

AKUSTICKÉ LISTY

České akustické společnosti

ročník 10, číslo 1

březen 2004

Obsah

Usnesení valné hromady České akustické společnosti	3
8. duben 2004 – Den „uvědomění si“ hluku <i>Jan Stěnička</i>	5
Pozvánka na přednášku Time-Frequency Analysis for Acoustic Signals	5
Pozvánka na WORKSHOP 2004 <i>Libor Husník</i>	5
Poznátky z porovnávacího měření v interferometru Experiences from round-Robin test in an impedance tube <i>Milan Krňák</i>	6
Prosodic analysis and manipulation demonstrated on the Praat software Prozodická analýza a manipulace v programu Praat <i>Tomáš Duběda and Radek Skarnitzl</i>	12

ČESKÁ AKUSTICKÁ SPOLEČNOST

Usnesení valné hromady České akustické společnosti, konané dne 29. ledna 2004 v prostorech Fakulty elektrotechnické ČVUT

Valná hromada ČsAS bere na vědomí:

1. zprávu o činnosti Rady ČsAS;
2. zprávy o činnosti jednotlivých odborných skupin a o jejich dalším zaměření;
3. zprávu o přípravách společnosti na pořádání kongresu Inter-noise 2004 v Praze a vyzývá členy k účasti;
4. zprávu o výsledcích revize hospodaření společnosti;
5. výsledky voleb do Rady společnosti a výsledky voleb předsedů odborných skupin;
6. zprávu o usnášední schopnosti valné hromady, 105/78;
7. informace o plnění úkolů a poslání Akustických listů.

Pro funkční období roku 2004 byli v jednotlivých odborných skupinách zvoleni:

- A. Obecná a lineární akustika
předseda – BEDNAŘÍK zástupce – ČERVENKA
- B. Ultrazvuk a akustické emise
předseda – ŠLEGROVÁ zástupce – PLOCEK
- C. Hluk a vibrace
předseda – STRNAD zástupce – HELLMUTH
- D. Prostorová, stavební a urbanistická akustika
předseda – MELLER zástupce – RYNDOVÁ
- E. Zpracování a záznam akustických signálů
předseda – SALAVA
- F. Psychoakustika, fyziologická akustika a akustika hudby a řeči
předseda – ŠTĚPÁNEK
- G. Elektroakustika
předseda – KEŠNER zástupce – SÝKORA

Valná hromada ČsAS schvaluje:

1. zprávu o činnosti Rady za kalendářní rok 2003 a uděluje Radě absolutorium;
2. zprávu o hospodaření společnosti za kalendářní rok 2003;
3. výši členských příspěvků na rok 2004 ve shodné výši jako v roce 2003 (350 Kč pro členy, 100 Kč pro studenty a důchodce);
4. činnost Rady a odborných skupin v roce 2003;
5. odměnu hospodáři ČsAS za rok 2003 ve výši 4000 Kč;
6. ukončení činnosti odborné skupiny H – Certifikace pracovníků v akustice.

Valná hromada ukládá nově zvolené Radě společnosti na kalendářní rok 2004:

1. pokračovat v odborné a organizační činnosti i v zahraničních kontaktech, rozvíjení spolupráce s Českou maticí technickou, Slovenskou akustickou společností, společností Elektra, českou sekci AES, Českým centrem IEE a československé sekce IEEE;
2. věnovat zvýšenou pozornost přípravě konference Inter-noise 2004 v Praze;
3. nadále rozvíjet vydávání Akustických listů.

Valná hromada ukládá nově zvoleným předsedům odborných skupin na kalendářní rok 2004:

1. publikovat informace o připravovaných aktivitách skupin v Akustických listech a na webové stránce s předstihem tak, aby se zájemci mohli včas na akce přihlašovat.

Valná hromada doporučuje Radě ČsAS:

1. pravidelně se zabývat činností a plánem akcí odborných skupin;
2. pravidelně se zabývat plánem a zaměřením konaných akustických konferencí.

Výsledky voleb do Rady České akustické společnosti

Počet vydaných platných hlasovacích lístků: 79

Odevzdáno platných hlasovacích lístků: 78

předseda:	V. KUNZL	78
místopředsedové:	J. KOZÁK	74
	J. ŠTĚPÁNEK	76
sekretář:	O. JIRÍČEK	77
hospodář:	O. KUDĚJOVÁ	78
revizní komise:	D. VÁPENÍKOVÁ	78
	T. HELLMUTH	76
	J. NOVÁK	76

Návrh usnesení sestavila návrhová komise ve složení M. Meller a P. Urban.

Vážení kolegové,

jak se můžete dočíst v usnesení valné hromady otištěném v tomto čísle Akustických listů, i v letošním roce zůstávají členské příspěvky nezměněny. Pro výdělečně činné členy tedy činí 350 Kč a pro důchodce a studenty 100 Kč. Pokud jste ještě nezaplatili, proveďte prosím platbu převodem na účet společnosti. Majitel účtu je Česká akustická společnost, č. ú. 17838061/0100 Komerční banka, a. s., Praha 6. Variabilní symbol uvádějte ve tvaru XXX03, kde XXX je Vaše osobní číslo, které naleznete nad svým jménem na obálce.

Zároveň bychom rádi aktualizovali databázi e-mailových adres. Proto prosím pošlete zprávu na e-mailovou adresu jiricek@fel.cvut.cz, která bude mít předmět (subject) Databaze a bude obsahovat aktuální telefonické či faxové spojení eventuálně další informace o Vás, které se změnily.

Ondřej Jiríček
sekretář společnosti

8. duben 2004 – Den „uvědomění si“ hluku

Hudebně ekologické sdružení při České hudební společnosti HUDEKOS vyhlašuje jako každý rok Den „uvědomění si“ hluku ve spolupráci s Americkou ligou neslyšících v New Yorku. K této výzvě se připojilo Německo, Polsko a Rakousko. Tento den se může kdokoli z veřejnosti obrátit na provozovatele hudebního hluku, odvolat se na naše sdružení a požádat o vypnutí produkce.

Informace naleznete na webových stránkách <http://www.musicecology.cz/>.

Sekretariát: ČHS, Radlická 99, 150 00 Praha 5
tel.: 251 552 382
fax: 251 552 453
e-mail: music@musicecology.cz

Jan Stěnička

Katedra radioelektroniky Fakulty elektrotechnické ČVUT
ve spolupráci s českou sekcí
Audio Engineering Society
si Vás dovoluje pozvat na přednášku

Time-Frequency Analysis for Acoustic Signals,

kteřou přednese dr. Laurent Simon z Université du Maine, Francie
(přednáška bude proslovena v angličtině).

Akce se koná dne 2. dubna 2004 v 11.00 hodin, v místnosti číslo 434 v prostorách ČVUT–FEL.

Dr. Ing. Libor Husník
předseda české sekce AES

prof. Ing. František Vejražka, CSc.
vedoucí katedry radioelektroniky ČVUT–FEL

České vysoké učení technické v Praze pořádá ve dnech **22. až 26. března 2004** odborný seminář **WORKSHOP 2004**, kde se formou posterů představí výsledky výzkumné činnosti v širokém spektru technických disciplín. Zájemci z praxe mohou navázat kontakty přímo s autory. Bližší informace na <http://workshop.cvut.cz>

Za organizační výbor
Libor Husník

Poznatky z porovnávacího měření v interferometru

Milan Krňák

Kafkova 20, 160 00 Praha

During the realisation of a Czech round-Robin test concerning the determination of the sound absorption coefficient in an impedance tube, additional measurements were made with small modifient test specimens. It was found that the change in the material has, greater influence than the mounting. It was also confirmed that the mineral wool has at the middle frequency range, a dip in the values of the sound absorbtion coefficient.

1. Úvod

Impulsem k uskutečnění dále uvedených měření bylo uspořádání mezilaboratorní porovnávací zkoušky akreditovanou zkušebnou akustiky č. 1007.5. Tento kruhový pokus [1] je, pokud je mně známo, první týkající se měření v akustickém interferometru metodou stojatých vln. Proč k porovnávání výsledků měření nedošlo už dříve, když druhá metoda měření činitele zvukové pohltivosti, a to v dozvukových místnostech, byla podrobena srovnávání výsledků snad od prvních pokusů pana Sabina a nejrozsáhleji v šestnácti laboratořích už v roce 1959? Důvod tkví zřejmě v jednoduchém principu měření, použitém původně v 19. století panem Kundtem ke stanovení rychlosti šíření zvuku. Proto také nenastaly technické spory z různých výsledků v laboratořích. K výpočtům činitele pohltivosti a impedance se totiž používají jen poměrné veličiny, když se kmitočet sinusového signálu a atmosférické podmínky stanoví s vyhovující přesností. Chybné výsledky jsou snadno poznatelné a většinou způsobené pronikáním signálu nebo hluku pozadí do mikrofonní sondy. Předností je dále nutnost jen malého vzorku, nevýhodou jen kolmý dopad zvukových vln a obtíže s návrhem vzorku kmitajících desek.

2. Metodika a měřicí zařízení

Metodika měření činitele zvukové pohltivosti α a složek akustické impedance z' a z'' pro kolmý dopad zvukových vln se řídí „ČSN ISO 10534-1 Akustika: Určování činitele zvukové pohltivosti a akustické impedance v impedančních trubcích – Část 1: Metoda poměru stojaté vlny“, která nahrazuje dřívější československou normu 73 0501 [2].

Používá se pro stanovení výše uvedených hodnot u materiálů a akustických obkladů pohlcujících zvuk. Získané hodnoty slouží především pro vývoj a kontrolu výrobků. Nelze je použít pro návrhy akustických úprav prostředky prostorové akustiky ke korekci doby dozvuku, anebo ke snížení hluku v uzavřených prostorech. Tam lze použít pouze činitel α_S získaný dozvukovou metodou.

Hlavní součástí měřicího zařízení je impedanční trubice, ve které se vytváří stojaté zvukové pole. V použitém zařízení je to svislá trubice čtvercového průřezu 75×75 mm a délky rovné části 930 mm. Je zhotovená z vodovzdorné překližky tlusté 10 mm, opatřené hladkým nátěrem a slepené v rozích a přilepené k přírubě držáku vzorku epoxy-

dovou pryskyřicí. Horní konec trubice pokračuje 190 mm dlouhou, o 45° skloněnou částí uvnitř s porézním klínem a svisle orietovaným otvorem pro mikrofonní sondu. Na konci šikmého úseku je příruba pro připevnění reproduktoru v uzavřené ozvučnici o objemu 3 dm^3 .

Mikrofonní sonda je trubička z lehkého kovu o vnějším průměru 8 mm, pružně napojená na tlakovou komůrku dynamického mikrofonu, umístěného ve zvukoizolačním krytu. Sonda je zavěšena na tenkém ocelovém lanku s protizávažím na druhém konci.

Celá konstrukce má montážní styky řešeny s ohledem na snížení nežádoucího přenosu budicího signálu nebo hluku pozadí do trubice nebo do sondy.

Vzorek tvoří souprava různých vysokých rámečků stejného průřezu jako trubice, což umožňuje vytvoření modelu akustického obkladu do největší tloušťky 105 mm. Celý vzorek včetně olověného, 9 mm tlustého dna, přitlačuje pohyblivá příruba pomocí čtyř ocelových svorníků k přírubě trubice.

Zdrojem signálu je přesný RC generátor Tesla BM 344. Mikrofonní zesilovač je kopií prvního zesilovače v analyzátoru B&K následovaný třetinooktáovým filtrem B&K, milivoltmetrem Tesla BM 384 s osciloskopem Křižík T 565.

Protože délka trubice neumožňuje v celém kmitočtovém rozsahu 200 až 2000 Hz stanovit akustický tlak ve dvou uzlech, používá se výpočet rychlosti šíření c_0 z nastaveného kmitočtu a teploty trubice podle rovnice (A.6) normy. Korekce na útlum v trubici nebývá uplatňována. Akustické vlastnosti prázdné trubice jsou podle tabulky 1 vyhovující, kromě činitele zvukové pohltivosti α a poměru stojatých vln s , u kmitočtu 2000 Hz.

3. Výsledky měření VIDETEXu, výrobce Vigona, Svitavy

3.1. Popis materiálu a vzorků

Tato plst' je vyrobena patrně z rozvlákněného odpadu hořlavých tkanin nebo nití pojeného do desek o tloušťce 12 mm. Na její líc a rub je bodově nalepena velmi řídká netkaná textilie o tloušťce asi 0,15 mm a plošné hmotnosti $0,14 \text{ kg} \cdot \text{m}^{-2}$. Tato textilie vytváří na měkkém povrchu jemné plsti v jednom směru souběžná, nepravidelná „údolíčka“ o hloubce až asi 1 mm. Akustické vlastnosti díky řídkosti a nízké plošné hmotnosti neovlivňuje.

f [Hz]	200	250	315	400	500	630	800	1000	1250	1600	2000
α [-]	0,011	0,013	0,015	0,017	0,019	0,01	0,013	0,006	0,006	0,017	0,060*
s [dB]	51	49	48	47	50	52	50	55	55	47	36*
L_0 [dB]	69	63	72	72	70	74	69	70	69	66	65

Tabulka 1: Akustické vlastnosti prázdné trubice – naměřené hodnoty činitele zvukové pohltivosti α pro kolmý dopad zvuku, poměru stojatých vln s a hluku pozadí L_0

Veličina	Hmotnost	Plošná hmotnost	Objemová hmotnost	Reálná složka ak. impedance pro $f = 2000$ Hz	Teplota vzduchu	Tlak vzduchu	Rychlost šíření
	[g]	[kg.m ⁻²]	[kg.m ⁻³]		[°C]	[hPa]	[m.s ⁻¹]
Vzorek B1	5,77	1,025	85,5	1,29	22,5	1012	345,5
B2	6,08	1,081	90,1	0,99	23,5	1012	345,5
B3	5,93	1,054	88,0	1,02	23,5	1012	345,5
Průměr	5,93	1,050	87,9				

Tabulka 2: Materiál VIDETEX – tabulka fyzikálních hodnot

f [Hz]	200	250	315	400	500	630	800	1000	1250	1600	2000
Vzorek: B1	0,038	0,034	0,043	0,044	0,068	0,096	0,12	0,16	0,27	0,36	0,42
B2	0,050	0,048	0,056	0,060	0,074	0,087	0,14	0,18	0,25	0,35	0,39
B3	0,041	0,036	0,045	0,058	0,071	0,094	0,13	0,18	0,25	0,35	0,37
Průměr	0,043	0,039	0,048	0,054	0,071	0,091	0,13	0,17	0,25	0,35	0,39
α_{pi} (oktávový)		0,043			0,072			0,18			0,37*
α_{str} (NRC)						0,150**					
α_w (vážený)						0,150**					
Poznámky:	* chybí hodnota pro 2500 Hz				** podle norem mají být určeny z hodnot α_S						

Tabulka 3: Materiál VIDETEX – tabulka naměřených hodnot činitele zvukové pohltivosti pro kolmý dopad zvuku

Zhotovení vzorků B1, B2, B3, rozměrů 75 × 75 mm z pásu jen 80 mm širokého bylo obtížnější kvůli vysoké pevnosti vláken. Plst' musela být z obou stran naříznuta a nakonec ustřižena nůžkami. Na jedné straně vzorku B1 zůstala u jedné hrany štěrbina asi 1 mm široká, která musela být zaplněna plastelinou. Kvůli nízké ohybové tuhosti plsti nebyly vzorky zcela rovinné. Vzorek byl zasunut do rámečku držáku vzorku o výšce 10 mm a vložky o tloušťce 1,3 mm, nepatrně a nepravidelně poněkud vyčníval do trubice.

3.2. Zhodnocení

Naměřené hodnoty činitele zvukové pohltivosti vykazují u všech vzorků plynulý nárůst hodnot od velmi nízkých hodnot kolem $\alpha = 0,05$ až k hodnotám středním $\alpha = 0,4$ na konci měřeného pásma. Mají malý rozptyl hodnot díky velké homogenitě plsti. Reálná složka akustické impedance se u vyšších kmitočtů blíží optimální hodnotě 1,6.

4. Výsledky měření ORSTECH 90 – minerální plsti vyrobené v Saint-Gobain Orsil, s. r. o., Častolovice

4.1. Popis materiálu a vzorků

Měřený materiál je minerální plst' pojená z minerálních vláken do desek rozměrů 500 × 1000 mm, různé tloušťky a objemové hmotnosti. Lícový i rubový povrch vykazuje pravidelné zvlnění v jednom směru v rozteči 6 mm a o hloubce až 1 mm, čímž přesahuje nominální tloušťku 40 mm. Předložený vzorek měl udánu nominální objemovou hmotnost 90 kg.m⁻³. Jednotlivé vzorky byly zhotoveny z 80 mm širokého pásu vyříznutého z desky napříč. Z náhodně vybraných míst byly tenkou jemnou pilkou podle ocelových přílohek vyříznuty vzorky A1, A2, A3 o průřezu 75 × 75 mm. Zvážením vzorků byly získány fyzikální hodnoty uvedené v tabulce, které ukazují na poměrně velkou nehomogenitu materiálu, kterou naznačují „toky“ vláken pozorovatelné na řezech vzorků. Se změnou hmotnosti zcela koresponduje změna reálné složky akustické impedance, což ukazuje na různé množství vláken a nikoliv nerozvlákněné kuličky minerální hmoty ve formě „hnízd“. Při měření α_S v dozvukové místnosti a při praktické apli-

Veličina:		Hmotnost	Plošná hmotnost	Objemová hmotnost	Reálná složka ak. impedance pro $f = 2000$ Hz	Teplota vzduchu	Tlak vzduchu	Rychlost šíření
		[g]	[kg.m ⁻²]	[kg.m ⁻³]		[°C]	[hPa]	[m.s ⁻¹]
Vzorek	A1	18,3	3,25	81,3	1,70	22,6	1018	345,6
	A2	20,0	3,56	88,9	1,83	22,6	1018	345,6
	A3	24,5	4,36	108,9	2,24	23,7	1018	346,2
Průměr		20,91	3,72	93,0				

Tabulka 4: Materiál ORSTECH 90 – tabulka fyzikálních hodnot

f [Hz]		200	250	315	400	500	630	800	1000	1250	1600	2000
Vzorek:	A1	0,21	0,29	0,34	0,46	0,55	0,68	0,83	0,86	0,89	0,90	0,88
	A2	0,24	0,31	0,39	0,49	0,55	0,40	0,84	0,84	0,86	0,86	0,85
	A3	0,26	0,34	0,41	0,48	0,54	0,50	0,72	0,77	0,75	0,80	0,79
Průměr		0,24	0,31	0,38	0,48	0,55	0,53	0,80	0,82	0,83	0,85	0,84
α_{pi} (oktávový)			0,31			0,52			0,84			0,84*
$\alpha_{stř}$ (NRC)							0,60**					
α_w (vážený)							0,55**					
Poznámky:		* chybí hodnota pro 2500 Hz					** podle norem mají být určeny z hodnot α_S					

Tabulka 5: Materiál ORSTECH 90 – tabulka naměřených hodnot činitele zvukové pohltivosti pro kolmý dopad zvuku

kaci se však tato nehomogenita neuplatní. Vzorky byly lehce, bez pozorovatelné deformace a štěrbin, postupně vtlačeny do 40 mm vysokého rámečku držáku vzorků. Aby vzorky nezasahovaly do trubice, byla mezi rámeček a čelo trubice dána ještě 1,3 mm tlustá vložka.

4.2. Zhodnocení

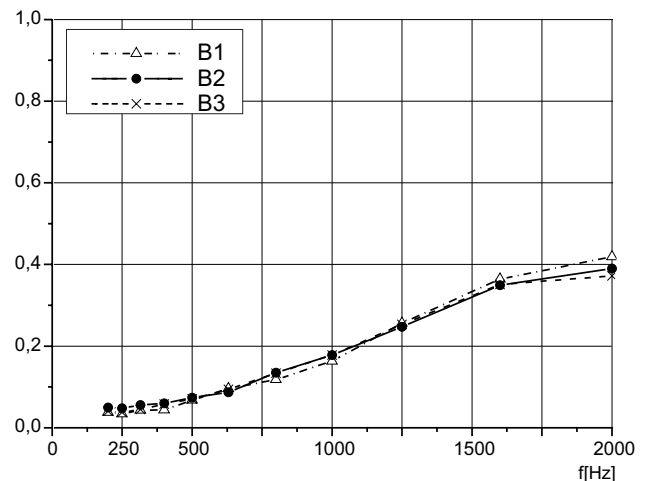
Naměřené hodnoty činitele zvukové pohltivosti vykazují u všech vzorků plynulý průběh kromě hodnot na kmitočtu 630 Hz, kde je průměrná hodnota α vzorku A1 a zejména A2 nižší oproti hodnotě u kmitočtu 500 Hz, ačkoliv by měla být mezi hodnotami pro 500 a 800 Hz. Podobné poklesy α se projeví už dříve u jiných vzorků minerální plsti a jsou podle jednoho zahraničního výrobce způsobeny kmitáním skeletu. Odchyly mezi vzorky jsou způsobeny spíše jejich nehomogenitou než odlišnou montáží nebo chybou metody měření.

5. Poznatky z hlavních a doplňovacích měření

Po naměření hodnot f , $U_{\max}$, $U_{\min}$, l_1 a zapsání byly ihned vypočítány výsledné hodnoty α_N , α_{stat} , z' , z'' , porovnány, a pak teprve zapsány i výsledné hodnoty k vyloučení hrubých chyb z omylů.

5.1. K měření VIDETEXu

Pro vizuální posouzení jsou na připojeném obrázku 1 vyneseny průběhy α tří vzorků B1, B2, B3. Významnější variační rozpětí 0,05 se objevuje jen u krajního kmitočtu 2000 Hz měřeného pásma. K ověření vlivu montáže byla

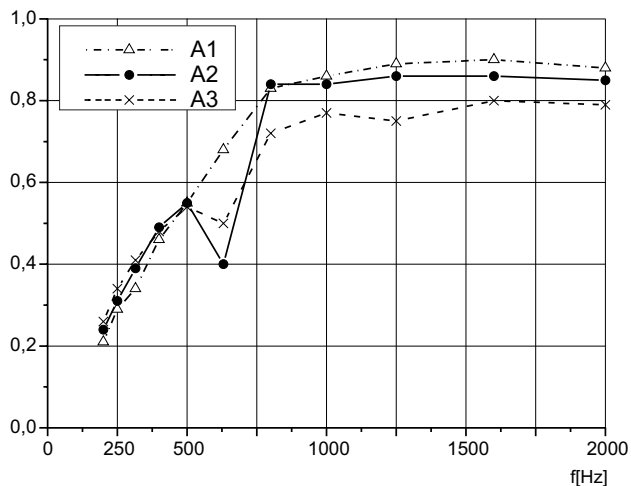
Obrázek 1: Kmitočtové charakteristiky činitele zvukové pohltivosti α pro VIDETEX

provedena se vzorkem B1 další dvě měření, a to po novém zasunutí vzorku do rámečku s respektováním omezení štěrbin na jedné straně plastelinou, a pak bez jakékoliv změny jen nové proměření. Můžeme tedy porovnat variační rozpětí tří různých vzorků a jednoho vzorku třikrát měřeného. V prvním případě nacházíme nejmenší rozpětí 0,009 a průměrné rozpětí ze všech kmitočtů 0,0154, ve druhém případě nejmenší rozpětí 0,001 a průměrné 0,0121, tedy asi o 21 % menší. Největší rozpětí je v obou případech u krajního kmitočtu 2000 Hz.

Vidíme, že i u homogenního materiálu mají drobné rozdíly v materiálu vzorků větší vliv než chyby montáže a měření.

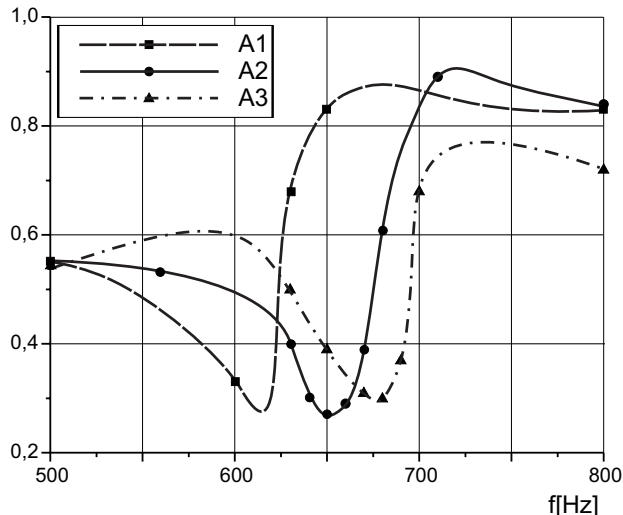
5.2. K měření ORSTECHu 90

Shodou náhod jsem vůbec jako první měřil vzorek A2, u kterého se projevil atypický pokles α při vzrůstu kmitočtu z 500 Hz na 800 Hz, jak je vidět na grafu na obrázku 2. Snažil jsem se tento dip dalšími měřeními až



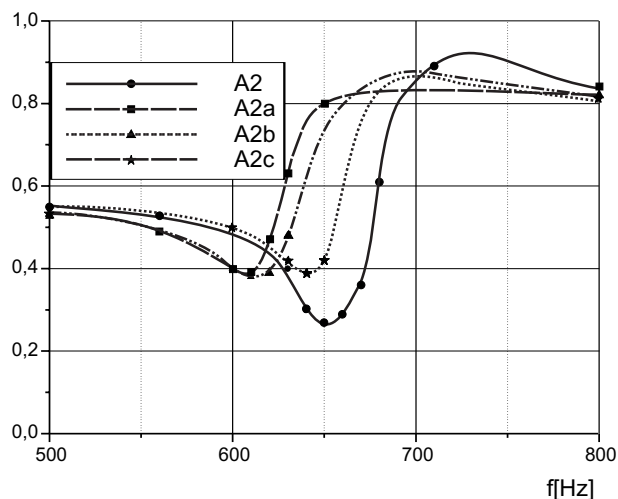
Obrázek 2: Kmitočtové charakteristiky činitele zvukové pohltivosti α pro ORSTECH 90

s odstupem jen 10 Hz zpřesnit a výsledek je vidět na podrobném grafu na obrázku 3. Dip se objevuje u všech



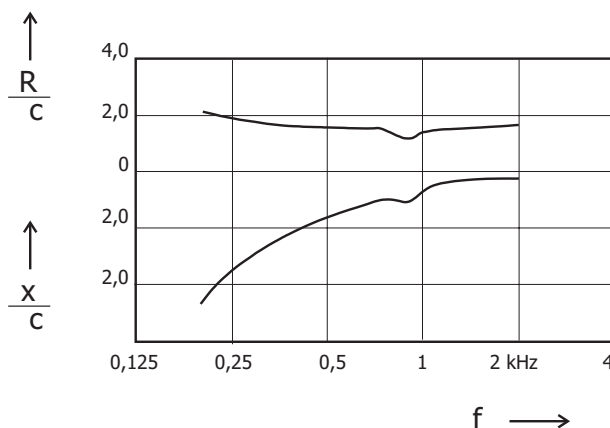
Obrázek 3: Podrobné charakteristiky činitele zvukové pohltivosti α pro tři různé vzorky, ORSTECH 90

vzorků, a to u vzorku A1 je $\alpha = 0,33$ u 600 Hz, u vzorku A2 je to 0,27 u 650 Hz a u A3 činí dip 0,30 u 680 Hz. Dip se tedy posouvá souhlasně se zvětšující se plošnou hmotností, a tedy se stoupajícím množstvím vláken ve vzorku. K získání dalších poznatků jsem udělal ještě tři další měření se vzorkem A2. Měření A2a bylo po opětném, tedy druhém zasunutí vzorku do rámečku držáku vzorků, měření A2b jen po povolení a utažení



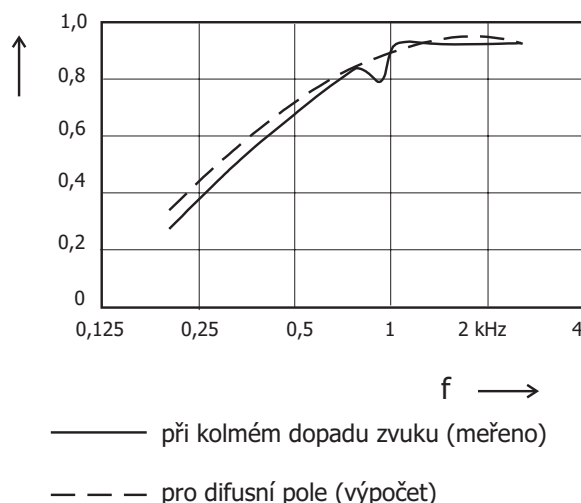
Obrázek 4: Podrobné charakteristiky činitele zvukové pohltivosti α pro vzorek A2, ORSTECH 90

svorníků držáku vzorků a měření A2c po obrácení rámečku se vzorkem, tedy po výměně lícové a rubové plochy (výsledek je vidět na podrobném grafu na obrázku 4). Všechna tři opakování mají společnou novou montáž vzorku a stejný pokles velikosti dipu α o 0,11. Na kmitočty dipu však montáž vzorku vliv má, pochopitelně nejmenší při prostém povolení a utažení svorníků. Pokud si vzpomínám, byl podobný, byť i méně výrazný dip, zjištěn při předepsaném kontrolním měření SILLANU SP100 (100 kg.m^{-3} a 50 mm tloušťky, tedy podobný materiál jako ORSTECH) od firmy Grünzweig a Hartmann, použitý při mezinárodním srovnávacím měření v dozvukových místnostech v letech 1959 a 1960. Výsledky jsou uvedeny na obrázku 5 a na obrázku 6. Výrobce tuto anomálii vysvětlil kmitáním skeletu vláknitého materiálu. Dále jsem našel měření Isoflexu 50 a 100 kg.m^{-3} celkové tloušťky 105 mm z roku 1999, které vykazuje dip o velikosti 0,09



Obrázek 5: Kmitočtové charakteristiky činitele zvukové pohltivosti Sillanu o tloušťce 50 mm pro kolmý dopad (převzato z [3])

na kmitočtu 800 Hz, když na kmitočtech 710 a 900 Hz jsou hodnoty spojitě vyšší.



Obrázek 6: Kmitočtová charakteristika činitele zvukové pohltivosti Sillanu o tloušťce 50 mm (převzato z [3])

Zda lze toto chování minerální plsti vysvětlit kmitáním skeletu, jak udával tehdy výrobce, si nejsem zcela jist. Protože mám z dřívější výroby Orsilu kolekci vzorků čtyř objemových hmotností a tří tloušťek, možná zkusím dalšími měřeními tuto věc osvětlit.

Možnost proložení naměřených bodových hodnot α plynulými křivkami svědčí o dobré způsobilosti měřicího zařízení.

Reference

- [1] Schwarz, J.: *Vyhodnocení mezilaboratorní porovnávací zkoušky činitele zvukové pohltivosti podle ČSN ISO 10534-1,2 a ČSN EN ISO 11654*, Centrum stavebního inženýrství – zkušebna akustiky, Akreditovaná zkušební laboratoř č. 1007.5, Praha, 2003.
- [2] Krňák, M.: *Akustika: Měření činitele zvukové pohltivosti a akustické impedance v interferometru*, ČSN 75 0501, ÚNM, Praha, 1989.
- [3] Kolmer, F., Krňák, M.: Beitrag zur Problematik der Diffusität des Schallfeldes im Hallraum, *Akustica Vol.10*, No.5/6, p. 36, S. Hirzel Verlag, Stuttgart, 1960.

Prosodic analysis and manipulation demonstrated on the Praat software

Tomáš Duběda and Radek Skarnitzl

Institute of Phonetics, Charles University in Prague

Príspevek se zamýšlí nad motivací a metodologií různých typů zkoumání a ovlivňování prozodie. V úvodu je představen volně dostupný program Praat, jenž velmi dobře uspokojuje nároky na komplexní fonetickou analýzu, a otázky prozodické analýzy z hlediska technického postupu a volby jednotek.

Mezi způsoby prozodické manipulace se řadí egalizace prozodických hodnot (v oblasti základní frekvence jde o monotonizaci). Další důležitou manipulací je delexikalizace, tedy zbavení signálu segmentálních charakteristik. V případě editace prozodických parametrů na materiálu přirozené řeči se postupuje cestou resyntézy. Lze tak např. přenášet (transplantovat) prozodické konfigurace z jedné věty na druhou. Konkrétní příklad transplantace základní frekvence v programu Praat je představen v závěru článku, včetně použitého skriptu.

1. Introduction

Under prosody we understand changes of those acoustic parameters which are usually examined at a higher than segmental level. Prosodic features include mainly intonation (fundamental frequency, or f_0), duration of speech segments and their intensity, in a broader sense also spectral appearance. To illustrate the contrast between the segmental and suprasegmental (or prosodic) structure, we may say, for instance, that the timbre of the vowel [i:] in the word *speaking*, as well as the contrast between the duration of this long vowel and that of short vowels, are segmental problems, while the intonation pattern of the word or the speaking rate of its production pertain to the suprasegmental area.

Some acoustic theories and procedures are based on the separation of the segmental and suprasegmental components, as for example the source-filter theory (Fant, 1960) or linear predictive coding (LPC; see e. g. Ladefoged, 1996). In concatenative speech synthesis, the segmental component is usually accounted for by choosing appropriate units (e. g., diphones) and is then, in most cases, not changed anymore. However, the units must remain “open” to suprasegmental changes which are applied on the given segmental base.

It is also useful to separate the two components in acoustic and perceptual analysis, or at least to restrict (neutralize, equalize) one of them. If the subject of our experiment is, for instance, the perception of voicing in the consonants [p] and [b], the samples used should be more or less balanced in terms of duration, intensity, and fundamental frequency, so as to avoid the distortion of the results (unless the correlation between the perception of voicing and these prosodic parameters is what we want to examine). On the other hand, when examining intonation contours it is often beneficial to use comparable segmental base, especially in terms of the alternation of voiced and voiceless segments.

A radical separation of the two components, such as overall monotonization of a sentence (equalization of fun-

damental frequency to a single value, while retaining the segmental structure) or – the other way round – its delexicalization (removing of the perceptible segmental phonological structure, while retaining prosodic characteristics), may on the one hand be informative, but on the other hand misleading, since examining such an isolated parameter is close to a mere artefact obtained from natural speech, and not to natural speech itself. However, this approach is often the only way to perform an analysis.

In the following paragraphs, we will consider the possibilities of prosodic analysis, editing, and modelling offered by the freeware Praat program. The outlined problems, which have been taken from phonetic practice, are accompanied by specific examples.

2. Praat

Praat, the whole name being “Praat – doing phonetics by computer”, is a program developed by the Dutch phoneticians Paul Boersma and David Weenink from the Institute of Phonetic Sciences, University of Amsterdam. Praat – the word “praat” means *I speak* (as well as *you speak* and *he/she speaks*) in Dutch – is a freeware program downloadable from the site <http://www.praat.org/>, and it works on all of the most important platforms (Windows, Linux, and Macintosh). The program was introduced in 1992 and it is continuously updated; a new version is available on the Internet approximately every one or two weeks. At the time of the preparation of this paper, the most recent version released was Praat 4.1.14 (October 29, 2003).

In this section, we are going to introduce the basic characteristics of Praat, and mention some phonetic analyses that we can perform with this program. Spectrographic analysis, which operates on the basis of Fourier transformation, offers the choice of various settings (e. g., window length and shape, or time step). The methods for the analysis of fundamental frequency include autocorrelation, cross-correlation, and spectral compression. When tracking f_0 , it is possible to set the criteria for specifying the

goodness of individual candidates. With other analytic components, we can examine formants, intensity or voice quality, we can filter and manipulate the speech signal. There are two possibilities to synthesize speech: Praat offers synthesis by means of formants, f_0 and intensity contours, or articulatory synthesis.

It is very easy in Praat to label sound files in multiple tiers – we may, for instance, label words in one tier, speech-sounds in the second tier, and subsegmental phenomena in the third (e.g., when examining plosives, it may be beneficial to represent separately the hold phase, the plosion, and the transition toward the following speech-sound). It is possible to use the phonetic symbols of the IPA for labelling.

It is possible to export data into a binary file, into the *.txt or *.xls formats. Similarly, graphical outputs may be exported into the *.eps and *.wmf formats, as well as directly into a graphics editor via clipboard.

An indisputable advantage of the Praat software is the possibility to create one's own scripts which facilitate and speed up research. The scripting language is relatively easy and intuitive (see Section 6 for an example), and it is possible to insert commands by the so-called history mechanism (i.e., to save a series of commands performed by means of buttons and, subsequently, to paste it in the script editor in the correct syntactic form). Moreover, a tutorial is available, where the user may verify the syntax of the given commands. When we do not find desired help in the tutorial, there is a so-called *Praat User List* (<http://uk.groups.yahoo.com/group/praat-users/>) where users share their experience, and the author of the program P. Boersma sometimes enters the discussions as well.

We believe that Praat is – thanks to the wide scope of possibilities, to its freeware character, user-friendly environment, and continuous development – a truly universal program for phonetic analysis.

3. Prosodic analysis

As we mentioned in the introduction, the three main components of prosody are intonation, intensity, and duration. The intonation contour is given by changes of fundamental frequency (f_0). Fundamental frequency is typically measured in Hertz (Hz). The unit Hz is a physical, i.e., objective unit, which does not reflect the way our hearing system perceives changes in pitch (for instance, we do not hear the difference between 100 and 200 Hz on the one hand, and between 1100 and 1200 Hz on the other as equal pitch intervals).

That is why various psychoacoustic units have been proposed which better reflect our perception (for a more detailed overview see, for example, Neppert, 1999). The unit *mel* was derived from experiments with frequency intervals; the subjects were asked to determine when they perceive the pitch of a probe tone to be double or half that of a given signal tone. When the interval between two sounds in mels is the same, the intervals are also perceived

as the same. The unit *Bark* was later derived from experiments which examined the auditory bandwidth. Based on experiments with masking, the concept of the so-called critical bandwidth was developed, whose bandwidth corresponds to 1 Bark. Both mels and Barks are nearly logarithmical functions. Barks have several disadvantages, though, the most important of which renders the unit inappropriate for the analysis of intonation: the resolution of Barks is too rough, especially in the low frequencies under 500 Hz (which is the frequency domain of intonation). That is why Barks are only used for formant analysis in Praat. The drawbacks of Barks have been overcome by the unit ERB (Equivalent Rectangular Bandwidth) which was also derived from experiments with the bandwidth of the auditory filter. Another unit, which stems from the Western musical tradition, is semitones. It is a purely logarithmical scale in which 12 semitones correspond to one octave, and a one-octave increase corresponds to a twofold increase in Hz. The semitone scale is sometimes also called the harmonic scale, as opposed to the melodic (or mel) scale. It turns out (Nolan, 2003) that semitones, and only slightly less ERBs, are closest to perception when analysing intonation. The units Bark and Hertz are from the perceptual viewpoint least appropriate for intonation analysis.

In the Praat program, the f_0 contour can be viewed in the editor window, together with the waveform and spectrogram of a given sound (*Pitch – Show pitch*). It is also possible in the editor window to select the desired units; apart from Hertz, it is possible to view f_0 in mels, ERBs, and semitones (menu *Pitch – Pitch settings...* in the editor window). In the same menu, we may extract the f_0 contour into an independent object, which enables us to perform various manipulations with it (*Pitch – Extract visible pitch contour*). We can achieve the same result directly from the object window (*Periodicity – To pitch...*).

Intensity corresponds to what we may call the magnitude of sound. Intensity is a physical quantity, defined as energy which arrives onto an area of 1 m^2 within 1 second, and is therefore given in W/m^2 . However, it is not convenient to operate in these units, which is why the logarithmical unit decibel (dB) was developed. The decibel was originally conceived as a relative unit. When we use it as an absolute unit, the generally acknowledged reference value of intensity is 10^{-12} W/m^2 , which corresponds to 0 dB. When we use decibels in this absolute sense, we talk about the so-called sound pressure level, and intensity values are then given in dB SPL.

Since the decibel – as a physical unit – does not correspond to the way we perceive the magnitude of sound, two principal subjective units were devised. The first, loudness level, whose unit is the phon (Ph), is only relevant for specifying the subjective magnitude of pure (sinusoidal) tones. A value of loudness level in phons specifies what the intensity in dB SPL of a 1000 Hz tone must be to be perceived as loud as a given signal tone. By applying this procedure for the whole frequency scale, we will arrive at

the so-called equal-loudness contours. Since loudness level does not express how we perceive complex tones, another quantity, called loudness, was introduced, and it concerns the overall percept of sound magnitude. Its unit is called the sone, and 1 sone corresponds to the loudness of a tone whose intensity is 40 dB SPL. It was found that we perceive an increase in intensity by 10 dB as a twofold increase in loudness. That means that with an exponential growth of intensity, the subjectively perceived loudness grows linearly.

In Praat, the intensity contour may be viewed in the editor window (*Intensity – Show intensity*). As with intonation, the intensity contour can be exported into an independent object from the editor window (*Intensity – Extract visible intensity contour*) or from the object window (*To Intensity...*). When we convert this new object into an intensity tier (*Down to IntensityTier*), we will be able to edit the dynamic contour. At the level of intensity, Praat operates only with decibels.

As far as the third main prosodic dimension, duration, is concerned, the situation here is relatively easy. Duration values in speech analysis are specified in milliseconds (ms). In the most common analyses – in the spectrogram, as well as in the waveform – duration corresponds to the horizontal axis.

4. Equalization of prosodic parameters

Some types of phonetic research require the prosodic values to be adapted in a way to reduce or even completely exclude prosodic variability in the given speech unit. This procedure may be referred to as *equalization* or *normalization*.

Monotonization is the application of this principle in the domain of fundamental frequency: the result of the processing should be a constant F0, corresponding to a monotonous prosody. This manipulation gives us the possibility to exclude e. g. the intonational component of stress – the prosodic residuum, corresponding solely to intensity and duration excursions, can then be subject to independent tests.

In the Praat software, the monotonization can be achieved by applying a newly created *PitchTier* object (*New – Tiers – Create PitchTier*) on the selected sound. The *PitchTier* object, whose length should be equal to that of the sound, contains two points having the same frequency value; the first one must lie before the beginning of the effective signal, the second one after its end. The sound itself has to be prepared for manipulation by the command *To Manipulation*. The *Manipulation* object is then combined with the *PitchTier* by selecting *Replace pitch tier*.

The category of prosodic equalization also includes temporal and dynamic normalization of speech products. However, in contrast to the fundamental frequency, these two processes cannot be understood as an application of a single constant on the speechsounds in question, since

each speechsound has its intrinsic duration and intensity. Acoustic studies claim e. g. that the duration of vowels increases with their openness (i, u < e, o < a), that unvoiced sounds tend to be longer than the voiced ones, and that fricatives mostly have longer duration than plosives (Palková, 1994). As far as intrinsic intensity is concerned, we can establish a rough descending scale as follows: open vowels > closed vowels > sonorants > nasals > fricatives > plosives. The intensity of a given segment may be represented by its maximum, or by the mean over the whole speechsound – the selected value then determines further manipulations. A survey of intrinsic intensity for English is given e. g. in Lehiste (1970).

To obtain rhythmically neutral and linear-sounding speech, in which all prosodic boundaries are blurred, we have to adapt the duration and intensity with respect to these intrinsic values. These can be obtained for each speaker by extracting a reasonably large, labelled corpus. In the Praat application, the values can be best applied by means of a script which should transform the labelled sound in accordance with the given table of values (further on prosodic modifications in section 6).

5. Delexicalization

When investigating the perceptual impact of prosody, we have several possibilities to reduce the segmental structure of the utterance, i. e. to make its wording unintelligible. This allows us to “denude” the prosodic structure; however, it is not true that all three parameters can exist independently of the segments: while the information about fundamental frequency can be fully preserved, the temporal and dynamic relations between sounds disappear proportionally to the spectral impoverishment of the signal.

The highest degree of delexicalization is the simple extraction of f_0 (*Periodicity – To Pitch...*), which can be played back either in the form of pulses (*Play – Play pulses*), or as a neutral-sounding pseudo-vowel (*Play – Hum*).

If no such radical delexicalization is required, one can use a low-pass filter. The selected threshold determines the amount of spectral information to be removed. A gradual lowering of this threshold will first neutralize the differences between fricatives; by contrast, the difference of voicing as well as the specific character of plosives will probably be resistant enough to this manipulation. In the case of vowels, the setting of the threshold frequency should be guided by the position of formants, especially by the perceptually most significant second formant (F2). In Czech, F2 mostly varies between 600–2800 Hz.

The third way of delexicalizing speech consists in replacing individual segments by others. This method preserves the naturalness of speech, but is not technically trivial. It can be described as a “transplantation” of prosody from real speech stretch upon a sequence of sound without meaning, either pronounced by a real speaker or syn-

thesized. Technical aspects of this procedure, which are described in more detail in section 6, are no less crucial than the choice of segmental support for non-sense speech. In a study by F. Ramus and J. Mehler (1999), three degrees of segmental substitutions are mentioned: the “saltanaj” method roughly preserves the syllabic structure and sound phonotactics (combinatory properties of speechsounds); the substitution rules are the following: plosives $\rightarrow t$, fricatives $\rightarrow s$, liquids $\rightarrow l$, nasal consonants $\rightarrow n$, glides $\rightarrow j$, vowels $\rightarrow a$. The “sasasa” method, preserving the contrast between consonants and vowels, and so, roughly, the syllabic structure, only works with the sounds s and a . The third degree of substitution, called “aaa”, consists in copying the prosody on an a vowel of appropriate duration; the results is comparable to the extraction of f_0 followed by its playback with the spectrum of a neutral vowel (see above in this section). An example of f_0 transfer from one word to another is showed in section 6. A print-out of the script which performs this prosodic transformation is quoted thereafter.

It has been said that fundamental frequency is, among the three prosodic parameters, the one which has the weakest link to segmental structure. Should the research object be primarily duration or intensity, we have to perform an operation where the f_0 transfer is accompanied by adequate durational and dynamic changes. In the case of the “saltanaj” method, we could do with a rule according to which every speechsound representative in the output signal would inherit the duration of its counterpart in the source signal; however, it is probable that the same value would be perceived differently according to whether it belongs to a t or a k , an l or an r etc. The intensity problem is even more complicated, since a direct transfer of values is excluded; one has to work with coefficients expressing the mean-related intensity of the speechsounds. These more complex manipulations mostly require *ad hoc* scripts.

6. Prosodic modifications

Since the first half of the 20th century, different methods of speech synthesis have been in use in phonetic research. Although synthetic speech exhibits many unnatural features, it is of high value for phoneticians because it offers the possibility of precise modelling of a wide range of prosodic configurations, which would not be obtainable from natural speech. So, to test e. g. the perception of stress, we are able to synthesize stimuli in which a given prosodic parameter will attain precisely defined values or patterns. The perception response will then determine what the role of the parameter in stress perception is.

Among pioneer studies in prosody perception, we should especially mention the study of synthesized sentences with different final intonation patterns (Hadding-Koch and Studdert-Kennedy, 1964). The listeners’ response to the stimuli, formulated in a binary way (do you hear a statement or a question?) showed a non-linear and categorial relationship between acoustic values and their perceptual correlates.

Prosodic modelling can be performed both on synthetic and natural speech. For the latter case, we can reserve the term *resynthesis*. Full synthesis is more practical, since any utterance can be synthesized in real time, while resynthesis mostly exhibits a higher degree of naturalness. The problem of monotonization by means of prosody transfer on non-sense speech, as described in the preceding section, can be thus generalized and termed *prosodic transplantation*. When choosing appropriate anchoring points (e. g. sound or syllable), we can e. g. transform the phrase *in their garden* into a question (*in their garden?*) using the prosody of the sentence *Will you help me?*

To illustrate the points made above, we provide a specific example of f_0 transplantation from a real word upon a naturally pronounced pseudo-word (see Section 5 – delexicalization by means of sound substitution) along with the Praat script which leads to the result presented.

7. Conclusion

The present paper examines various ways of prosody manipulation, and how they can be performed in the free Praat program. We argue that it is beneficial in certain situations in phonetic research to separate the segmental and suprasegmental components. We describe two extreme ways of this separation – monotonization of a sentence (equalization to a single f_0 value) and the opposite process, delexicalization (stripping a sentence off its segmental structure, while preserving the prosodic component only) – and the so-called prosodic transplantation (for example, taking the intonation of a question and applying it onto a statement). We argue that Praat is a useful tool not only for the analysis of suprasegmental area of speech, but for phonetic analysis in general.

Acknowledgements

This research was supported by the grant of the GACR, No. 102/02/0124. The authors of the article express their gratitude to Jan Votrúbec for his patient help in exploring the Praat program.

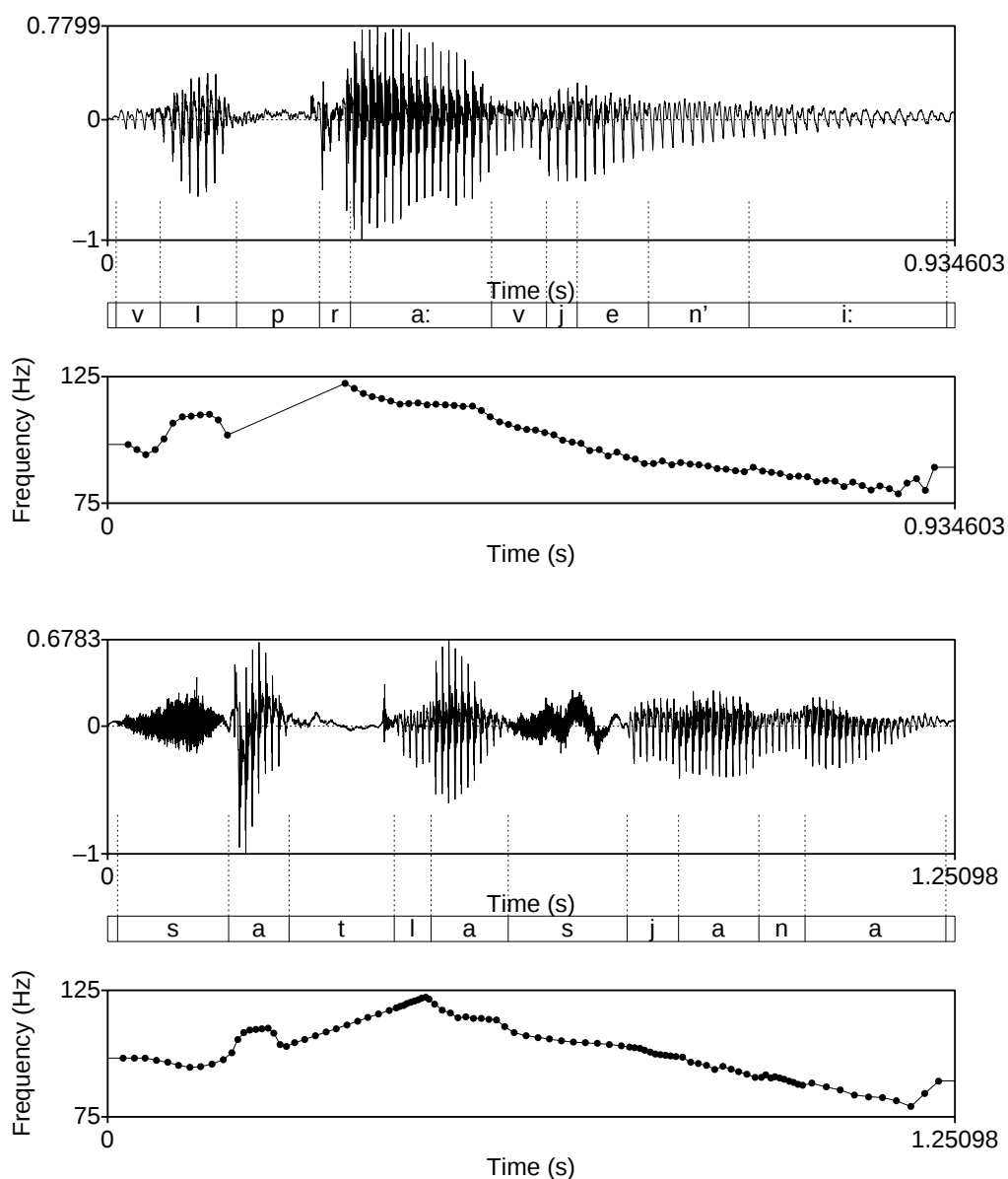


Figure 1: An example of f_0 transfer from the word *vyprávění* “tale, narration” (upper part of the figure) on the pseudo-word *satlasjana* (lower part of the figure). The sounds are of unequal length, and the time axis is normalized on the diagram (see time indications in the lower right corner of the frames). According to the boundaries set during labelling, each sound was divided into 10 equidistant points at which the f_0 values were measured. These values were then transferred to 10 equidistant points anchored in the target sound in the same way

```

snd_source = selected ("Sound", 1)           /source sound
grd_source = selected ("TextGrid", 1)        /source sound labelling
snd_target = selected ("Sound", 2)           /target sound
grd_target = selected ("TextGrid", 2)        /target sound labelling
select 'snd_source'
To Pitch (ac)... 0.01 75 15 no 0.03 0.45 0.01 0.35 0.14 600 /f0 computation
Down to PitchTier
ptch_source = selected ("PitchTier", 1)
select 'grd_source'
n_lab = Get number of intervals... 1         /establishing the number of speechsounds
select 'snd_target'
l_target = Get duration                      /establishing the duration of the target sound
Create PitchTier... ptch_target 0 'l_target' /creating an empty PitchTier (target PitchTier)
ptch_target = selected ("PitchTier", 1)
for i from 2 to n_lab-1                     /step-by-step speechsound modifications
    select 'grd_source'
    start1 = Get starting point... 1 i
    end1 = Get end point... 1 i
    dur1 = end1 - start1                     /establishing the duration of the ith
speechsound in the source sound
    select 'grd_target'
    start2 = Get starting point... 1 i
    end2 = Get end point... 1 i
    dur2 = end2 - start2                     /establishing the duration of the ith
speechsound in the target sound
    for j from 1 to 10                       /transfer of 10 f0 values for each speechsound
        time1 = start1 + (j*dur1/10)-dur1/20
        select 'ptch_source'
        value = Get value at time... 'time1' /getting the f0 in the source sound
        time2 = start2 + (j*dur2/10)-dur2/20
        select 'ptch_target'
        Add point... 'time2' 'value'         /adding a point to the target PitchTier
    endfor
endfor
select 'snd_target'
To Manipulation... 0.01 75 600              /preparing the target sound for f0 changes
manip = selected ("Manipulation", 1)
select 'manip'
plus 'ptch_target'
Replace pitch tier                           /f0 substitution

```

References

- [1] Boersma, P. & Weenink, D. (2003). Praat, version 4.1.11. <http://www.praat.org/>
- [2] Crystal, D. (1997). *The Cambridge Encyclopedia of language*. Cambridge: Cambridge University Press.
- [3] Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts: MIT Press.
- [4] Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton.
- [5] Hadding-Koch, K., Studdert-Kennedy, M. (1964). *An Experimental Study of some Intonation Contours*, *Phonetica*, 11, 175-185.
- [6] Ladefoged, P. (1996). *Elements Of Acoustic Phonetics*. Chicago: The University of Chicago Press.
- [7] Liberman, A. M., Delattre, P. C., Cooper, F. S. & Gerstman, L. J. (1976). The Role of Consonant Vowel Transitions in the Perception of the Stop and Nasal Consonants. In Fry, D. B. (ed.), *Acoustic Phonetics. A Course Of Basic Readings*. 315–331.
- [8] Neppert, J. M. H. (1999). *Elemente einer Akustischen Phonetik*. Hamburg: Helmut Buske Verlag.
- [9] Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales, *Proceedings of the XVth ICPHS*, Barcelona, 771–774.
- [10] Palková, Z. (1994). *Fonetika a fonologie češtiny*. Praha: Karolinum.
- [11] Ptáček, M. (1993). *Úvod do fonetické akustiky*. Praha: Karolinum.
- [12] Ramus, F. & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis, *Journal of the Acoustical Society of America*, 105, 512–521.
- [13] Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, Massachusetts: MIT Press.

Pokyny pro přípravu příspěvků do Akustických listů

Příspěvky do Akustických listů je možné napsat v jazyce českém, slovenském nebo anglickém. Příspěvky se dodávají v elektronické podobě na e-mailovou adresu *brothan@fel.cvut.cz* nebo *jiricek@fel.cvut.cz*.

Pro vypracování příspěvku je možné použít textový editor:

- **L^AT_EX** (do něj jsou příspěvky převáděny)
- **Word**
- **jiný** – zvlášť text a obrázky. Doporučujeme předem konzultovat s redakcí.

Na webové stránce

http://www.czakustika.cz/csas_cz.htm

jsou umístěny ukázky příspěvků pro L^AT_EX, resp. Word včetně použitého classu, resp. šablony. Jejich použití velice zjednoduší tvorbu vlastního příspěvku.

redakce

Akustické listy: ročník 10, číslo 1 březen 2004

ISSN: 1212-4702

Vydavatel: Česká akustická společnost, Technická 2, 166 27 Praha 6

Vytisklo: Ediční středisko ČVUT

Počet stran: 20

Počet výtisků: 200

Redakční rada: M. Brothánek, O. Jiříček, J. Kozák, R. Čmejla, F. Kadlec, J. Štěpánek, P. Urban

Jazyková úprava: R. Štěchová

Uzávěrka příštího čísla Akustických listů je 31. května 2004.

© ČsAS
NEPRODEJNÉ!