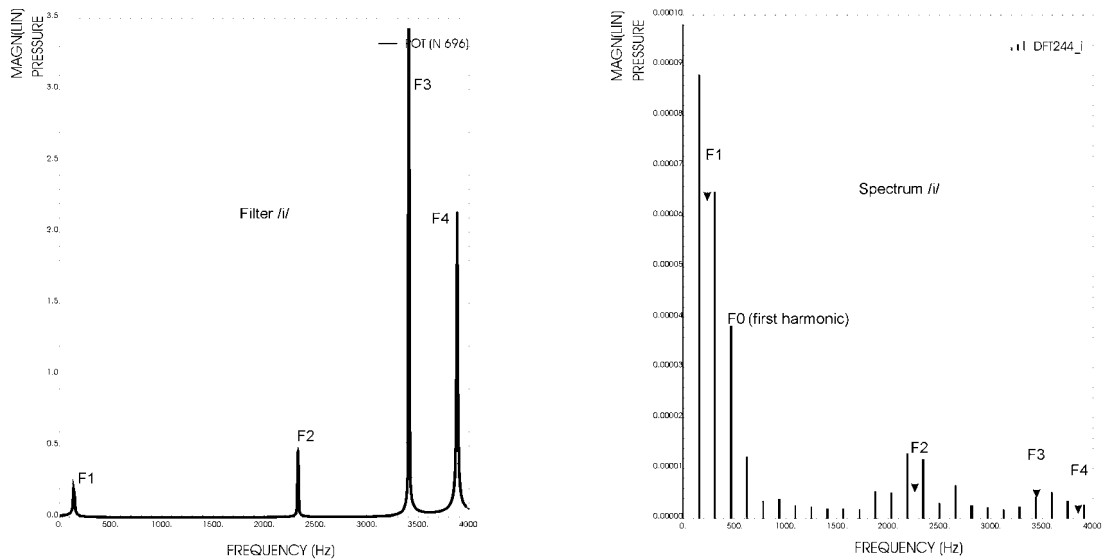
Z obr. 8 je vidět, že první harmonická zdroje v místě hlasivek jejichž funkce je modelovaná pomocí „bublín“ je 487.8 Hz. Je to frekvence hlasivek F_0 . Je funkcí předpětí a materiálových konstant modelu hlasivek.



Obr. 9 Filtr samohlásky /a/ a její spektrum na výstupu z vokálního traktu v místě úst

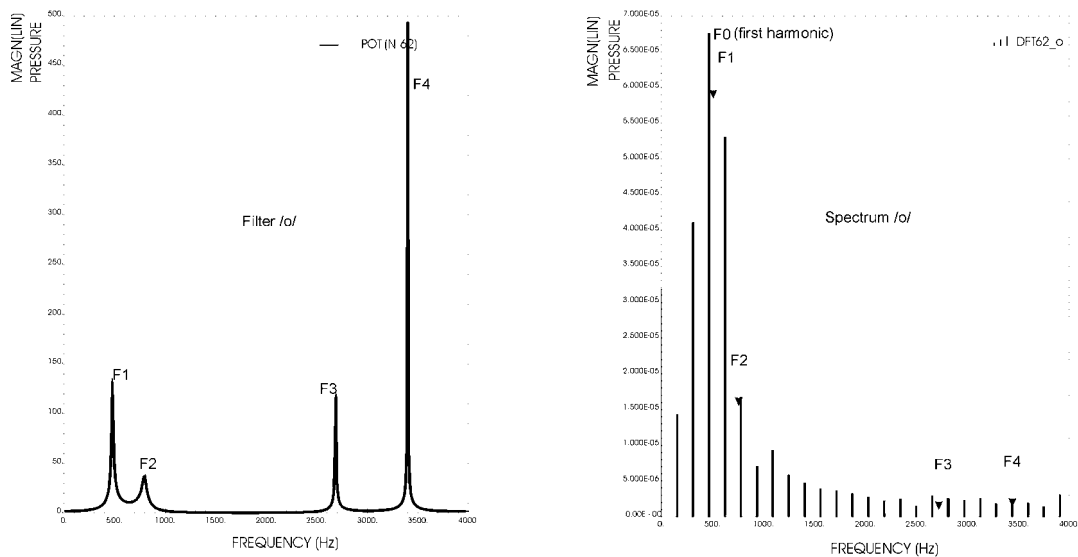
Vlastní frekvence vokálního traktu (formanty) při samohlásce /a/ jsou podle obr. 9. $F_1=612$ Hz, $F_2=1127$ Hz, $F_3=2860$ Hz a $F_4=3990$ Hz. Tyto frekvence mají podle teorie výrazně zesilovat harmonické složky zdroje z obr. 8. I přesto, že žádná z nich není shodná s žádnou harmonickou zdroje, dochází k zesílení a to nejvýrazněji první harmonické, která je zesílena

formantem F1 a druhá harmonická formantem F2. Harmonické složky zdroje jsou na obr. 9. vyznačeny tučně. Podobně byla vytvořena samohláska /i/. Výsledky analýzy jsou na obr. 10.



Obr. 10 Filtř samohlásky /i/ a její spektrum na výstupu z vokálního traktu v místě úst

První formant samohlásky /i/ je $F1=246$ Hz, a druhý formant $F2=2275$ Hz jsou podle obr. 10 výrazně vzdáleny. Formanty $F3=3431$ Hz a $F4=3844$ Hz jsou blízko u sebe a mohou se vzájemně ovlivňovat. První harmonická složka zdroje nemůže být filtrem /i/ zvýrazněna, protože je velmi vzdálena od prvního formantu F1. Rozdíl činí podle obr.10 až 221 Hz.

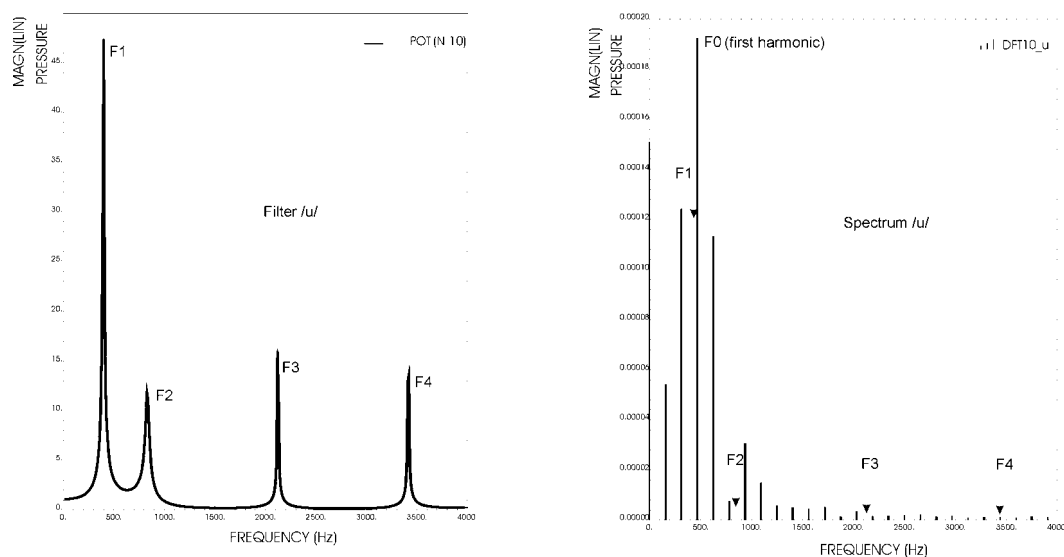


Obr. 11 Filtř samohlásky /o/ a její spektrum na výstupu z vokálního traktu v místě úst

Filtř další samohlásky /o/ byl získán jako u všech samohlásek harmonickou analýzou v požadovaném rozsahu frekvencí od 0 Hz do 4000 Hz. U každé samohlásky byla použita amplituda rychlosti 1ms^{-1} . Z harmonické analýzy vyplynuly pro samohlásku /o/ vlastní frekvence $F1=516$ Hz, $F2=798$ Hz, $F3=2721$ Hz a $F4=3437$ Hz. První harmonická zdroje je

487.8 Hz. První formant samohlásky /o/ je 516 Hz. Obě frekvence jsou velmi blízké a proto na obr. 10 je tato harmonická velmi zvýrazněna. Je vidět, že druhá vlastní frekvence $F_2=798$ Hz zvýrazňuje druhou harmonickou zdroje. Ostatní harmonické mají klesající tendenci.

Jako poslední byla provedeno modelování vytváření samohlásky /u/. Spektrum zdroje bylo opět podle obr. 8. Grafy jsou na obr. 12.



Obr. 12 Filtr samohlásky /u/ a její spektrum na výstupu z vokálního traktu v místě úst

Z harmonické analýzy kavity samohlásky /u/ vyplývá, že její formanty jsou $F_1=440$ Hz, $F_2=850$ Hz, $F_3=2135$ Hz a $F_4=3416$ Hz. První harmonická zdroje je výrazně zvýrazněna prvním formantem $F_1=440$ Hz. Druhá vlastní frekvence $F_2=850$ Hz zvýrazňuje druhou harmonickou zdroje.

Ačkoliv všechny samohlásky jsou buzeny stejnou základní frekvencí $F_0=487.8$ Hz, jejich spektra se v důsledku různých tvarů vokálního traktu na výstupu liší. Spektrum na výstupu je primárně určeno rezonančními frekvencemi kavit a tedy jejich tvarem při artikulaci. Harmonické složky zdroje jsou v blízkosti úst zesilovány rezonančními frekvencemi a ty složky, které jsou dále od rezonančních ztrácejí energii. Např. pro samohlásku /o/ na obr. 11. je první harmonická zdroje 487.8 Hz a první vlastní frekvence 516 Hz. Protože jsou velmi blízké, je zvýraznění velmi značné. Z průběhů spekter na ostatních obrázcích je také vidět, že vlastní frekvence ne vždy souhlasí s harmonickými zdroje. Pro lepší shodu u ostatních samohlásek by bylo třeba ladit model hlasivek.

5. Závěr

Primárním zdrojem při vyslovování samohlásek nahlas jsou periodické pohyby hlasivek. Jejich model, který byl vytvořen pomocí konečných prvků byl zatěžován subglotickým tlakem v závislosti na minimální mezeře mezi pohybuujícími se hlasivkami. Závislost je zobrazena na obr. 2. Tímto předpisem se podařilo dosáhnout toho, že fáze hlasivek a jejich celkový pohyb jsou velmi věrné pohybu pozorovanému stroboskopem. Vokální trakty byly modelovány pomocí modelů, které byly získány pomocí techniky magnetické resonance. Pro přechodovou analýzu a realizaci vytváření samohlásek pomocí filtru vokálního traktu byl

použit programový soubor SYSNOISE. Průběh vypočítávané minimální mezery mezi hlasivkami (ANSYS), při zatěžování subglotickým tlakem je na obr. 5. Tento průběh vlny byl využit k výpočtu příčného průřezu mezihlasivkové plochy A_g , která je v případě rovinného modelu bez uvažování media (vzduchu) podobná průběhu objemové rychlosti Q_v . Derivací Q_v obdržíme tvar vlny objemového zrychlení, které bylo použito ke zrychlování elementů modelů vokálních traktů v místě hlasivek. Průběh je na obr. 6. Vokální trakt je akustický systém zatížený na výstupu radiační impedancí. Ta je zahrnuta pomocí „Levine and Schwinger“ vztahem.

V teorii filtrů je třeba určit spektrum zdroje. To je provedeno pomocí diskretní Furierovy transformace vlny zdroje v místě hlasivek. Výsledek je na obr. 8. Z obrázku plyne, že základní frekvence hlasivek F_0 je 487.8 Hz. Harmonické jsou jejím násobkem. Tyto harmonické jsou zvýrazňovány filtrem, kterým je přenosová funkce vokálního traktu. V oblasti úst potom vznikne tlaková vlna se spektry, která odpovídají jednotlivým samohláskám. Lze konstatovat, že modelování vytváření samohlásek pomocí definovaného modelu odpovídá známým spektrům řešených samohlásek.

6. Poděkování

Autoři děkují Doc. MuDr. Petrovi Krupovi z nemocnice u Sv. Anny v Brně za umožnění zobrazování vokálních traktů pomocí Magnetické resonance. Práce vznikla za podpory projektu GAČR čís.: 106/98/K019 Matematicko-fyzikální modelování vibroakustických systémů důležitých v biomechanice hlasu a sluchu, a za podpory VZ.: MSM 262100001 s názvem Výpočtové a fyzikální modelování inženýrských problémů termofluidní mechaniky, mechaniky těles a fázových přeměn.

7. Literatura

- Burnett G.C.: The physiological basis of glottal electromagnetic micropower sensors (GEMS) and their use in defining excitation function for the human vocal tract.
<http://speech.lnl.gov/thesis/index.htm>
- Kršek P.: Design of FE models of vocal tract for Czech vowels. In Interaction and Feedbacks 2000, Institute of Thermomechanics ASCR, Prague, 28-29. Nov.2000, pp.103-110. (in Czech)
- Mišun V.: The principle of pressure air bubble within the fonation, Proc.Conf. Interaction of dynamic systems, Prague, 2001
- Přikryl K.: Numeric modelling of production of Czech vowels, 4th International conf. Mechatronic, Robotic and Biomechanics 2003, Hrotovice, March 2003
- Přikryl K., Mišun V.: Phonation simulation of 2D vocal folds model using the finite-element method. Proc. 3th International conf. Mechatronic, Robotic and Biomechanics 2001, Třešť, September 2001
- Story B.H.: An overview of the physiology, physics and modeling of the sound source for vowels. Acoustic. Sci & Tech. 23, 4 (2000)
- SYSNOISE, User's manual, Leuven, Belgium 1995
- Titze I.R. Principles of voice production, Prentice Hall, Engelwood Cliffs, New Jersey 07632, 1994