

Výroční zpráva projektu specifického výzkumu na rok 2017 - zakázka č. 2124

Název projektu: Závislost kvality řečového signálu pro vzdělávací účely na supersměrové charakteristice akustického přijímače a verifikace indexu srozumitelnosti analýzou signálu logatomických testů

Specifikace řešitelského týmu

Odpovědný řešitel:

doc. dr. René Drtina, Ph.D.

Studenti magisterského studia na PdF UHK:

Bc. Radka Vohralíková, P16P0860, MA-ZT

Externí spolupráce:

doc. Ing. Jaroslav Lokvenc, CSc.

Mgr. Pavel Fejfar, P16P0904, DR-INT

Akustické laboratoře FEL ZČU Plzeň:

Ing. Oldřich Tureček, Ph.D., Ing. Ladislav Zuzjak,

Ing. Jan Karel.

Celková částka přidělené dotace

240 779 Kč (vč. navýšení rozpočtu projektu) viz. RD 13/2017

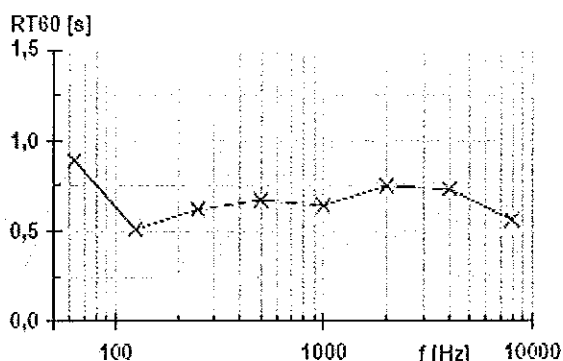
Východiska projektu

Digitalizace školství, sdílená videa a sociální sítě jsou oblasti do nichž směřují pokusy o tvorbu tzv. multimediálních studijních opor, jako jsou např. záznamy přednášek. Audiovizuální záznam se používá i jako pomůcka pro zpětnou vazbu a rozbor činnosti studentů při pedagogických praxích. Pro obě oblasti použití je nezbytné zajistit dokonalou srozumitelnost řeči a minimalizovat rušivé pozadí. Pro zvyšování kvality a efektivity vzdělávání, musí elektronické studijní materiály uvedený cíl podporovat. Veřejně dostupné materiály, stejně jako materiály používané při výuce, mají mít jistou profesionální úroveň a především vysokou vypovídací hodnotu, která přinese požadovaný didaktický efekt.

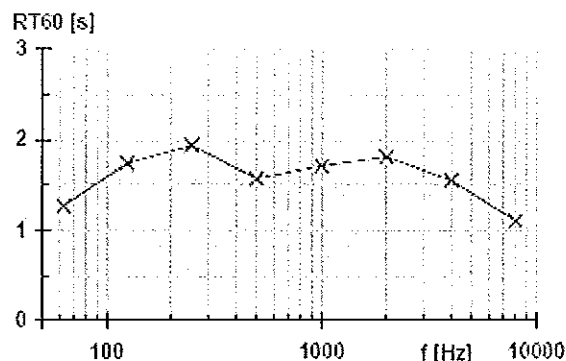
Postup při řešení projektu

Výzkum navazoval na projekty specifického výzkumu z let 2011-2017 a zabýval se měřením srozumitelnosti pomocí logatomických testů v závislosti na snímacích podmínkách [1] a jejich porovnání s výsledky zjištěnými normovanou metodou STIPA [2]. Použitá metodika vychází z normativních a souvisejících požadavků opakovatelnosti měření ve smyslu ČSN ISO 5725-1 [3], ČSN ISO 5725-2 [4], ČSN ISO 21748 [5], ČSN ISO 2602 [6], ČSN ISO 2854 [7], i v případě logatomických testů, které jsou primárně vyhodnocovány poslechem. Testy logatomické poznatelnosti představují nejpřísnější kritérium kvality přenosu řečového signálu s největší rozlišovací schopností, jsou mimořádně citlivé na frekvenční omezení v přenosovém kanálu, nelineární zkreslení, dobu dozvuku a odstup užitečného signálu od rušivého pozadí. Při dodržení metodiky a správném provedení dávají testy logatomické poznatelnosti zcela objektivní výsledky.

Pro všechna měření byl použit referenční zářič NTI TalkBox a 100% opakovatelnost zajistily logatomické testy [8] vyrobené ve spolupráci se studiem Českého rozhlasu Hradec Králové. Studiová kvalita zajišťuje 100% logatomickou poznatelnost a index přenosu řeči, měřený přes celý studiový řetěz, na maximální možné úrovni $STI = 1$, tj. srozumitelnost řeči $CIS = 100\%$ [1]. Zaručena je i technická kvalita A+ podle ČSN EN 60268-16 [2] a 100% kvalitativně-technická opakovatelnost podle požadavků ČSN ISO 5725-1 [3] a ČSN ISO 5725-2 [4]. Logatomické testy byly snímány v akusticky rozdílných prostředích. V učebně S51 s objemem $144,6\text{ m}^3$ a střední dobou dozvuku 670 ms (obr.1) a v učebně C4 s objemem $273,6\text{ m}^3$ a střední dobou dozvuku 1,72 s (obr.2). Hladina ambientního hluku se v době měření pohybovala na úrovni $39 \pm 3\text{ dB(Z)}$.



Obr.1 Doba dozvuku S51



Obr.2 Doba dozvuku C4

Měřicí signály STIPA, FFT i signály logatomických testů byly zaznamenávány souběžně z pěti snímačů ve standardním formátu DAT 16 bit/48 kHz. Tím byly zaručeny identické snímací podmínky. Spektrální analýzy byly provedeny

rychlou Fourierovou transformací, logatomické testy byly zpracovány poslechem přes uzavřená studiová sluchátka a koaxiální poslechové monitory. Měřicí technika byla doplněna o vlnový snímač a teleskopický výložník s kompletním příslušenstvím pro interiérová i plenérová měření, osmikanálový mobilní záznam, propojovací kabely a nezbytné příslušenství.

Dosažené výsledky

Při známých parametrech prostoru, akustického zářiče a akustického snímače, můžeme predikovat ztrátu srozumitelnosti souhlásek ζ , která přímo souvisí se ztrátou srozumitelnosti řeči

$$\zeta = \frac{200T^2 l_{tr}^2}{lbhQ_t Q_r} [\%] \quad (1)$$

kde T je doba dozvuku v daném prostoru [s], l_{tr} je vzdálenost vysílač - přijímač [m], lbh délka, šířka a výška prostoru [m], Q_t činitel směrovosti vysílače a Q_r činitel směrovosti přijímače. Do rovnice je možné zavést řadu korekčních faktorů v závislosti na šířce pásma, odstupu rušivého pozadí, atd. Logatomickou poznatelnost lze potom odhadnout na hodnotu

$$LOG \cong 100 - \frac{400T^2 l_{tr}^2}{lbhQ_t Q_r} [\%] \quad (2)$$

V tabulce 1 jsou hodnoty indexu přenosu řeči STIPA a srozumitelnosti CIS změřené analyzátozem NTI XL2 podle ČSN EN 60268-16. Tabulka 2 uvádí průměrné výsledky logatomických testů a predikci srozumitelnosti podle rovnice 2 a konverzní funkce uvedené v [9]. Výsledky v tab.2 reprezentují vyhodnocení tzv. surového záznamu bez jakýchkoliv dalších úprav pomocí filtrů, korekcí, dynamických procesorů, atd. Ve výsledcích nejsou korigovány ani fonémové chyby, které vznikají podobnou znělostí některých hlásek a slabik (například d-t, ž-š, y-i, mně-mě...).

Tab.1 Výsledky měření podle ČSN EN 60268-16

Učebna l_{tr}	Snímač	Srozumitelnost podle ČSN EN 60268-16		
		STIPA	Class	CIS [%]
S51 1,5 m	MKH70	0,93	A+	97
	NTI 2210	0,77	A+	88
	ME36	0,84	A+	93
	2MD441	0,84	A+	93
S51 6,5 m	MKH70	0,75	A	88
	NTI 2210	0,68	C	83
	ME36	0,74	A	87
	2MD441	0,75	A	88
C4 1,5 m	MKH70	0,85	A+	93
	NTI 2210	0,67	C	83
	ME36	0,79	A+	90
	2MD441	0,83	A+	92
C4 7,25 m	MKH70	0,61	D	79
	NTI 2210	0,51	G	71
	ME36	0,53	F	72
	2MD441	0,58	E	77

l_{tr} je vzdálenost mezi referenčním zdrojem a snímačem

Tab.2 Logatomická poznatelnost a predikce

Učebna l_{tr}	Snímač	Srozumitelnost [%]				
		Hlasový signál			predikce podle (2)	predikce podle [14]
		J	L	K		
S51 1,5 m	MKH70	97	97	96	100	99
	NTI 2210	97	97	95	99	96
	ME36	97	96	95	100	98
	2MD441	96	95	95	100	98
S51 6,5 m	MKH70	96	96	94	99	92
	NTI 2210	92	91	90	83	92
	ME36	96	93	93	95	95
	2MD441	95	96	97	97	95
C4 1,5 m	MKH70	96	96	96	100	98
	NTI 2210	96	95	96	97	91
	ME36	96	95	92	99	96
	2MD441	96	96	97	100	97
C4 7,25 m	MKH70	95	97	96	93	86
	NTI 2210	82	82	80	37	73
	ME36	91	87	91	71	76
	2MD441	95	95	96	88	83

(J) vyšší mužský hlas, (L) ženský hlas, (K) nižší mužský hlas

Pro výpočty podle (2) byly použity hodnoty uvedené v tab.3, činitel směrovosti NTI TalkBox $Q_t = 2$.

Tab.3 Činitelé směrovosti a dozvukové vzdálenosti ve zkušebních prostorech

Snímač	Činitel směrovosti Q_t	Dozvuková vzdálenost [m]	
		S51	C4
MKH70	18	4,99	4,22
NTI 2210	1,78	1,58	1,34
ME36	4	2,36	2,00
2MD441	10	3,74	3,17
posluchač	1,7-2,5	1,40-1,87	1,19-1,58

Z výsledků vyhodnocených logatomických testů (tab.2) můžeme pro záznamové podmínky ve dvou zkušebních prostorech odvodit, že v případě krátkých snímacích vzdáleností a při maximální dosažitelné technické kvalitě reprodukce i záznamu jsou rozdíly v závislosti na směrových charakteristikách snímačů zanedbatelné. Podobně je při dokonalé výslovnosti řečníka zanedbatelný i vliv spektrálního složení řeči. Údaje v tabulce 3 ukazují, že záznamy pořízené vysoce směrovými mikrofony (MKH70 a dvojice MD441) v první řadě učeben ve vzdálenosti 1,5 m od zdroje zvuku bezpečně spadají do tzv. volného pole. Naproti tomu všesměrový mikrofon a posluchač jsou při této vzdálenosti již na hranici volného a difúzního pole. V obou prostorech jsou poslední řady (6,5 m a 7,25 m) výhradně v difúzním poli. Stejně tak i vysoce směrové mikrofony jsou za hranici dosahu volného pole a v pořízeném záznamu se výrazně uplatňuje dozvuk prostoru.

Výsledky měření v konkrétních zkušebních prostorech ukázaly problematiku predikce reálně dosažitelné logatomické poznatelnosti, a to jak její stanovení výpočtem z parametrů prostoru, tak přepočtem ze změřeného indexu přenosu řeči. Přiměřená shoda prediktivní a reálné logatomické poznatelnosti je při frekvenčně neomezeném přenosu, krátké době dozvuku nebo snímání ve volném poli pouze v rozsahu 95-100 %. Musíme rovněž respektovat vysokou adaptabilitu sluchu na různá akustická prostředí.

Mimořádně příznivé výsledky poslechových testů jsou dány použitím špičkové profesionální snímací a záznamové techniky, nekomprimovaným záznamem a poslechem přes uzavřená studiová sluchátka nebo špičkové profesionální koaxiální monitory. Cílem dalšího rozsáhlého výzkumu je stanovení konverzních křivek mezi indexem přenosu řeči a logatomickou poznatelností při uvažování omezených frekvenčních charakteristik přenosového kanálu a dynamické komprese.

Výstupy projektu

V rámci řešení projektu jsou připraveny k publikování následující články:

VOHRALÍKOVÁ, R. - FEJFAR, P. - DRTINA, R. - LOKVENC, J. Dependence of the quality of speech signal on recording conditions - Using of high-directivity microphones. Conference IEEE2018 - Signal processing.

VOHRALÍKOVÁ, R. - FEJFAR, P. - DRTINA, R. - LOKVENC, J. *Závislost kvality řečového signálu elektronicky distribuovaných studijních materiálů na snímacích podmínkách - Část 1: Použití vysoce směrových akustických snímačů*. Media4u Magazine, 1/2018.

Přehled realizovaných výdajů

a) osobní náklady nebyly pro administrativní náročnost a nutné odvody nárokovány.

b) po schválení změn a navýšení rozpočtu byla navýšena částka na stipendia. Do stipendií byly přesunuty také plánované osobní náklady spolupracovníků a rozšířené cestovní náklady. Stipendium v celkové výši 28 037 Kč bylo přiznáno studentce P16P0860 - Bc. Radka Vohralíková za spolupráci při řešení úkolu: zajištění podpory alternativní technické specifikace po nutných typových změnách měřicí techniky, zajištění přepravy řešitelů na technickou přejímkou a adjustaci měřicí techniky (2× Janz Audio Praha), kontrola kompletnosti dodávky, vybalení a uložení měřicí techniky v nových prostorách KTP, příprava na provozní zkoušky před kalibrací a měřením, zajištění drobných nákupů doplňkových dílů, formátování záznamových médií, funkční zkoušky, dílčí příprava měřicích signálů pro logatomické testy, přeprava řešitelů (4× akustické laboratoře FEL ZČU v Plzni), kalibrace měřicí techniky, měření frekvenčních a směrových charakteristik, signálové analýzy odezvy prostoru, výpočet přenosu, instalace snímačů, záznam zkušebních signálů a logatomických testů, transport techniky, přepis záznamů logatomických testů, sumární vyhodnocení logatomických testů, stanovení dosažené srozumitelnosti, grafické zpracování směrových a frekvenčních charakteristik, příprava dat, zpracování úvodní části publikačního výstupu, spolupráce na textu článků a terminologická příprava překladu do angličtiny.

c) materiálové náklady dosáhly výše 176 352,95 Kč.

d) náklady na služby (akustické laboratoře, překlad článku) byly ve výši 36 390,00 Kč.

Přidělená dotace ve výši 240 779 Kč byla po schválených změnách v rozpočtu projektu zcela vyčerpána, rozpočet byl překročen o 95 haléřů. Podrobný přehled čerpání je uveden v tabulce 4.

Použité zdroje

[1] ČSN EN ISO 9921. *Ergonomie - Hodnocení řečové komunikace*. Praha. ČNI. 2004.

[2] ČSN EN 60268-16 ed.2. *Elektroakustická zařízení - Část 16: Objektivní hodnocení srozumitelnosti řeči indexem přenosu řeči*. Praha. ČNI. 2004.

[3] ČSN ISO 5725-1. *Přesnost (správnost a shodnost) metod a výsledků měření - Část 1: Obecné zásady a definice*. Praha. ÚNMZ.1997

[4] ČSN ISO 5725-2. *Přesnost (správnost a shodnost) metod a výsledků měření - Část 2: Základní metoda pro stanovení opakovatelnosti a reprodukovatelnosti normalizované metody měření*. Praha. ÚNMZ. 1997.

[5] ČSN ISO 21748. *Návod pro použití odhadů opakovatelnosti, reprodukovatelnosti a pravdivosti při odhadování nejistoty měření*. Praha. ÚNMZ. 2012.

[6] ČSN ISO 2602. *Statistická interpretace výsledků zkoušek. Odhad průměru. Konfidenční interval*. Praha. ČNI. 1993.

[7] ČSN ISO 2854. *Statistická interpretace údajů. Odhady a testy středních hodnot a rozptylů*. Praha. ČNI. 1994.

- [8] DRTINA, R. - SEDIVY, J. - SCHLOSSER M. *The Objectivization of the Entry Conditions of the Pedagogical Research*. IEEE Education and modern educational technologies (EMET 2015). IEEE Press. 2015. pp.25-28.
- [9] DRTINA, R. et al. *STIPA method in public adress sound systems and voice alarm systems - Part 2: Conversion curves and bandwidth effect*. Trans Tech Publications. Advanced Materials Research Vol. 1082 (2015). pp.574-580.

Tab.4 Přehled nákladů projektu 2124

akustický snímač long-wave MKH70	38 513,09 Kč
instalační upevnění MZQ40	912,34 Kč
antivibrační upevnění MZS20	8 692,64 Kč
plenérová ochrana MZW70	7 638,73 Kč
protihluková ochrana MZH70	5 795,90 Kč
interiérová ochrana MZW71	3 686,87 Kč
8kanálové registrační zařízení ZF8	27 567,43 Kč
transportní obal PCF8	3 014,11 Kč
ovladač ZFRC8	9 788,90 Kč
spojovací kabel XLR/3F-XLR/3M 5 m - 2 ks	1 447,16 Kč
zakázkový kabel GB2R-direct out/ZF8	17 280,01 Kč
maticový multiplex GB2R-12/2	25 981,12 Kč
výložník WDB XL-QT	19 438,65 Kč
speakon	3 480,00 Kč
SD-XC	1 800,00 Kč
zdroj pro FRC8	1 316,00 Kč
překlad článku	6 390,00 Kč
akustické laboratoře FEL ZČU	30 000,00 Kč
stipendium vč. cestovních nákladů	28 037,00 Kč
celkem	240 779,95 Kč

Datum: 29. prosince 2017

Podpis odpovědného řešitele



Radka Vohralíková - Pavel Fejfar - René Drtina - Jaroslav Lokvenec

Katedra technických předmětů, Pedagogická fakulta, Univerzita Hradec Králové
Department of Technical subjects, Faculty of Education, University of Hradec Králové

Abstrakt: Kvalitní pracovní prostředí vzdělávacích institucí, škol, ale i domácích pracovišť (při využívání e-learningových kurzů) je podmíněno řadou hygienických i ergonomických požadavků. Přestože se stupňují požadavky na multimediální podporu výuky, na zvyšování kvality a efektivity vzdělávání, je tvorba pracovního prostředí v kontextu požadavků ISO 9001 v podstatě minimální. Článek přináší výsledky vlivu snímacích podmínek a použitých snímačů na kvalitu zvukového záznamu a logatomickou poznatelnost, což je klíčový faktor pro srozumitelnost řečového signálu.

Abstract: A good level of the quality of the working environment of educating institutions, schools and also of home environment (in case of e-learning courses) is specified by a number of sanitary and ergonomic requirements. Although requirements on multimedia support of teaching, and on increase in the quality and effectiveness of education are escalating, there are more or less zero requirements imposed on the creation of the working environment in the context of the ISO 9001 requirements. The paper presents the results of the analysis of the impact of the recording conditions and the used sensors on the quality of audio recordings and logatonic recognisability, the key factor for the speech intelligibility.

Klíčová slova: zvuk; snímač; záznam; logatomická poznatelnost; srozumitelnost řeči; STIPA; CIS.

Keywords: sound; sensor; recording; logatonic recognizability; speech intelligibility; STIPA; CIS.

ÚVOD

Na veletrhu spotřební a výpočetní techniky International Consumer Electronics Show v roce 2015 prohlásil Werner Struth, šéf americké pobočky společnosti Bosch, že „analogový svět tak, jak ho známe, končí“ [1]. Přesto naše vnímání zůstává i nadále analogové. Obrazové a zvukové receptory, zrak a sluch jsou stále dominantními smysly při zprostředkování poznání okolního světa. Stejně tak přetrvává transfer informací k našim smyslům analogovým přenosovým kanálem, do něhož mohou zasahovat i omezení vyplývající z používaných digitálních technologií [2].

Snahy o zvyšování kvality a efektivity vzdělávání, rozvoj distančního vzdělávání a e-learningu zvyšují požadavky na elektronickou distribuci nejrozličnějších materiálů, včetně materiálů multimediálních. Přestože se požadavky na multimediální podporu výuky stupňují, požadavky na od-

povídající tvorbu pracovního prostředí v kontextu ISO 9001 [3] jsou velmi často opomíjeny.

Elektrotechnické laboratoře katedry technických předmětů Univerzity Hradec Králové se v dlouhodobé spolupráci s akustickými laboratořemi Fakulty elektrotechnické Západočeské univerzity v Plzni věnují problematice vlivu parametrů přenosového kanálu (frekvenční a směrové charakteristiky akustických snímačů a akustických zářičů, přenosová funkce prostoru) na dosažitelnou srozumitelnost řeči a objektivizaci jednočíselného parametru pro měření srozumitelnosti, který by odpovídal výsledkům logatomických testů.

Článek přináší výsledky měření srozumitelnosti pomocí logatomických testů v závislosti na snímacích podmínkách [4] a jejich porovnání s výsledky zjištěnými normovanou metodou STIPA [5].

1 METODIKA MĚŘENÍ A TESTŮ

Z hlediska prováděných výzkumů musíme být schopni zajistit potřebnou opakovatelnost měření ve smyslu norem ČSN ISO 5725-1 [6], ČSN ISO 5725-2 [7], ČSN ISO 21748 [8], ČSN ISO 2602 [9], ČSN ISO 2854 [10] a norem souvisejících, a to i v případě logatomických testů, které jsou primárně vyhodnocovány poslechem. V tabulce 1 je základní stoslábikový logatomický test pro český jazyk, z něhož lze permutací bez opakování teoreticky odvodit až $9,33 \cdot 10^{157}$ variant testů podle rovnice

$$P_{(100)} = 100! \quad (1)$$

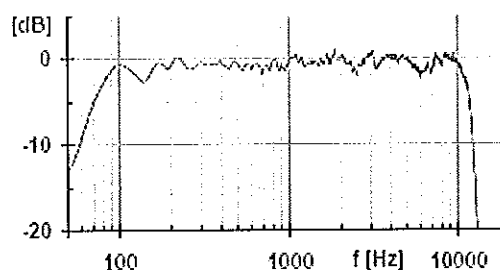
Reálně je v praxi využitelná kombinace 120 variant stoslábikových logatomických testů pro jeden hlas. Pro smíšené testy umožňuje kombinace dvacetislábikových skupin ze všech tří hlasů teoreticky vytvořit přes 360 tisíc testů.

Tab.1 Stoslábikový test pro český jazyk

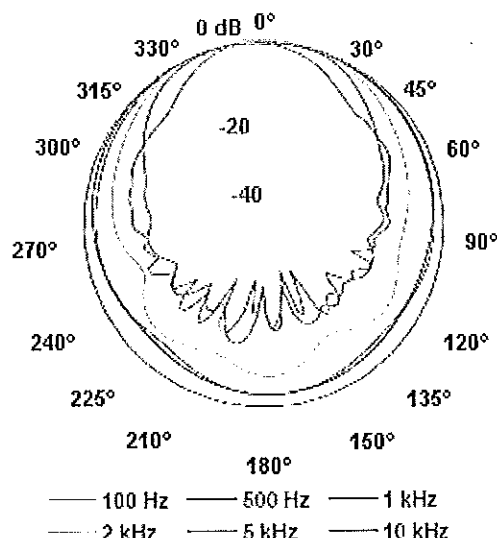
1 strach	21 to	41 te	61 nu	81 vln
2 po	22 tvař	42 dým	62 sí	82 teď
3 lé	23 ho	43 si	63 du	83 ví
4 dí	24 pad	44 tec	64 le	84 ně
5 pla	25 lou	45 něž	65 sko	85 žnou
6 ste	26 sbí	46 li	66 stří	86 smrt
7 a	27 což	47 je	67 tom	87 a
8 mo	28 se	48 krk	68 u	88 dě
9 zdě	29 ry	49 ta	69 ja	89 te
10 já	30 ru	50 sné	70 šu	90 ho
11 vů	31 o	51 vždy	71 jiš	91 za
12 ca	32 má	52 ne	72 to	92 lo
13 ve	33 de	53 vás	73 kdo	93 ma
14 ko	34 kra	54 či	74 šli	94 bem
15 str	35 dy	55 ši	75 zba	95 la
16 žeš	36 štěs	56 bych	76 vu	96 od
17 ner	37 kvů	57 tí	77 be	97 náct
18 ní	38 ně	58 snad	78 ce	98 pl
19 ba	39 pa	59 tak	79 na	99 nel
20 ce	40 ur	60 žá	80 ak	100 pou

Jako referenční akustický zříč byl pro všechna měření použit NTi TalkBox s referenční hladinou akustického tlaku 70 dB ve vzdálenosti 1 m. Frekvenční a směrové charakteristiky TalkBoxu jsou na obrázcích 1 a 2.

Pro 100% opakovatelnost logatomických testů [11] za stejných vstupních podmínek jsme v roce 2015 vyrobili ve spolupráci se studiem Českého rozhlasu Hradec Králové testy pro český jazyk ve třech hlasových podobách: nižší mužský hlas (baryton, K), vyšší mužský hlas (tenor, J) a ženský hlas (mezzosoprán, L).



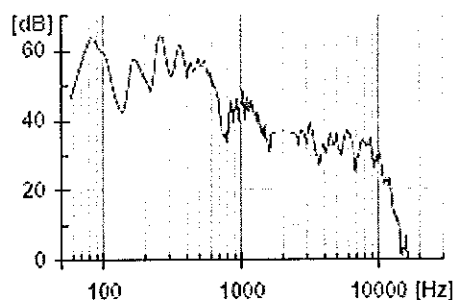
Obr.1 Frekvenční charakteristika NTi TalkBox



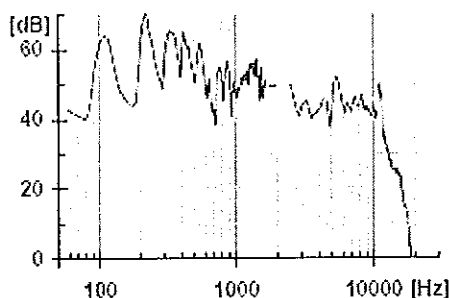
Obr.2 Směrová charakteristika NTi TalkBox

Studiová kvalita zajišťuje 100% logatomickou poznatelnost a index přenosu řeči měřený přes celý studiový řetěz na maximální možné úrovni $STI = 1$, tj. srozumitelnost řeči $CIS = 100\%$ [12]. Spektrální charakteristiky logatomických testů pro jednotlivé hlasy měřené pro celý logatomický test jsou na obrázcích 3, 4 a 5.

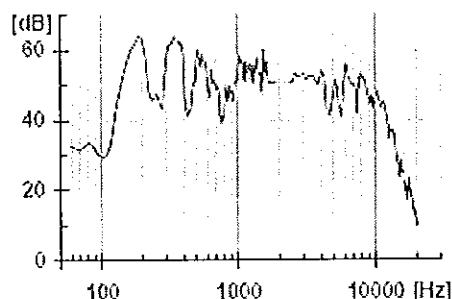
Měřicí signál STIPA je v současné době validován pouze pro mužskou řeč. Z porovnání spektrogramů (obr.6) je zřejmé, že měřicímu signálu STIPA téměř přesně odpovídá spektrum nižšího mužského hlasu (K).



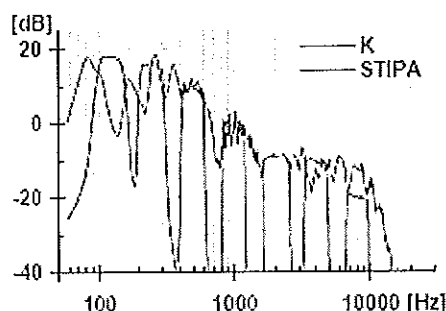
Obr.3 Spektrum nižšího mužského hlasu (K)



Obr.4 Spektrum vyššího mužského hlasu (J)



Obr.5 Spektrum ženského hlasu (L)



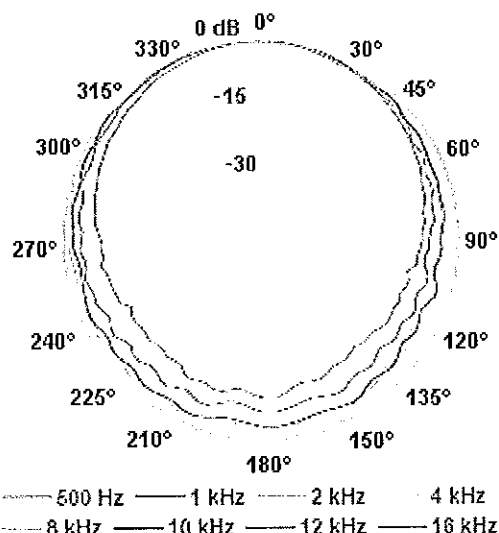
Obr.6 Porovnání spektra měřicího signálu STIPA a spektra nižšího mužského hlasu (K)

Ke snímání měřicích signálů bylo použito pět akustických snímačů s různými směrovými charakteristikami. Referenčním snímačem byl všesměrový měřicí mikrofon třídy I. od firmy NTi audio, typ NTi M2210 (obr.7).



Obr.7 Měřicí mikrofon NTi 2210

Směrové charakteristiky mikrofonu NTi 2210, změřené v bezodrazové komoře akustických laboratoří jsou FEL ZČU v Plzni, kde byly měřeny všechny směrové a frekvenční charakteristiky, jsou na obr.8.

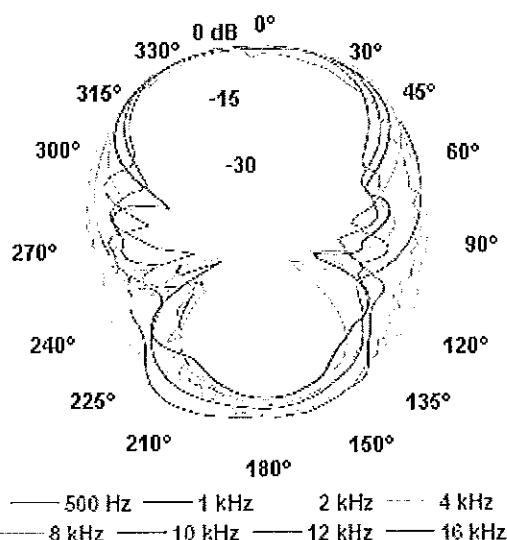


Obr.8 Směrové charakteristiky měřicího mikrofonu NTi 2210

Dalším snímačem byl krátký vlnový mikrofon, tzv. krátká puška, Sennheiser ME36 (obr.9). Jeho směrové charakteristiky jsou na obr.10.



Obr.9 Krátký vlnový mikrofon Sennheiser ME36

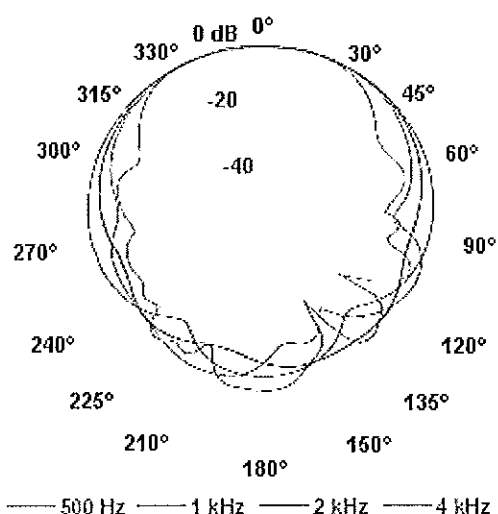


Obr.10 Směrové charakteristiky mikrofonu Sennheiser ME36

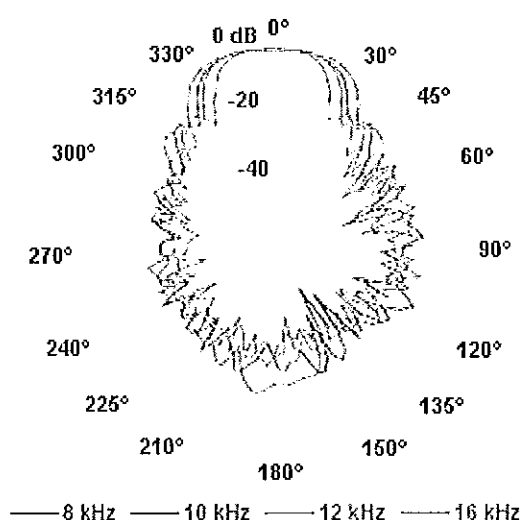
Z projektu specifického výzkumu SV PdF 2124 byla měřicí technika doplněna vysoce směrovým vlnovým mikrofonom Sennheiser MKH70, bývá označován jako tzv. dlouhá puška (obr.11). Směrové charakteristiky jsou na obr.12a, 12b.



Obr.9 Vlnový mikrofon Sennheiser MKH70

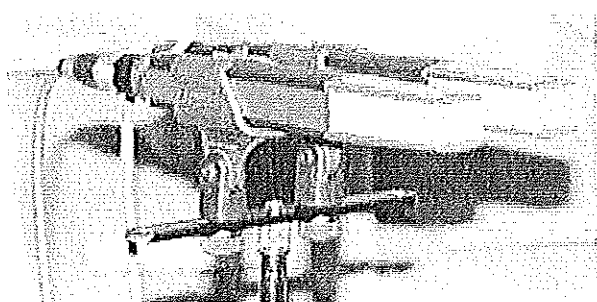


Obr.12a Směrové charakteristiky vlnového mikrofonu Sennheiser MKH70 (0,5-4 kHz)



Obr.12b Směrové charakteristiky vlnového mikrofonu Sennheiser MKH70 (8-16 kHz)

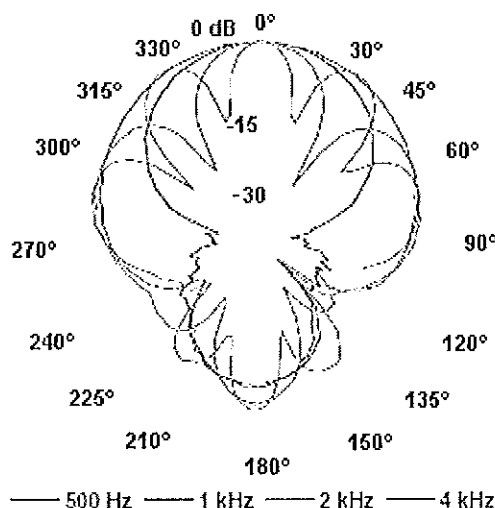
Dalšími akustickými snímači byla dvojice hyperkardiodních mikrofonů Sennheiser MD441 (obr. 13) s distanční vzdáleností 17 cm.



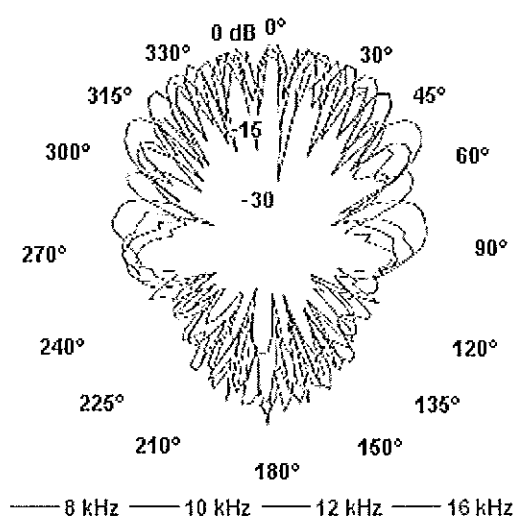
Obr.13 Dvojice mikrofonů Sennheiser MD441

Mikrofonní dvojice byla použita jednak jako dva monaurální snímače pro součtový a rozdílový signál a současně pořízený binaurální záznam simuluje poslech z pozice standardního posluchače.

Směrové charakteristiky mikrofonní dvojice jsou na obr.14a, 14b.



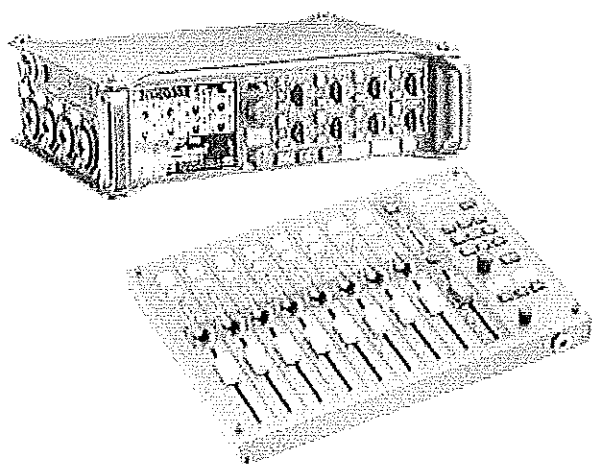
Obr.14a Směrové charakteristiky mikrofonní dvojice Sennheiser MD441 (0,5-4 kHz)



Obr.14b Směrové charakteristiky mikrofonní dvojice Sennheiser MD441 (8-16 kHz)

Měřicí signál STIPA i signály logatomických testů ze všech snímačů byly souběžně zaznamenávány do pěti stop rekordérem ZOOM-F8 (obr. 15) ve standardním formátu DAT 16 bit/48 kHz. To zaručilo identické snímací podmínky, jako je vlastní hluk učebny a rušivé hluky z okolí. Spektrální analýzy byly provedeny rychlou Fourierovou transformací analyzátozem NTi XL2, logatomické testy byly zpracovány poslechem přes

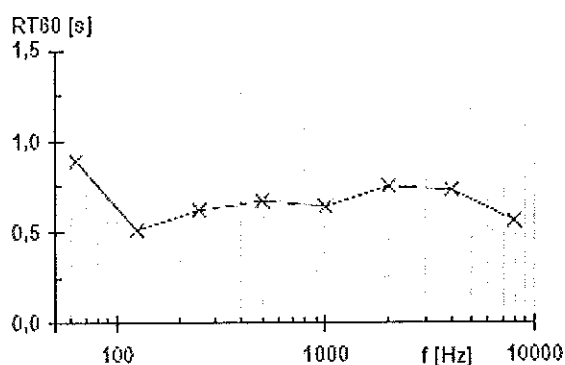
uzavřená studiová sluchátka ARF310 a koaxiální monitory RCF NX10-SMA.



Obr.15 Osmikanálový rekordér ZOOM-F8 s ovládacím pultem FRC-8

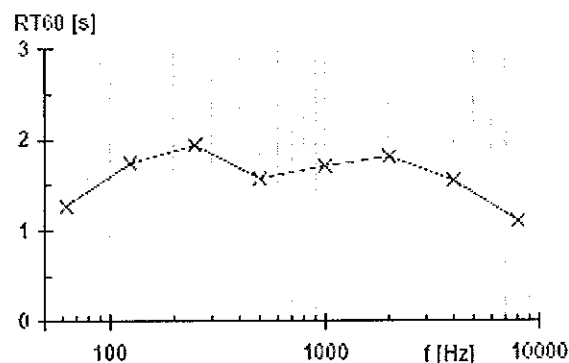
2 SNÍMACÍ PODMÍNKY

Logatomické testy byly snímány ve dvou akusticky rozdílných prostředích. Prostor S51 je seminární učebna s objemem $144,6 \text{ m}^3$ o rozměrech ($l \times b \times h$) $8,10 \times 5,95 \times 3,0 \text{ m}$ a střední dobou dozvuku 670 ms (obr.16) v nově postavené budově Přírodovědecké fakulty Univerzity Hradec Králové. Hladina ambientního hluku se v době měření pohybovala v rozmezí 36-40 dB(Z).



Obr.16 Doba dozvuku učebny S51

Prostor C4 je učebna v historické budově Pedagogické fakulty Univerzity Hradec Králové s objemem $273,6 \text{ m}^3$ o rozměrech ($l \times b \times h$) $9,25 \times 7,18 \times 4,12 \text{ m}$ a střední dobou dozvuku 1,72 s (obr.17). Hladina ambientního hluku se v době měření pohybovala v rozmezí 38-42 dB(Z).



Obr.17 Doba dozvuku učebny C4

Referenční akustický zářič NTi TalkBox s akustickou osou ve výšce 150 cm simuloval stojícího řečníka. Snímače byly instalovány ve vzdálenosti 1,5 m ve výšce uší sedícího posluchače, tj. ve výšce 115 cm. Následně byly instalovány v poslední řadě učeben, tj. ve vzdálenosti 6,5 m pro učebnu S51 a 7,25 m pro učebnu C4. Vlnový snímač MKH70 byl při snímání z blízka upevněn na mikrofonní tyči v odpruženém závěsu ve výšce 225 cm s akustickou osou skloněnou ve směru k akustickému zářiči pod úhlem 45° .

Všechny snímače byly připojeny symetrickým nízkofrekvenčním kabelem k rekordéru ZOOM-F8 a pro všechna měření byl splněn požadavek ISO 9921 [4], aby hladina akustického tlaku rušivého signálu byla minimálně o 26 dB nižší než hladina akustického tlaku užitečného (měřicího) signálu.

3 INDEX PŘENOSU ŘEČI STIPA A SROZUMITELNOST CIS

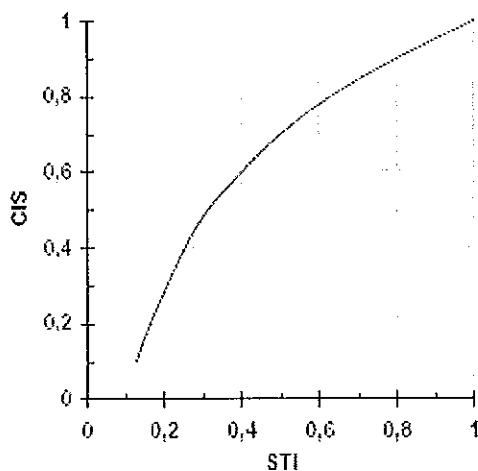
Pro měření indexu přenosu řeči pomocí elektroakustických přenosových systémů je určena metoda STIPA (speech transmission index for public address systems). STIPA používá sedm oktaвовých pásem a pro každé pásmo jsou současně použity dvě jedinečné modulační frekvence (tab.2) [5].

Tab.2 Modulační frekvence pro metodu STIPA

Modulační frekvence [Hz]	Oktávnové pásmo [Hz]						
	125	250	500	1k	2k	4k	8k
1.	1,60	1,00	0,63	2,00	1,25	0,80	2,50
2.	8,00	5,00	3,15	10,0	6,25	4,00	12,5

Metoda STIPA je v současné podobě validována pouze pro spektrum mužské řeči a doba měření je přibližně 15-20 s. Pro konverzi naměřených

hodnot STIPA na společné měřítko srozumitelnosti (speech intelligibility) CIS jsme podle grafu uvedeném v ČSN EN 60849 [13] odvodili analytické řešení pro konverzi STIPA/CIS [14] (obr. 17).



Obr.17 Konverze stupnic STI/CIS

Konverzní křivku v grafu na obr.17 lze podle logaritmické aproximace posadit rovnici (2) s hodnotou spolehlivosti $R^2 = 0,9998$.

$$CIS = 0,4387 \cdot \ln STI + 1,0018 \quad (2)$$

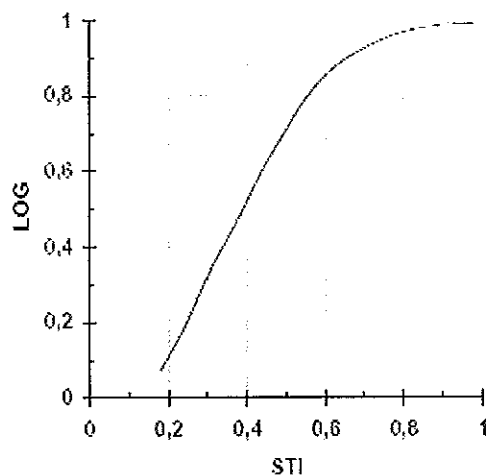
Pro hodnoty $STI \geq 0,25$ můžeme použít i polynom třetího stupně (3) s hodnotou spolehlivosti $R^2 = 0,9997$

$$CIS = 0,9647 \cdot STI^3 - 2,5056 \cdot STI^2 + 2,6678 \cdot STI - 0,1263 \quad (3)$$

Na základě normovaných hodnot konverzních křivek různých metod do CIS [13] jsme odvodili pro nás podstatné vzájemné konverze mezi logatomickou poznatelností (LOG) a indexem přenosu řeči (STI). Konverzi z indexu přenosu řeči STIPA na logatomickou poznatelnost (obr.18) lze s vysokou přesností aproximovat polynomem šestého stupně (4) s hodnotou spolehlivosti $R^2 = 0,9999$.

$$\begin{aligned} LOG = & -20,783 \cdot STI^6 + 70,377 \cdot STI^5 - \\ & -88,867 \cdot STI^4 + 50,558 \cdot STI^3 - \\ & -13,606 \cdot STI^2 + 3,6785 \cdot STI - 0,3668 \end{aligned} \quad (4)$$

V tabulce 3 jsou uvedeny výsledky měření metodou STIPA, s uvedením klasifikační třídy podle ČSN EN 60268-16 [5] a konverze srozumitelnosti CIS, která je v tabulce uvedena v procentech.



Obr.18 Konverze indexu přenosu řeči (STI) na logatomickou poznatelnost (LOG)

Tab.3 Výsledky měření podle ČSN EN 60268-16

Učebna l_r	Snímač	Srozumitelnost (ČSN EN 60268-16)		
		STIPA	Class	CIS [%]
S51 1,5 m	MKH70	0,93	A+	97
	NTI 2210	0,77	A+	88
	ME36	0,84	A+	93
	2MD441	0,84	A+	93
S51 6,5 m	MKH70	0,75	A	88
	NTI 2210	0,68	C	83
	ME36	0,74	A	87
	2MD441	0,75	A	88
C4 1,5 m	MKH70	0,85	A+	93
	NTI 2210	0,67	C	83
	ME36	0,79	A+	90
	2MD441	0,83	A+	92
C4 7,25 m	MKH70	0,61	D	79
	NTI 2210	0,51	G	71
	ME36	0,53	F	72
	2MD441	0,58	E	77

l_r je vzdálenost mezi referenčním zdrojem a snímačem

Z naměřených hodnot je zřejmé, že lepších výsledků dosahují snímače s velkou směrovostí jako je vlnový mikrofón Sennheiser MKH70 nebo dvojice superkardioidních mikrofónů Sennheiser MD441 s distanční vzdáleností 17 cm, která využívá zúžení směrové charakteristiky (obr.14a, 14b) v důsledku součtu fázově posunutých signálů podle normované rovnice (5). Její podrobné odvození je uvedeno v [15].

$$A = 0,25 + 0,75 \cos \beta_1 +$$

$$k_p (0,25 + 0,75 \cos \beta_2) \cos 2\pi \frac{f}{c} (r_2 - r_1) \quad (5)$$

Naměřené hodnoty potvrzují prediktivní odhad, že zúžení směrové charakteristiky snímacího dvojčete potlačí signály, které přicházejí z míst

mimo akustickou osu (zejména odražený zvuk z bočních směrů) [16].

4 VÝSLEDKY LOGATOMICKÝCH TESTŮ

Testy logatomické poznatelnosti představují nej-
přísnější kritérium kvality přenosu řečového sig-
nálu s největší rozlišovací schopností. Při dodrže-
ní jejich metodiky a správném provedení dávají
testy logatomické poznatelnosti zcela objektivní
výsledky. Logatomické testy jsou mimořádně cit-
livé na frekvenční omezení v přenosovém kaná-
lu, nelineární zkreslení, dobu dozvuku a odstup
užitečného signálu od rušivého pozadí. Při pou-
žití referenčního akustického zářiče a profesio-
nálně připravených testů je zaručena jejich 100%
opakovatelnost. Stejně tak záznam pořízený ně-
kolika snímači nebo mikrofonním polem zajišťuje
identické poslechové podmínky (dozvuk, hluk
okolí, atd.) pro hodnocení skupinou posluchačů,
což je důležité pro hodnocení kvality přenosu ře-
či při on-line přenosech nebo poslechu ze zázna-
mu.

Známe-li parametry prostoru, akustického zářiče
a akustického snímače, můžeme predikovat ztrá-
tu srozumitelnosti souhlásek ζ , která přímo sou-
visí se ztrátou srozumitelnosti řeči, podle rovnice
(6)

$$\zeta = \frac{200T^2 l_{tr}^2}{lbhQ_t Q_r} [\%] \quad (6)$$

kde T je doba dozvuku v daném prostoru [s], l_{tr}
je vzdálenost vysílač - přijímač [m], lbh délka,
šířka a výška prostoru [m], Q_t činitel směrovosti
vysílače a Q_r činitel směrovosti přijímače. Do
rovnice (6) je možné zavést řadu korekčních fak-
torů v závislosti na šířce pásma, odstupu rušivé-
ho pozadí, atd. Logatomickou poznatelnost (sla-
bikovou srozumitelnost) lze potom odhadnout na
hodnotu

$$LOG \cong 100 - \frac{400T^2 l_{tr}^2}{lbhQ_t Q_r} [\%] \quad (7)$$

Pro výpočet podle rovnice (7) uvažujeme pro
referenční zářič NTi TalkBox s činitelem směro-
vosti vysílače $Q_t = 2$. Činitelé směrovosti jedno-
tlivých snímačů byly určeny z technických údajů
výrobce a hodnot naměřených v bezodrazové
komoře akustických laboratoří FEL ZČU v Plzni
a jsou uvedeny v tabulce 4.

Tab.4 Činitelé směrovosti použitých snímačů

Snímač	Činitel směrovosti Q_r
MKH70	18
NTi 2210	1,78
ME36	4
2MD441	10

V tabulce 5 jsou uvedeny výsledky poslecho-
vých testů pro hlasové signály J (vyšší mužský
hlas), L (ženský hlas) a K (nižší mužský hlas).
Dále je uvedena predikce srozumitelnosti určená
výpočtem podle rovnice (7) pro $Q_t = 2,5$ a Q_r po-
dle tabulky 4 a predikce srozumitelnosti určená
přepočtem z indexu přenosu řeči STIPA (tab.3)
podle rovnice (4) a grafu na obr.18.

Tab.5 Porovnání logatomické poznatelnosti,
predikce podle rovnice (4) a indexu
přenosu řeči STIPA

Učebna l_{tr}	Snímač	Srozumitelnost [%]				
		Hlasový signál			predikce podle (7)	predikce podle (4)
		J	L	K		
S51 1,5 m	MKH70	97	97	96	100	99
	NTi 2210	97	97	95	99	96
	ME36	97	96	95	100	98
	2MD441	96	95	95	100	98
S51 6,5 m	MKH70	96	96	94	99	92
	NTi 2210	92	91	90	83	92
	ME36	96	93	93	95	95
	2MD441	95	96	97	97	95
C4 1,5 m	MKH70	96	96	96	100	98
	NTi 2210	96	95	96	97	91
	ME36	96	95	92	99	96
	2MD441	96	96	97	100	97
C4 7,25 m	MKH70	95	97	96	93	86
	NTi 2210	82	82	80	37	73
	ME36	91	87	91	71	76
	2MD441	95	95	96	88	83

Výsledky v tabulce 5 reprezentují vyhodnocení
tzv. surového záznamu, bez jakýchkoliv dalších
úprav pomocí filtrů, korekcí, dynamických pro-
cesorů, atd. Ve výsledcích nejsou korigovány ani
fonémové chyby, které vznikají z podobností ve
znělosti některých hlásek a slabik (například d-t,
ž-š, y-i, mně-mě...).

Přestože poslechové testy byly realizovány na
velmi malém vzorku posluchačů a s omezením
pouze na dva, do jisté míry typické, příklady vý-
ukových prostorů, můžeme z dosažených výsled-
ků odvodit, že je velmi obtížné dosáhnout poža-
dované kvality zvukového materiálu, který se má
dále distribuovat. Srozumitelnost řeči (logato-

mická poznatelnost) by zejména pro výukové materiály měla dosahovat minimálně 95 %, což v kontextovém vnímání slovního projevu zajistí téměř 100% větnou srozumitelnost. Přes zdánlivě nepatrné rozdíly ve výsledcích poslechových testů je výsledovatelná závislost, jak na typu řečového signálu, tak na směrových charakteristikách, respektive na činitelích směrovosti použitých snímačů. Nezanedbatelný vliv má také subjektivní vnímání úrovně a charakteru rušivého pozadí i to, jestli je záznam monaurální nebo binaurální (stereofonní, snímáný z pozice posluchače).

5 DISKUSE VÝSLEDKŮ - SUBJEKTIVNÍ HODNOCENÍ KVALITY ZÁZNAMŮ POSLUCHAČI

Z výsledků vyhodnocených logatomických testů (tab.5) můžeme pro dané záznamové podmínky ve dvou zkušebních prostorech odvodit (nikoliv však zevšeobecnit), že v případě krátkých snímacích vzdáleností a při maximální dosažitelné technické kvalitě reprodukce i záznamu jsou rozdíly v závislosti na směrových charakteristikách snímačů zanedbatelné. Podobně je při dokonalé výslovnosti řečníka zanedbatelný i vliv spektrálního složení řeči. Z pohledu referenčního zdroje zvuku můžeme s využitím poloměru dozvuku

$$r_d = 0,057 \sqrt{\frac{Ibh}{T}} \quad (8)$$

a známém činiteli směrovosti referenčního zářiče Q_t určit dozvukovou vzdálenost v daném prostoru pro všesměrový snímač

$$l_{d0} = 0,057 \sqrt{Q_t \cdot \frac{Ibh}{T}} \quad (9)$$

a s uvažováním činitele směrovosti přijímače Q_r potom pro jednotlivé snímače dostaneme dozvukové vzdálenosti

$$l_d = 0,057 \sqrt{Q_t Q_r \cdot \frac{Ibh}{T}} \quad (10)$$

Dosadíme-li do rovnice (10) $Q_t = 2$ a hodnoty Q_r z tabulky 4, dostaneme pro použité snímače dozvukové vzdálenosti podle tabulky 6. Doplněna je i dozvuková vzdálenost pro běžného posluchače s činitelem směrovosti $Q_r = 1,7-2,5$.

Tab.6 Dozvukové vzdálenosti ve zkušebních prostorech

Snímač	Činitel směrovosti Q_r	Dozvuková vzdálenost [m]	
		S51	C4
MKH70	18	4,99	4,22
NTI 2210	1,78	1,58	1,34
ME36	4	2,36	2,00
2MD441	10	3,74	3,17
posluchač	1,7-2,5	1,40-1,87	1,19-1,58

Z údajů v tabulce 6 je zřejmé, že záznamy pořízené vysoce směrovými mikrofony (MKH70 a dvojice MD441) ve vzdálenosti 1,5 m od zdroje zvuku (1. řady učeben) bezpečně spadají do tzv. volného pole. Naproti tomu všesměrový mikrofon a posluchač jsou při této vzdálenosti na hranici volného a difúzního pole. V obou prostorech jsou poslední řady (6,5 m a 7,25 m) výhradně v difúzním poli. Stejně tak i vysoce směrové mikrofony jsou za hranicí dosahu volného pole a v pořízeném záznamu se výrazně uplatňuje dozvuk prostoru.

Zajímavé výsledky přineslo subjektivní hodnocení kvality záznamu při různých typech modelového poslechu v domácích podmínkách: poslech na uzavřená studiová sluchátka, poslech v blízkém poli a poslech ve volném poli přes koaxiální monitory RCF NX10-SMA (tab.7). Posluchači hodnotili kvalitu podle stupnice ECTS ve stupních A (výborná) - F (nevyhovující).

Tab.7 Subjektivní hodnocení kvality záznamu

Poslech/Učebna/ l_r	Subjektivní hodnocení záznamu			
	Snímač			
	MKH70	NTI 2210	ME36	2MD441
ARF/S51/1,5 m	A	A	A	A
ARF/S51/6,5 m	A	C	C	B
ARF/C4/1,5 m	A	C	B	B
ARF/C4/7,25 m	C	E	D	C
RCF-b/S51/1,5 m	A	B	A	A
RCF-b/S51/6,5 m	A	C	B	B
RCF-b/C4/1,5 m	B	C	B	B
RCF-b/C4/7,25 m	C	E	D	C
RCF-v/S51/1,5 m	A	B	B	A
RCF-v/S51/6,5 m	A	C	B	B
RCF-v/C4/1,5 m	B	D	C	B
RCF-v/C4/7,25 m	C	F	E	D

ARF - sluchátka
RCF-b - blízké pole
RCF-v - volné pole

Pozitivní hodnocení záznamů pořízených vysoce směrovým mikrofonom Sennheiser MKH70 se dalo celkem logicky očekávat. Vysoce směrový

mikrofon dokáže i v difúzním poli částečně potlačit boční a zadní odrazy. Odstup dozvuku tak může dosáhnout až -2,5 dB oproti všesměrovému snímači.

Neočekávané bylo vcelku příznivé hodnocení binaurálního záznamu z dvojice superkardioidních mikrofonů Sennheiser MD441 s distanční vzdáleností 17 cm. Zejména při poslechu na sluchátka nebo při poslechu v blízkém poli, kdy byl subjektivní vjem při poslechu záznamu v prostoru C4/7,25 m hodnocen jako velmi dobrý i přes výraznou dozvukovou složku.

ZÁVĚR

Prvotní výsledky měření ve dvou konkrétních zkušebních prostorech ukázaly na problematiku predikce reálně dosažitelné logatomické poznatelnosti, a to jak její stanovení výpočtem z parametrů prostoru, tak přepočtem ze změřeného indexu přenosu řeči. Vzhledem k malému vzorku posluchačů a dvěma zkušebními prostorům nelze výsledky zevšeobecňovat. Musíme rovněž res-

pektovat vysokou adaptabilitu sluchu na různá akustická prostředí. Nicméně můžeme v daných mezích konstatovat následující fakta.

Přiměřená shoda prediktivní a reálné logatomické poznatelnosti je při frekvenčně neomezeném přenosu, krátké době dozvuku nebo snímání ve volném poli pouze v rozsahu 95-100 %.

Mimořádně příznivé výsledky poslechových testů jsou dány použitím špičkové profesionální snímací a záznamové techniky, nekomprimovaným záznamem a poslechem přes uzavřená studiová sluchátka nebo špičkové profesionální koaxiální monitory. Cílem dalšího rozsáhlého výzkumu je stanovení konverzních křivek mezi indexem přenosu řeči a logatomickou poznatelností při uvažování omezených frekvenčních charakteristik přenosového kanálu a dynamické komprese.

Článek byl zpracován s podporou projektu specifického výzkumu SV PdF 2124/2017.

Použité zdroje

- [1] *International Consumer Electronics Show (CES)*. Reuters. 14. January 2015.
- [2] DRTINA, R. *Možnosti a omezení elektronické podpory kvality vzdělávání*. Praha. ExtraSystem. 2011. ISBN 978-80-87570-01-2.
- [3] ČSN EN ISO 9001. *Systémy managementu kvality - Požadavky*. Praha. ČNI. 2016.
- [4] ČSN EN ISO 9921. *Ergonomie - Hodnocení řečové komunikace*. Praha. ČNI. 2004.
- [5] ČSN EN 60268-16 ed. 2. *Elektroakustická zařízení - Část 16: Objektivní hodnocení srozumitelnosti řeči indexem přenosu řeči*. Praha. ČNI. 2004.
- [6] ČSN ISO 5725-1. *Přesnost (správnost a shodnost) metod a výsledků měření - Část 1: Obecné zásady a definice*. Praha. ČNI. 1997.
- [7] ČSN ISO 5725-2. *Přesnost (správnost a shodnost) metod a výsledků měření - Část 2: Základní metoda pro stanovení opakovatelnosti a reprodukovatelnosti normalizované metody měření*. Praha. ČNI. 1997.
- [8] ČSN ISO 21748. *Návod pro použití odhadů opakovatelnosti, reprodukovatelnosti a pravdivosti při odhadování nejistoty měření*. Praha. ČNI. 2012.
- [9] ČSN ISO 2602. *Statistická interpretace výsledků zkoušek. Odhad průměru. Konfidenční interval*. Praha. ČNI. 1993.
- [10] ČSN ISO 2854. *Statistická interpretace údajů. Odhady a testy středních hodnot a rozptylů*. Praha. ČNI. 1994.
- [11] DRTINA, R. - SEDIVÝ, J. - SCHLOSSER M. *The Objectivization of the Entry Conditions of the Pedagogical Research*. IEEE Education and modern educational technologies (EMET 2015). IEEE Press. 2015. pp.25-28.
- [12] DRTINA, R. et al. *Objektivizace vstupních podmínek pedagogického výzkumu - Testy srozumitelnosti řeči pro zvukové výukové materiály*. Media4u Magazine. 1/2015. s.19-27. ISSN 1214-9187.
- [13] ČSN EN 60849. *Nouzové zvukové systémy*. Praha. ČNI. 1999.
- [14] DRTINA, R. et al. *STIPA method in public address sound systems and voice alarm systems - Part 2: Conversion curves and bandwidth effect*. Trans Tech Publications. Advanced Materials Research Vol. 1082 (2015). pp.574-580.
- [15] TUREČEK, O. - FEJFAR, P. - DRTINA, R. *Využití výsledků výzkumu a vývoje ve výuce. Část 5: Vliv směrových charakteristik mikrofonní dvojice MD441 na měřený index přenosu řeči STIPA a srozumitelnost CIS*. Media4u Magazine. 4/2015. s.84-95. ISSN 1214-9187.
- [16] DRTINA, R. et al. *Effect of the directional characteristics of the microphone pair MD441 on the measurement of the speech transmission index STIPA and comprehensibility CIS*. Electromechanical Control Technology and Transportation (ICECTT 2015). Atlantis Press. 2015. pp.291-294.

Kontaktní adresy

doc. dr. René Drtina, Ph.D.
doc. Ing. Jaroslav Lokvenc, CSc.
Bc. Radka Vohralíková
Mgr. Pavel Fejfar

e-mail: rene.drtina@uhk.cz
e-mail: jaroslav.lokvenc@uhk.cz
e-mail: radka.vohralikova.3@uhk.cz
e-mail: pavel.fejfar@uhk.cz

Katedra technických předmětů
Pedagogická fakulta
Univerzita Hradec Králové
Rokilanského 62
500 03 Hradec Králové