

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE

Brno, 2019

Martin Kobližka



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

## ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

## ANALÝZA MASKOVÁNÍ ŘEČI HLUKEM V HROMADNÝCH DOPRAVNÍCH PROSTŘEDCÍCH

ANALYSIS OF SPEECH MASKED BY NOISE IN PUBLIC TRANSPORT VEHICLES

### BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

### AUTOR PRÁCE

AUTHOR

Martin Koblížka

### VEDOUCÍ PRÁCE

SUPERVISOR

MgA. Mgr. Ondřej Jirásek, Ph.D.

BRNO 2019

# Bakalářská práce

bakalářský studijní obor **Audio inženýrství**  
Ústav telekomunikací

**Student:** Martin Kobližka

**ID:** 186617

**Ročník:** 3

**Akademický rok:** 2018/19

## NÁZEV TÉMATU:

### **Analýza maskování řeči hlukem v hromadných dopravních prostředcích**

#### **POKYNY PRO VYPRACOVÁNÍ:**

Zaznamenejte reálné hlukové pozadí nejméně ve dvou druzích hromadných dopravních prostředků (autobus, trolejbus, metro, vlak) a to různě sytého spektra a různé hlasitosti. Při nahrávání hlukového pozadí zároveň změřte jeho hladinu akustického tlaku v celém slyšitelném spektru a v oktávových a třetino-oktávových pásmech. Na základě těchto měření spočítejte hlasitost podle Stevense a Zwickera.

Prostudujte metody subjektivního hodnocení srozumitelnosti řeči pomocí logatomů, slov a vět, např. DRT, SPIN a HINT a zvolte metodu vhodnou pro daný účel. V bezodrazové místnosti nahrajte testovací promluvy podle metodiky vámi zvolené metody hodnocení srozumitelnosti řeči a to nejméně nejméně se třemi typy lidských hlasů v subjektivní dynamice mluvcích – tiché, střední a silně hlasité. Zároveň změřte hladinu akustického tlaku promluv. Nahraná hluková pozadí zkombinujte s nahranými promluvami v suché podobě a s různým dozvukem odpovídajícím rozměru a odrazivosti hromadného dopravního prostředku. Popište spektra hlukových pozadí i hlasů, vypočítejte maskovací práh hlukových pozadí a pro vytvořené kombinace hlukových pozadí a mluvených frází objektivně vyhodnoťte míru maskování frází.

#### **DOPORUČENÁ LITERATURA:**

- [1] MELKA, A. Základy experimentální psychoakustiky. Praha: Akademie múzických umění, 2005. ISBN 80-733-043-0
- [2] LOIZOU, P.C.: Speech Enhancement - Theory and Practice. CRC Press, Boca Raton, 2007. ISBN 978--8493-5032-0

**Termín zadání:** 1.2.2019

**Termín odevzdání:** 27.5.2019

**Vedoucí práce:** MgA. Mgr. Ondřej Jirásek, Ph.D.

**Konzultant:** doc. Ing. Jiří Schimmel, Ph.D.

**prof. Ing. Jiří Mišurec, CSc.**  
*předseda oborové rady*

#### **UPOZORNĚNÍ:**

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## **ABSTRAKT**

Tato práce se zabývá problematikou hluku v dopravních prostředcích a maskování řeči. Teoretická část se zabývá percepcí zvuku z hlediska hlasitosti a kmitočtu. Praktická část zahrnuje nahrávání hluku uvnitř dopravních prostředků a stanovení hlasitosti a maskovacích prahů metodou podle Zwickera s využitím prostředí Matlab. Dále nahrávání řeči v různých dynamikách, poslechové testy a jejich vyhodnocení z hlediska srozumitelnosti řeči.

## **KLÍČOVÁ SLOVA**

Hluk, maskování, spektrum, kritická pásma, hlasitost podle Zwickera, Matlab, srozumitelnost řeči, poslechové testy

## **ABSTRACT**

This work deals with the problems of noise and speech masking in public transport . The theoretical is focused on the perception of sound in terms of loudness and frequency. The practical part includes recording noise inside public transport vehicles and determining the loudness and masking thresholds by the Zwicker method using Matlab. Furthermore, speech recording in different dynamics, listening tests and their evaluation in terms of speech intelligibility.

## **KEYWORDS**

Noise, masking, spectrum, critical bands, Zwicker's loudness, Matlab, speech intelligibility, listening tests

KOBLIŽKA, Martin. *Analýza maskování řeči hlukem v hromadných dopravních prostředcích*. Brno, 2019, 43 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: MgA. Mgr. Ondřej Jirásek, Ph.D.

## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Analýza maskování řeči hlukem v hromadných dopravních prostředcích“ jsem vypracoval(a) samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....

podpis autora(-ky)

## PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu MgA. Mgr. Ondřeji Jiráskovi, Ph.D. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno .....

.....

podpis autora(-ky)



Faculty of Electrical Engineering  
and Communication  
Brno University of Technology  
Purkynova 118, CZ-61200 Brno  
Czech Republic  
<http://www.six.feec.vutbr.cz>

## PODĚKOVÁNÍ

Výzkum popsany v této bakalářské práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno .....

.....  
podpis autora(-ky)



EVROPSKÁ UNIE  
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ  
INVESTICE DO VAŠÍ BUDOUCNOSTI



# OBSAH

Úvod	10
<b>1 Zvuk - Fyziologie</b>	<b>11</b>
1.1 Ucho	11
1.1.1 Vnější ucho	11
1.1.2 Střední ucho	11
1.1.3 Vnitřní ucho	12
1.2 Řeč	13
1.3 Hlasitost	14
1.4 Maskování	16
1.4.1 Maskování kmitočtové	17
1.4.2 Maskování dočasné	17
1.5 Kritická pásma	18
1.6 Hlasitost podle Zwickera	19
1.7 Hluk	21
1.8 Hluková expozice	22
<b>2 Metody hodnocení srozumitelnosti řeči</b>	<b>23</b>
2.1 Subjektivní metody	23
2.1.1 DRT	23
2.1.2 SPIN(Speech Perception in Noise)	24
2.1.3 HINT(Hearing in Noise Test)	24
2.1.4 SRT	24
2.1.5 Hodnocení kvality	24
2.2 Objektivní metody	25
2.2.1 SNR (Signal to Noise Ratio)	25
2.2.2 STI	26
<b>3 Praktická část</b>	<b>27</b>
3.1 Nahrávání prostředí	27
3.1.1 Tramvaj	27
3.1.2 Metro	27
3.1.3 Autobus	28
3.2 Výsledky měření prostředí	28
3.2.1 Tramvaj	28
3.2.2 Metro	30
3.2.3 Autobus	32



3.3	Souhrn . . . . .	34
3.4	Nahrávání řeči . . . . .	35
3.5	Poslechové testy . . . . .	37
3.5.1	Měření sluchátek . . . . .	37
3.5.2	Výsledky testů . . . . .	38
<b>4</b>	<b>Závěr</b>	<b>40</b>
	<b>Literatura</b>	<b>41</b>
<b>A</b>	<b>Obsah přiloženého CD</b>	<b>43</b>

# SEZNAM OBRÁZKŮ

1.1	Lidské ucho [6]	11
1.2	Frekvenční citlivost a maxima kmitání bazilární membrány [6]	13
1.3	Oblast lidského slyšení [6]	14
1.4	Křivky stejné hlasitosti[1]	15
1.5	Závislost hlasitosti tónu 1000 Hz na hladině akustického tlaku[1]	16
1.6	Hladiny maskování [4]	17
1.7	Časové maskování [4]	18
1.8	Tabulka kritických pásem	19
1.9	Příklad grafického výstupu Zwickerovy metody [4]	20
1.10	Graf specifických hlasitostí kritických pásem [8]	20
1.11	Graf kmitočtových charakteristik váhových filtrů[13]	22
3.1	Nahrávací řetězec	27
3.2	Tramvaj T3R.P	28
3.3	Tramvaj 15T na úseku po bezžlábkových kolejích	29
3.4	Zvukové energetické spektrum prostředí tramvaje při jízdě přes výhybky	30
3.5	Metro typ 81-71M	31
3.6	Zvukové energetické spektrum uvnitř metra	32
3.7	Energetické spektrum hnědého šumu	32
3.8	Bus BN12	33
3.9	Energetické spektrum autobusu	34
3.10	Tabulka naměřených maximálních hodnot	34
3.11	Tabulka naměřených průměrných hodnot	35
3.12	Spektrum žena (nahore) a muž	36
3.13	Srovnání spekter při rozdílných dynamikách	37
3.14	Kmitočtová odezva sluchátek k poslechovému testu	38
3.15	Tabulka průměrů rozpoznávaných slov	38

# ÚVOD

Hluk je nedílnou součástí našeho života. Ať už se jedná o stavební činnosti, dopravu nebo zábavu. Je hluk spjat s technikou, která se vyvíjí mílovými kroky kupředu už od průmyslové revoluce v 19.století. Bylo vypracováno mnoho studií na téma negativního vlivu hluku na člověka, z těch pak vzešly limity, které jsou v ČR stanoveny zákonem. Tyto limity jsou ale především stanoveny pro obytné oblasti, pracovní prostředí školy a nemocnice. Netýkají se interiéru dopravních prostředků. Právě dopravním prostředkům se ale v současnosti prakticky nevyhneme. Automobilky už se na tento problém zaměřily, avšak hromadná doprava má stále v této oblasti nedostatky.

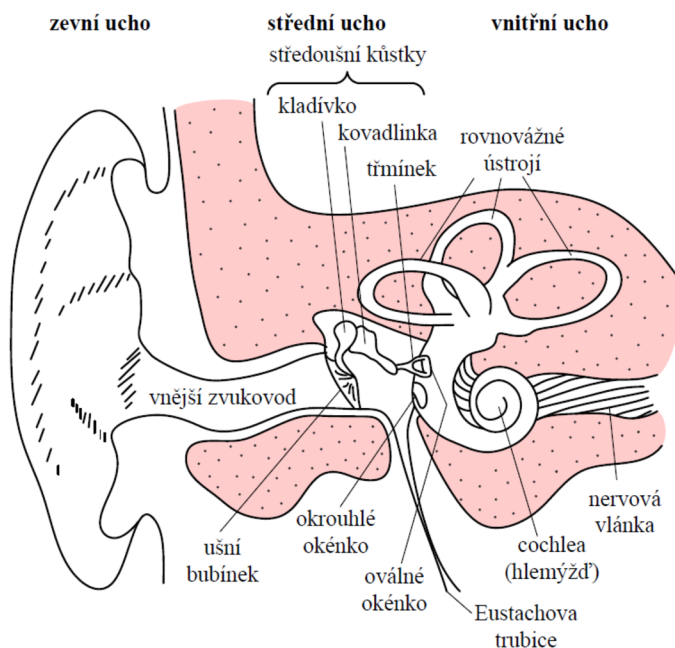
Tato práce se věnuje nahrávání a měření hluku v dopravních prostředcích a následného vyhodnocení z hlediska hlasitosti a maskování zvuku v oblasti lidské řeči. Na začátku nejprve popíšeme lidské slyšení a v jakých frekvenčních pásmech se pohybuje lidská řeč. Dále rozebereme co je to hlasitost, jaké známe typy maskování a popíšeme si model kritických pásem, který využívá metoda měření hlasitosti podle Zwickera. V další části si rozdělíme metody hodnocení srozumitelnosti řeči.

Praktická část bude zahrnovat zachycení reálného hlukového pozadí v různých dopravních prostředcích za pomoci mikrofonní techniky. Dále také proběhne nahrávání lidské řeči v různé dynamice. Popíšeme spektra jak zachyceného hluku, tak nahraného lidského hlasu. Spočítáme hlasitost zvuku prostředí podle Zwickera. Na základě poslechových testů vyhodnotíme míru maskování řeči v nahrávkách prostředí smíchaných s lidským hlasem.

# 1 ZVUK - FYZIOLOGIE

## 1.1 Ucho

Je párový sluchový orgán. Skládá se ze tří částí. Vnější, střední a vnitřní ucho. Tato kapitola čerpá z [6],[9].



Obr. 1.1: Lidské ucho [6]

### 1.1.1 Vnější ucho

Vnější ucho se skládá z boltce a zvukovodu a je ukončeno bubínkem. Zvukové vlny jsou soustředované boltcem do zvukovodu, jehož nesymetrický tvar způsobuje, že zvukové vlny přicházející z různých směrů jsou mírně odlišné. To je důsledek odrazů od různých povrchů, které pozměňují jeho fázi, amplitudu i frekvenční spektrum. To pomáhá k orientaci v prostoru, protože mozek je potom schopen určit, který zvuk je odražený a který přichází přímo od zdroje. Zvukové vlny potom pokračují dále zvukovodem až k bubínku.

### 1.1.2 Střední ucho

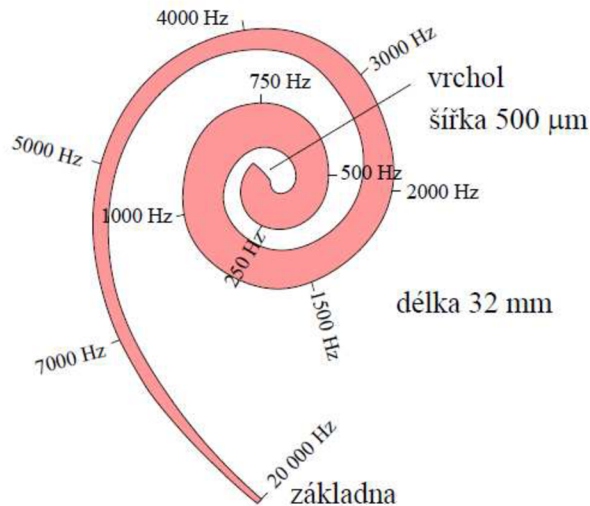
Střední ucho je úzká štěrbinová dutina v kosti spánkové, která je vpředu spojená s nosohltanem Eustachovou trubicí, která vyrovnává tlak na bubínek. Střední část tvoří dutina

bubínková, obsahující tři kůstky – kladívko, kovadlinku a třmínek. Kladívko je svými výběžky spojeno s bubínkem a kovadlinkou, ta je spojena s třmínkem; třmínek je zasazen do předsíňového okénka, vedoucího již do vnitřního ucha. Zvukové vlny procházejí zvukovodem, rozkmitají bubínek a tyto kmity se přenášejí sluchovými kůstkami do vnitřního ucha.

### 1.1.3 Vnitřní ucho

Vnitřní ucho je uloženo v kostěném pouzdru spánkové kosti. Skládá se ze dvou smyslových orgánů, z ústrojí rovnovážného pro vnímání polohy těla a pohybů v prostoru a z vlastního ústrojí pro vnímání zvuků. Ten je uložen v kanálku stočeném do tvaru hlemýždě a má 2,5 závitů délky asi 32 mm. Jeho prostor je rozdělen na dvě patra. V jeho vrcholu jsou obě patra mezi sebou spojena. Dělicí část obou pater je tvořena vlastním smyslovým ústrojím, tzv. Cortiho orgánem. Jeho stavba se podobá trubici. Dolní část je tvořena bazilární membránou, na které jsou ve čtyřech řadách smyslové vláskové buňky. Shora na jejich vlásky dosedá krycí membrána. Proti hornímu patru je Cortiho orgán oddělen Reisnerovou membránou. Všechna ústrojí vnitřního ucha jsou uložena v labyrintové tekutině. Přenos zvukových kmitů se děje touto tekutinou, kterou uvede v kmitání pohyb třmínku v oválném okénku. Při přenosu kmitů do hlemýždě dochází v tekutině ke vzniku postupné vlny, jejíž rychlost se mění. Tato postupná vlna přechází vestibulem na horní patro hlemýždě až do vrcholu, kde přechází do dolního patra. Chování vlny je dáno tlumením v tekutině, pružností tkáně a tvarem prostoru hlemýždě. Pomalá vlna při podnětu tónem nízkého kmitočtu proběhne celým hlemýžděm, rychlá vlna při podnětu tónem vyššího kmitočtu se vyrovná dříve. Slabé podněty dráždí jen malý, silné podněty větší úsek bazilární membrány. Tlaková změna tekutiny se vyrovnává na blance okrouhlého okénka.

Ve vnitřním uchu dochází ke kmitočtové analýze zvuku. Podráždění smyslových buněk je způsobeno výchylkou bazilární membrány, kterou vyvolá pohyb tekutiny při průchodu hlemýžděm. Tento pohyb vede ke změně polohy vlásků vzhledem ke krycí membráně, která je vlastním impulsem pro smyslový vjem. Kmitání třmínku vyvolává kmitání bazilární membrány a pozice maximální výchylky tohoto kmitání je závislá na kmitočtu: hluboké tóny vyvolají maximální výchylku v její vrcholové části, vysoké naopak při základně hlemýždě. Pozice maxima kmitání bazilární membrány v závislosti na kmitočtu není lineární.

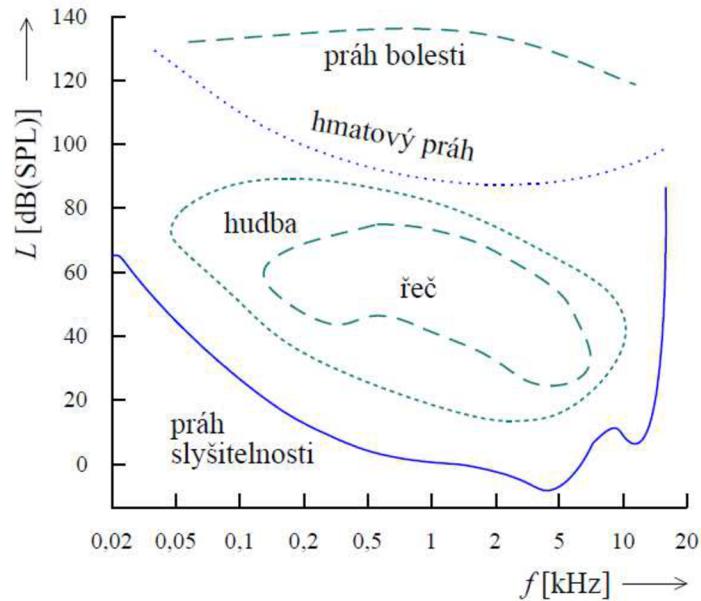


Obr. 1.2: Frekvenční citlivost a maxima kmitání bazilární membrány [6]

Pohybem vláskových buněk dochází ke změně jejich elektrického potenciálu a přeměně na elektrické impulzy, které dále zpracovává mozek a tím vzniká sluchový vjem. Tyto signály jsou mozkiem srovnávány a vyhodnocovány z obou uší současně.

## 1.2 Řeč

Princip tvorby řeči je podobný jako tvoření tónu u dechových hudebních nástrojů. Proud vzduchu dodávaný plícemi prochází skrz hlasivkovou štěrbinu, ve které se rozkmitávají hlasivky, a tím přeměňují proud vzduchu na budící signál. Tento signál je pak zpracován v hlasovém traktu (hrtan, dutina ústní, dutina nosní) pohybem řečových orgánů.[5] Frekvenční rozsah lidského hlasu je od cca 60 Hz, v případě zpěvu, alikvotní složky pak mohou dosahovat až 10 kHz. Řeč se běžně pohybuje v oblasti 100 Hz - 5 kHz.



Obr. 1.3: Oblast lidského slyšení [6]

### 1.3 Hlasitost

Hlasitost definujeme jako subjektivní počitek, odezvu nervového systému, na zvyšující se intenzitu zvukového podnětu. Hlasitost není přesně úměrná intenzitě, i když se vzrůstající intenzitou také roste. Závisí také na kmitočtu, spektrálním obsahu a době trvání zvuku.[4] Hlasitost čistého tónu závisí nejvíce na velikosti jeho akustického tlaku, dále také na jeho kmitočtu a době trvání. Nejstarší veličinou pro určování hlasitosti libovolného zvuku je **hladina hlasitosti** s jednotkou **fón**. Hladina hlasitosti ve fónech se číselně rovná hladině akustického tlaku v dB referenčního zvuku, tvořeného čelně dopadající sinusovou rovinnou postupnou vlnou o kmitočtu 1000 Hz, která je posouzená jako stejně hlasitá jako daný zvuk. Pro hladinu hlasitosti platí vztah

$$L_N = L_{1000} = 20 \log \frac{p_{1000}}{p_0} \quad (1.1)$$

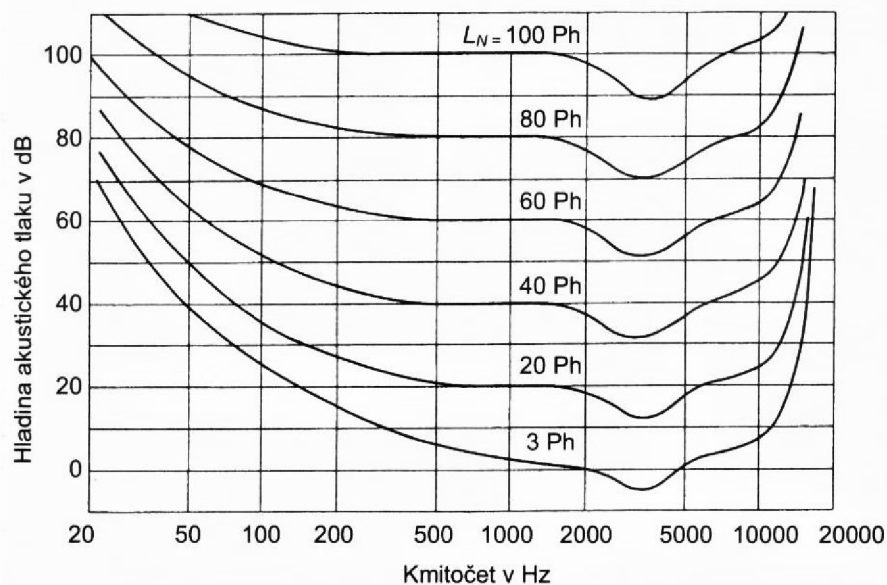
$L_N$  – hladina hlasitosti posuzovaného zvuku ve fónech

$L_{1000}$  – hladina akustického tlaku čistého tónu o kmitočtu 1000 Hz v dB

$p_{1000}$  – akustický tlak čistého tónu o kmitočtu 1000 Hz v pascálech

$p_0$  – referenční akustický tlak 20 Pa

Z toho vyplývá, že stupnice hladiny hlasitosti je napůl fyzikální a napůl psychologickou. Hladinu hlasitosti lze změřit pro jakýkoli zvuk. Podrobně byla změřena pro čisté tóny B.A.Kingsburym, brzy poté H. Fletcherem a W.A. Munsonem a dalšími. Průměrné hodnoty hladiny hlasitosti, získané měřením velkých populací otologicky normálních osob se zobrazují v závislosti na kmitočtu jako tzv. křivky stejné hlasitosti ("Kingsburyho křivky" nebo také "Fletcherovy-Munsonovy křivky") obr.1.4



Obr. 1.4: Křivky stejné hlasitosti[1]

Zatímco hladina hlasitosti se značí  $L$  a měří ve fónech, psychoakustickou veličinu **hlasitost** značíme symbolem  $N$ . Jednotka pak byla nazvána **son**. 1 son je definován jako hlasitost čistého tónu přicházejícího čelně jako rovinná vlna o kmitočtu 1000 Hz a hladině akustického tlaku 40 dB při referenčním akustickém tlaku  $20 \mu\text{Pa}$ [7].

Experimentálně zjištěná, mezinárodně normalizovaná závislost hlasitosti čistého tónu o kmitočtu 1000 Hz v sonech na hladině jeho akustického tlaku v dB je zobrazena v grafu na obr.1.5

Při zobrazení sonové stupnice v logaritmickém měřítku je průběh této závislosti pro hodnoty hladin větší než 40 dB lineární a každému vzrůstu hladiny o 10 dB odpovídá dvojnásobné zvětšení hlasitosti v sonech. Tuto část křivky lze matematicky popsat rovnicí

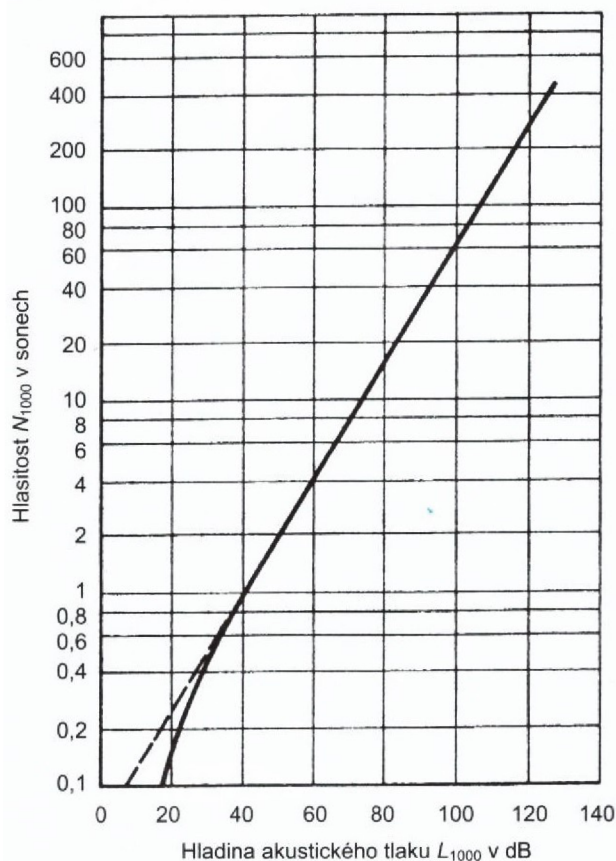
$$N_{1000} = 2^{\frac{L_{1000}-40}{10}} \quad (1.2)$$

$N_{1000}$  – hlasitost zvuku čistého tónu o kmitočtu 1000 Hz v sonech



$L_{1000}$  – hladina jeho akustického tlaku v dB

Tento vztah zobrazuje v obr.1.5 čárkovaná přímka. Při hladinách nižších než 40 dB se empiricky zjištěné hodnoty hlasitosti od tohoto průběhu liší. Křivka hlasitosti se zmenšováním hladiny klesá strměji a její asymptotické pokračování směrem k  $-\infty$  by dosáhlo 0 při prahu slyšení, jemuž z definice odpovídá hlasitost 0 sonů. Jelikož pro hladinu hlasitosti je sinusový tón 1000 Hz také referenčním tónem, lze na grafu na obr.1.5 uvádět na horizontální ose místo hladiny akustického tlaku v dB hladinu hlasitosti ve fónech. Hlasitost tónu o jiném kmitočtu pak lze stanovit z křivek stejné hlasitosti.



Obr. 1.5: Závislost hlasitosti tónu 1000 Hz na hladině akustického tlaku[1]

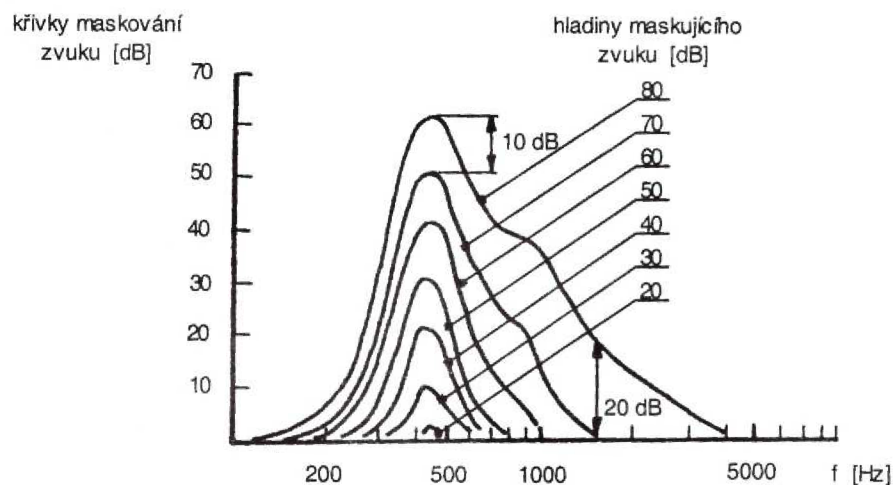
## 1.4 Maskování

Tato sekce vychází z[4], [6] a [10]. Při znění určitého signálu se práh sluchu posouvá k vyšším hladinám. Tóny, které byly slyšet dříve, jsou maskovány. Maskování je jev, při kterém je vjem nějakého zvuku potlačen jiným, obvykle hlasitějším, který se označuje

jako maskující. Maskování sluchových vjemů je dáno způsobem činnosti nervových buněk a vláken. Jsou-li nervové prvky ve stavu podráždění, nejsou schopny po určité době přijmout další podráždění. Maskování se měří posunem sluchového prahu maskovaného zvuku. Maskování je závislé na rozdílu kmitočtů poslouchaných zvuků a na jejich intenzitě. Rozlišujeme základní dva druhy maskování.

### 1.4.1 Maskování kmitočtové

Kmitočtové maskování je jev, při kterém se vyskytuje maskující zvuk ve stejném okamžiku jako maskovaný. V případě tónového zvuku se maskování projevuje v kmitočtovém okolí maskujícího signálu. To, jak velké je maskovací okolí, do velké míry ovlivňuje intenzita maskujícího zvuku, což je způsobeno větší výchylkou baziální membrány. Na obr. 1.6 jsou znázorněny maskovací hladiny stejného signálu se zvyšující se intenzitou.

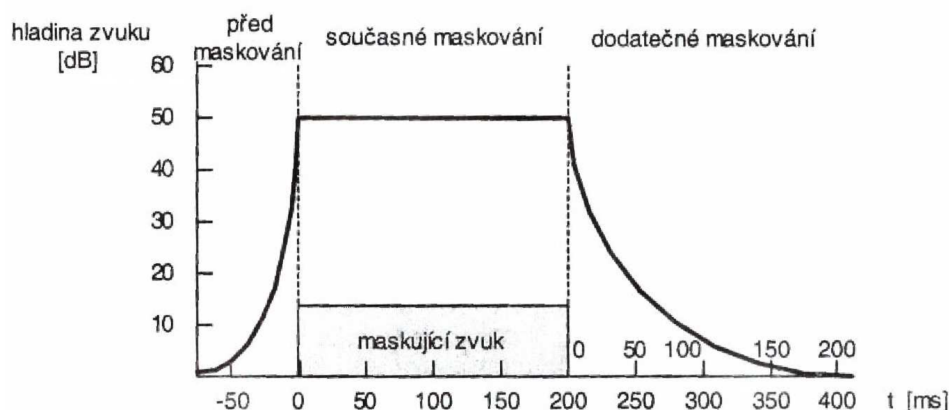


Obr. 1.6: Hladiny maskování [4]

### 1.4.2 Maskování dočasné

Efekt maskování se projevuje nejen při současném znění zvuků, ale také v případě, kdy maskující zvuk zní před a po znění zvuku maskovaného. První případ nazýváme pre-masking (na obr.1.7 před maskování) . Tento efekt se projevuje v řádu několika ms a je způsoben rychlejším zpracováním silnějšího podnětu. Když maskující zvuk časově předchází maskovaný, mluvíme o post-masking (také dodatečné maskování).

Tento jev lze vysvětlit jako dobu, za kterou se nervové buňky mohou znovu vybudit a projevuje se v řádu několika desítek ms.



Obr. 1.7: Časové maskování [4]

## 1.5 Kritická pásma

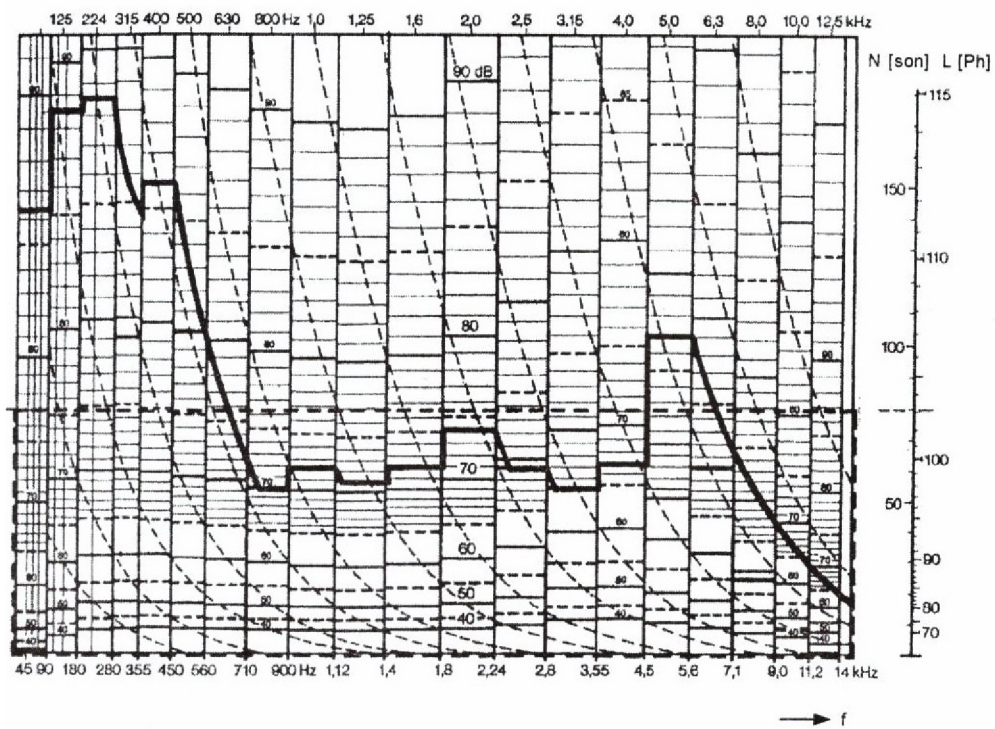
Koncept kritických pásem byl poprvé navržen Harvey Fletcherem. Vychází z předpokladu, že při maskování tónu širokopásmovým šumem se uplatňuje pouze ta část spektra šumu, která leží v okolí maskovaného tónu. Experimenty ukázaly, že při frekvenčním maskování se maskovací efekt v okolí maskujícího prvku zužuje směrem k nižším frekvencím. Bylo zjištěno, že lidské ucho rozděluje zvuk do 24 takzvaných kritických pásem nebo také barků. Uvnitř těchto pásem dochází k nejvýraznějším efektům kmitočtového maskování, protože lidské ucho nedokáže uvnitř těchto kritických pásem jasně rozlišit vícero kmitočtů. Kritická pásma jsou základním rysem lidského slyšení.[6]

barková stupnice	dolní mezní kmitočet pásma	horní mezní kmitočet pásma	střední kmitočet pásma	šířka pásma
[bark]	[Hz]	[Hz]	[Hz]	[Hz]
	0	100	50	100
2	100	200	150	100
3	200	300	250	100
4	300	400	350	100
5	400	510	450	110
6	510	630	570	120
7	630	770	700	140
8	770	920	840	150
9	920	1080	1000	160
10	1080	1270	1170	190
11	1270	1480	1370	210
12	1480	1720	1600	240
13	1720	2000	1850	280
14	2000	2320	2150	320
15	2320	2700	2500	380
16	2700	3150	2900	450
17	3150	3700	3400	550
18	3700	4400	4000	700
19	4400	5300	4800	900
20	5300	6400	5800	1100
21	6400	7700	7000	1300
22	7700	9500	8500	1800
23	9500	12000	10500	2500
24	12000	15500	13500	3500

Obr. 1.8: Tabulka kritických pásem

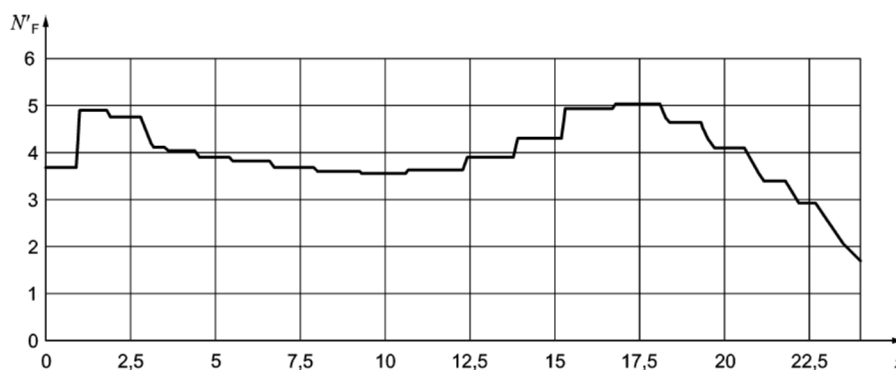
## 1.6 Hlasitost podle Zwickera

Tato metoda je standardizována v mezinárodní normě ISO 532-1[8]. Výchozí veličinou pro výpočet hlasitosti složeného zvuku podle Zwickera je hladina akustického tlaku v kritických pásmech. Používá model hlasitosti navržený Eberhardem Zwickerem, který zohledňuje kritická pásma a maskování. Vychází se se změřených hladin akustického tlaku v třetinooktávních pásmech, které se v dolní části spektra sčítají, aby lépe odpovídaly hladinám akustického tlaku v kritických pásmech. Tyto hodnoty se vynesou do předtištěných grafů 1.9 a stanoví se obsah plochy pod takto vzniklou křivkou. Subjektivní hladina hlasitosti hluku podle tohoto modelu se získá podělením této hodnoty šířkou plochy.



Obr. 1.9: Příklad grafického výstupu Zwickerovy metody [4]

Na obr. 1.10 je zobrazen další možný výstup hlasitosti podle Zwickera. Osa x popisuje jednotlivá kritická pásma (barky). Na ose y je specifická hlasitost jednotlivých barků s jednotkou son/bark, také je možné pro lepší orientaci na ose x uvádět kmitočet v Hz.

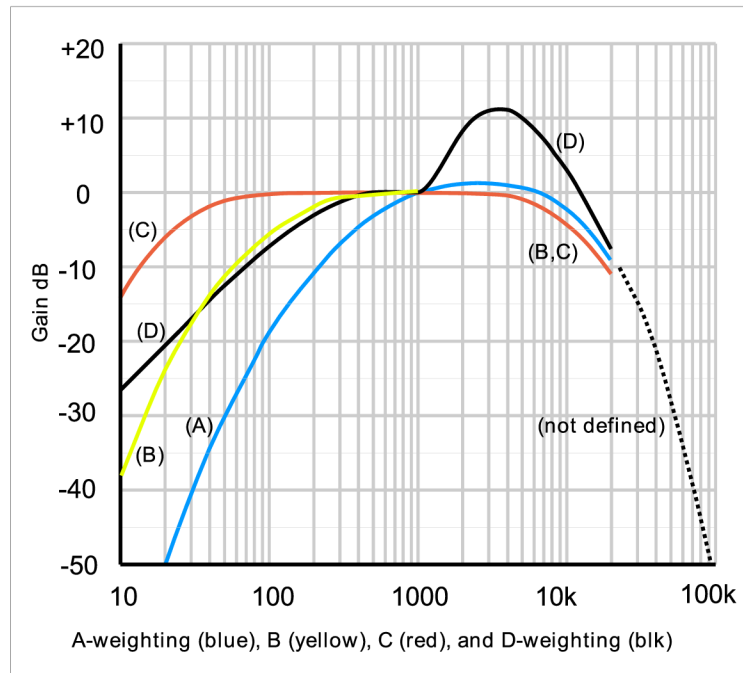


Obr. 1.10: Graf specifických hlasitostí kritických pásem [8]

## 1.7 Hluk

Hlukem označujeme každý nežádoucí, rušivý, nepříjemný anebo škodlivý zvuk, který svým působením vyvolává biologickou reakci člověka nebo jiných živočichů.[3] Svou podstatou je hluk děj stochastický. Dle statistických vlastností jej proto můžeme dělit na stacionární a nestacionární. Stacionární proces je ten, jehož statistické vlastnosti se v čase nemění. Stacionární hluk je např. zvuk motoru při konstantních otáčkách nebo zvuk způsobený valivým odporem při jízdě na kole. Nestacionární hluk pak můž být např. hluk na nádraží způsobený velkým množstvím lidí, veškeré náhodné zvuky, šumy, déšť atd. Z toho vyplývá, že kategorie nestacionárních jevů je rozsáhlejší a vyskytuje se v běžném životě častěji.

Hygienické limity pro hluk a vibrace jsou v České republice stanoveny v Nařízení vlády č. 272/2011 Sb. K měření hluku se většinou používají zvukoměry. Jejich součástí jsou váhové filtry, které upravují naměřené hodnoty podle křivek stejné hlasitosti tak, aby odpovídaly lidskému slyšení. Protože se křivky stejné hlasitosti liší tvarem v závislosti na hlasitosti, mají váhové typy lehce odlišné tvary kmitočtových charakteristik. Nejčastěji se používá typ A a C. Hodnota změřené hlasitosti se pak udává v dB s přidaným A nebo C, podle použitého filtru. Výhodou je, že hodnota hlasitosti změřené s váhovým filtrem je vlastně jedno číslo. Nevýhodou je, že nám jedno číslo nic neřekne o spektru daného zvuku. O rozložení energie ve spektru mají větší výpovědní hodnotu, změřené hodnoty akustických hladin v daných frekvenčních oktávových nebo 1/3 oktávových pásmech.



Obr. 1.11: Graf kmitočtových charakteristik váhových filtrů[13]

## 1.8 Hluková expozice

Hluková expozice je míra dlouhodobého vystavení lidského sluchu hlukům. Určuje, jestli sluch má po expozici hlukem dostatečný čas pro návrat k normálním vlastnostem, tj. zda dochází k vratným změnám (sluchová únava) nebo k nevratným změnám (akustické trauma). Pro 8 hodinovou hlukovou expozici je hranice mezi těmito změnami  $LA_{eq} = 85$  dBA, kde  $LA_{eq}$  je ekvivalentní hladina zvuku A. Hluková expozice se měří zvukovým expozimetrem, který přímo vyčísluje hlukovou expozici vzorkováním hodnot hladin akustického tlaku v pravidelných časových intervalech.[6]

Pokud cestujeme hromadnou dopravou 8 hodinová zátěž se nás většinou netýká, ovšem sluchová únava nastává v řádu jednotek minut, která se projevuje snížením pozornosti, zhoršením rozlišitelnosti změn kmitočtu, hlasitosti a snížením citlivosti a tím změnou v maskování. Přitom nejlevnější špunty do uší, které jsou běžně k zakoupení v knihovnách, zajistí v uchu útlum okolo 25 dB a to v celém pásmu slyšení.

## 2 METODY HODNOCENÍ SROZUMITELNOSTI ŘEČI

Srozumitelnost řeči vyjadřuje schopnost zjistit význam a obsah toho, co mluvčí řekl. Srozumitelnost není subjektivní a lze ji snadněji hodnotit. Srozumitelnost je velmi důležitá a zásadní vlastnost řeči. Tuto vlastnost řeči je žádoucí za každou cenu zachovat, popř. zlepšit. Pro hodnocení srozumitelnosti řeči byla vyvinuta řada metod a testů. Tato kapitola vychází z [2],[3],

### 2.1 Subjektivní metody

Jsou založené na poslechových testech otologicky normálních osob. Tyto testy by měli splňovat několik požadavků:

- *Dobrá reprezentace všech hlavních fonémů řeči.*

Tento požadavek vyjadřuje potřebu, aby v každém testu srozumitelnosti byly obsaženy všechny fonémy. V ideálním případě by frekvence výskytu fonémů v testu měla vyjadřovat reálnou frekvenci výskytu fonémů v běžné řeči.

- *Stejná složitost a náročnost testů.*

Z tohoto důvodu je nezbytné mít relativně velkou sadu testů obsahujících slova, popř. věty stejné složitosti.

- *Řízení informace vyplývající z kontextu věty.*

Slova obsažená ve větách mají obecně daleko vyšší srozumitelnost než slova izolovaná. Toto je důvodem, proč mají být zkušební věty sestaveny tak, aby neumožňovaly snadným způsobem domýšlet slova a aby z tohoto pohledu byly jednotlivé věty a testy vzájemně srovnatelné.

#### 2.1.1 DRT

Jedná se o test srozumitelnosti, který je používán pro hodnocení řečových kodérů. Vstupem do testu je celkem šest skupin dvojic slov. Všechny skupiny obsahují sadu několika dvojic rýmujících se slov. Jednotlivé skupiny se liší svými významnými vlastnostmi (tj. rozdílným spektrálním charakterem slov ve skupině). Rýmující se slova jsou volena tak, aby se lišila počáteční souhláskou a rovněž jednou význačnou vlastností hlavního fonému slova. Úkolem posluchače je určit ze seznamu dvojic slov, které slovo z dané dvojice bylo vysloveno. Jako příklad lze například pro angličtinu uvést tyto dvojice slov: meat – beat, knock – dock, atd.



### **2.1.2 SPIN(Speech Perception in Noise)**

Test SPIN používá k hodnocení věty složené z 5 až 8 slov. Používá se 8 skupin vět, každá skupina je tvořena 50 větami. Jednotlivé věty jsou prezentovány skupině posluchačů. Test je postaven tak, aby po prezentaci věty posluchač odpovídal jedním slovem. Úkolem pro posluchače může být například určit poslední slovo v každé větě. Opět je nezbytné zajistit, aby jednotlivé skupiny vět měly stejnou úroveň složitosti z hlediska hodnocení srozumitelnosti. V praxi je toho dosaženo tak, že polovina vět ze skupiny, tj. 25 vět z 50 obsahuje slova ve větě, která lze z věty snadno odvodit. Naproti tomu druhá polovina vět (25) je koncipována tak, aby nebylo možno konkrétní slovo z kontextu věty odvodit. Výsledkem testu je potom procento správně rozpoznaných slov ve větě.

### **2.1.3 HINT(Hearing in Noise Test)**

Test HINT obsahuje 25 foneticky vyvážených skupin vět, každá skupina je složena z 10 vět. Všechny věty musí být shodné z hlediska jejich délky a srozumitelnosti. Na rozdíl od testu SPIN, kde se hodnotí správné rozpoznání jednoho slova v každé větě, se u testu HINT vyhodnocují všechna slova ve větě. Výsledkem testu HINT je potom průměrná procentuální hodnota správně rozpoznaných slov.

### **2.1.4 SRT**

SRT může být používán jako alternativa procentuálního vyjádření srozumitelnosti. SRT lze měřit buď v tichém prostředí, nebo v hluku. Tomu potom ale odpovídají různé definice SRT. Jestliže je SRT měřeno v tichém prostředí, je pak definováno jako úroveň řeči, při které posluchači rozpoznávají slova s pravděpodobností 50 %. Toho může být dosaženo prezentováním řeči při různých úrovních hlasitosti, od nízké po vysokou hlasitost, a určením procenta správně porozuměných slov. Tato závislost je následně vynesena do grafu. Z takto vneseného grafu se určí úroveň hlasitosti řeči, při které bylo dosaženo srozumitelnosti 50 %. Při měření SRT v hluku prostředí, je SRT definováno jako poměr signálu k šumu, při kterém rozpoznávají posluchači jednotlivá slova s pravděpodobností 50 %. SRT získáme tak, že je řeč prezentována při různých úrovních SNR (Signal-noise ratio) a vynesením procent správně rozpoznaných slov. Z takto vneseného grafu lze určit, jaká hodnota SNR odpovídá 50 % srozumitelnosti.

### **2.1.5 Hodnocení kvality**

Tato metoda je založena na prezentaci zkušebního signálu skupině posluchačů, kteří následně hodnotí kvalitu tohoto zkušebního signálu pomocí pětibodové stupnice od

1 do 5. Stupeň 5 značí nejvyšší kvalitu, stupeň 1 pak značí nejhorší kvalitu. Výsledek poté udává vážený průměr těchto hodnot.

## 2.2 Objektivní metody

Nevýhodou při použití subjektivních metod je jejich časová náročnost a potřeba zapojit do poslechových testů zkušené posluchače. Toto je důvodem, proč se začaly používat kromě testů subjektivních rovněž testy objektivní. Při objektivním hodnocení řeči se řeč rozdělí do rámců v délce trvání 10-30 ms a následně se výpočtem určí zkreslení mezi vstupní a výstupní řečí. Celkové hodnocení zkreslení řeči je následně dáno průměrem zkreslení jednotlivých rámců. Obecně může být výpočet proveden v časové, nebo kmitočtové oblasti. Pro kritéria, která jsou založena na kmitočtové oblasti, se předpokládá, že zkreslení nebo rozdíly, které se určí v modulovém spektru, jsou v korelaci s kvalitou řeči. Objektivní metody hodnocení kvality hledají vztah mezi originálním a zpracovaným signálem. V následujícím textu jsou uvedeny nejdůležitější z objektivních metod hodnocení kvality řeči. Objektivní metody lze obecně rozdělit na dvě skupiny:

- Metody jednokanálové. Tyto metody jsou charakterizovány tím, že u nich je k dispozici pouze jeden, tj. degradovaný, nebo zvýrazněný signál.

- Metody dvoukanálové. U metod dvoukanálových je porovnáván signál originální se signálem, který byl zpracován testovaným systémem.[3]

### 2.2.1 SNR (Signal to Noise Ratio)

Odstup signálu od šumu (Signal to Noise Ratio) je velmi často používaná veličina v elektronice. Z hlediska nejčastější aplikace měření odstupů dvou akustických signálů vychází definice SNR z vlastností vnímání intenzity zvuku lidským sluchovým ústrojím, které není lineární, ale logaritmické. Je definován jako poměr výkonu užitečného signálu, čímž se rozumí řečový signál, a výkonu neužitečného signálu, tedy hluku. Čím větší odstup mezi těmito signály bude, tím bude výsledný signál kvalitnější, bude obsahovat méně šumu. Jednotkou poměru signálu od šumu je decibel (dB). Výhodnou této metody je, že je výpočetně jednoduchá. Metoda SNR není příliš vhodná pro určení kvality řečového signálu. Výsledky se příliš neshodují s výsledky dosaženými subjektivním měřením. Hlavním důvodem je, že řečový signál je nestacionární. [14]

### 2.2.2 STI

Dalším způsobem měření je tzv. Index přenosu řeči (STI). Jde o měření kvality přenosu hlasu od řečníka směrem k posluchači. Řídí se normou IEC 60268-16. Užívá se k měření akustických vlastností sálů, učeben a místností, využívaných především akusticky. Když je přenos řeči perfektní, tak  $STI = 1$ . V učebně o normální velikosti by tato hodnota neměla překročit 0,75. Při nedokonalém přenosu řeči dochází ke zhoršenému vnímání hlasových variací posluchačem. Faktory, které negativně ovlivňují přenos řeči a podílejí se na snížení faktoru STI, mohou být například hluk pozadí, dlouhá doba dozvuku nebo ozvěny.

## 3 PRAKTICKÁ ČÁST

Pro prostředí nahrávání byli zvoleny dopravní prostředky hl. města Prahy, jakožto města s jednou z celosvětově nejhustších sítí městské hromadné dopravy, a to tramvaj, metro a autobus.

### 3.1 Nahrávání prostředí

Nahráváno bylo stereofonní technikou AB. Při procesu kmitočtové a hlukové analýzy se pak využil jen jeden kanál. Jako zdroj referenčního zvuku byl použit kalibrátor Brüel & Kjær typ 4231.



Obr. 3.1: Nahrávací řetězec

#### 3.1.1 Tramvaj

U tramvaje je mnoho faktorů, které ovlivňují hluk. Většinu z nich řadíme do kategorie nestacionárních zdrojů hluku např. výhybky, stav kolejí, lidé. Po konzultaci s řidičem tramvaje jsem ze 6 základních typů souprav vybral ke změření dva typy souprav, kterých je v provozu největší počet. Starší modernizovaná Tatra T3R.P a novější Škoda 15T, tyto modely tvoří více než 3/4 pražského vozového parku (817 souprav z toho 247 x 15T, 387 T3R.P). Jako pozici jsem určil místo v zadní části vozu naproti dveřím. Nahrávalo se v úseku se standartními žlábkovými tramvajovými kolejemi a u tramvaje 15T i v úseku s kolejemi bezžlábkovými.

#### 3.1.2 Metro

V Praze v současnosti jezdí 2 typy souprav. Novější M1 české výroby, jenž se zatím používá pouze na lince C. A starší 81-71M, což je modernizovaná souprava typu 81-71 sovětské výroby, která pokrývá provoz linek A a B. Z důvodu většího využití jsem zvolil k proměření modernizovanou soupravu 81-71M.

### 3.1.3 Autobus

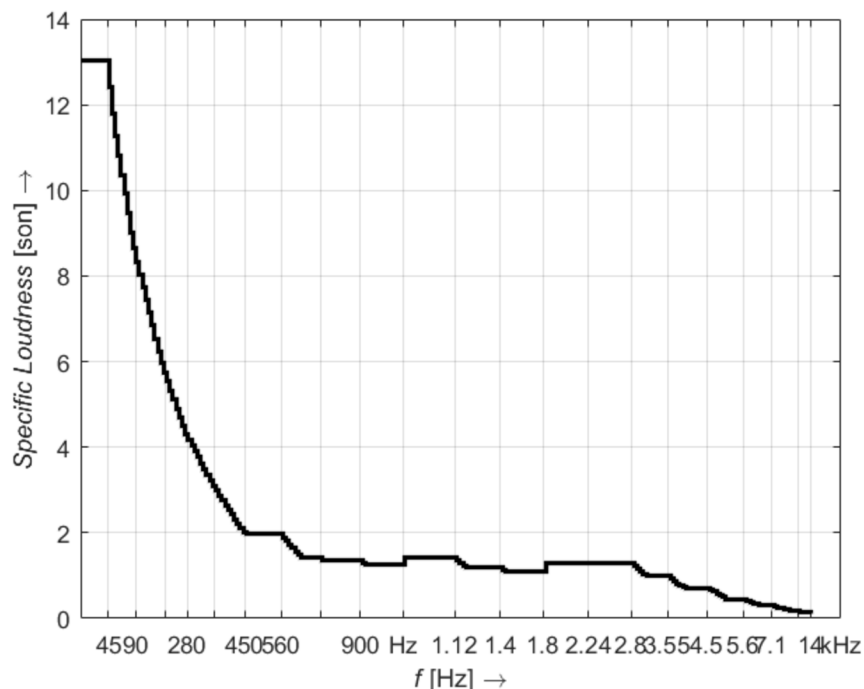
Autobusů jezdí celá řada. Jen v Praze je v provozu 46 různých typů autobusů 9-ti různých značek. Největší vozové zastoupení má český výrobce SOR. Proto jsem k měření vybral 2 typy. Kratší BN12 a dlouhý kloubový NB 15.

## 3.2 Výsledky měření prostředí

Pro vyhodnocení nahrávek jsem v prostředí MATLAB verze R2017a využil již napsaný program [12], který jsem upravil pro mou potřebu. Program načte zvuk ze souboru wav, kalibruje data podle kalibrační nahrávky, filtruje do 1/3 oktávových pásem a následně přepočítá na kritická pásma a vypočítá hlasitost. Všechny zdrojové kódy jsou na přiloženém CD a dají se spustit ve vývojovém prostředí MATLAB.

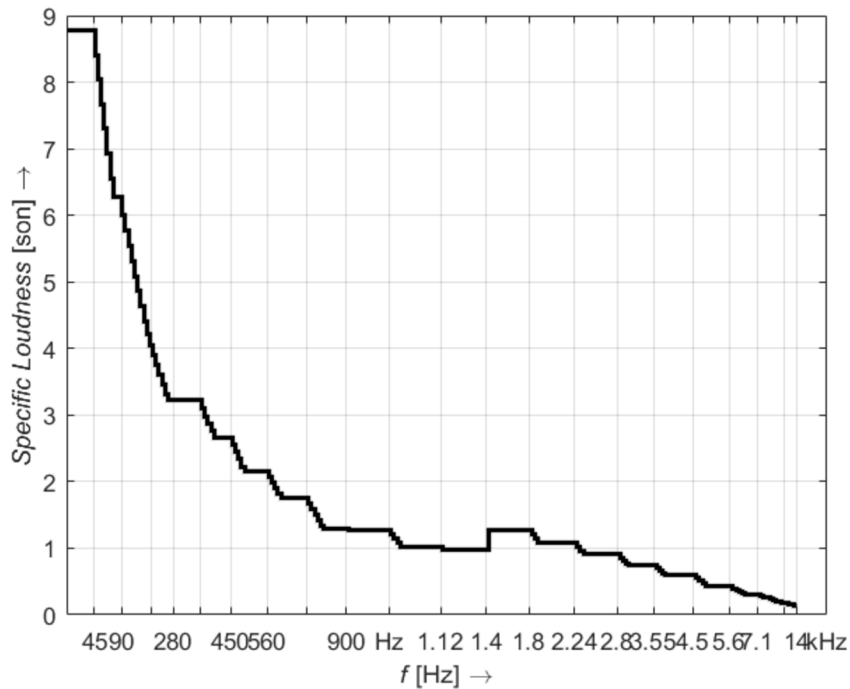
### 3.2.1 Tramvaj

Nejvíce u tramvaje ovlivňuje hlasitost typ použitých kolejí a jejich stav v obloucích. U stejného vozu v úseku, kde jsou položeny bezžlábkové koleje (jako u vlaků) klesla hlasitost ze 64 na 47 sonů. Dále také technický stav souprav (vybržděné plošky na kolech, povolené kryty kol,...).



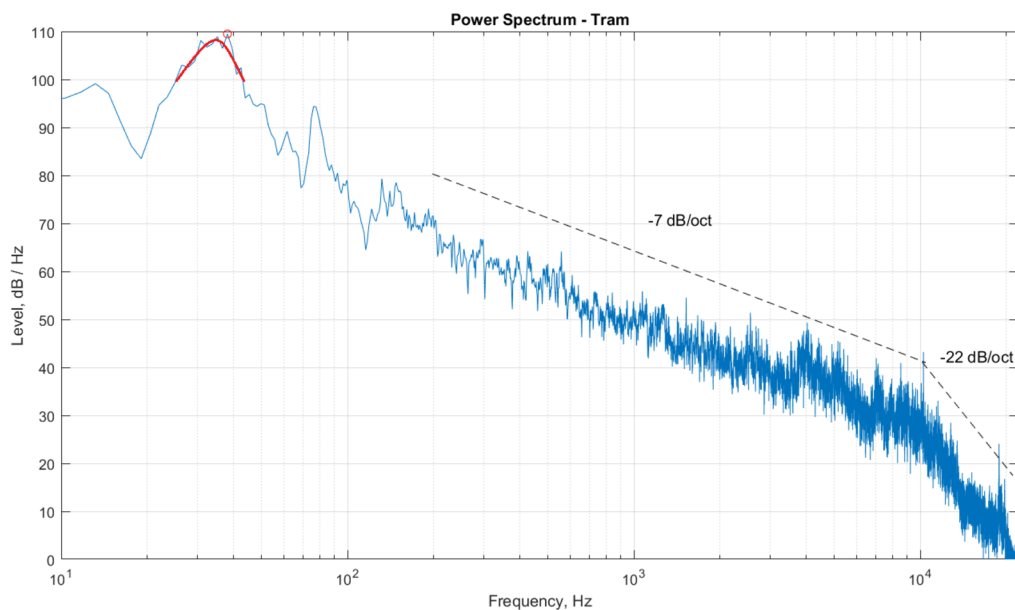
Obr. 3.2: Tramvaj T3R.P

Obr.3.2 je graf specifických hlasitostí a zároveň symbolizuje práh maskování v jednotlivých kritických pásmech, osa x má běžně jednotku bark, pro přehlednost je na obrázku uveden kmitočet, jednotlivá pásma jsou viditelná díky schodovitému charakteru grafu. U tramvaje je největší energie v prvních 3 kritických pásmech, tedy 0 - 300 Hz. Pásmo řeči se tak maskuje jen z malé části a v nižší dynamice.



Obr. 3.3: Tramvaj 15T na úseku po bezžlábkových kolejích

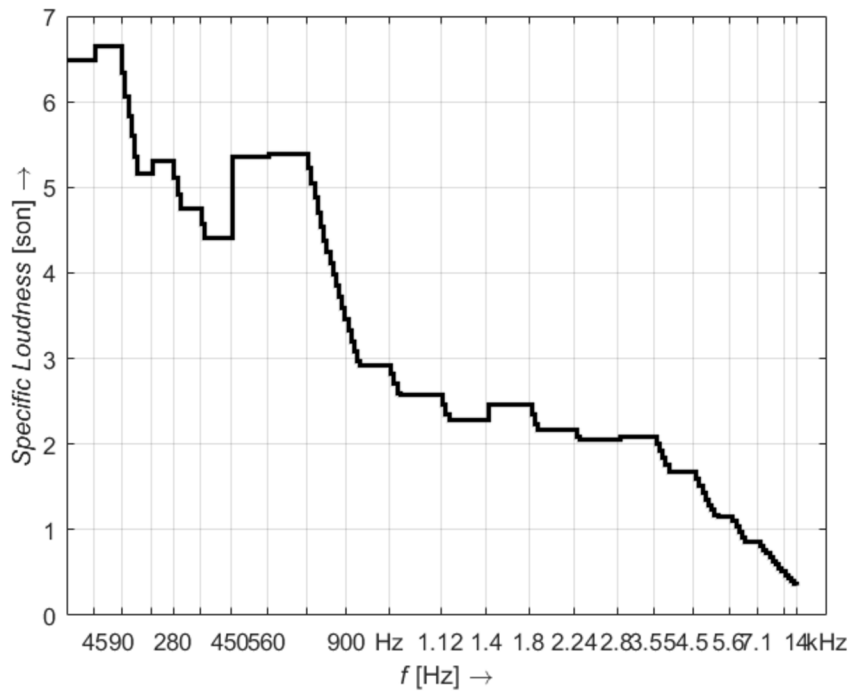
Maximální hlasitost 78 sonů byla naměřena, když souprava Škoda 15T (souprava s nejvyšší hmotností) přejížděla přes výhybky. Na obr.3.4 je vidět výrazně největší energie v červěně zvýrazněné oblasti 25 - 40 Hz. Při normální jízdě byla hlasitost v průměru 64 sonů po žlábkových a 47 sonů po bezžlábkových kolejích, které jsou navíc v mezikolejí zatravněné, hluk se tedy odráží v menší míře než od betonu nebo dláždění.



Obr. 3.4: Zvukové energetické spektrum prostředí tramvaje při jízdě přes výhybky

### 3.2.2 Metro

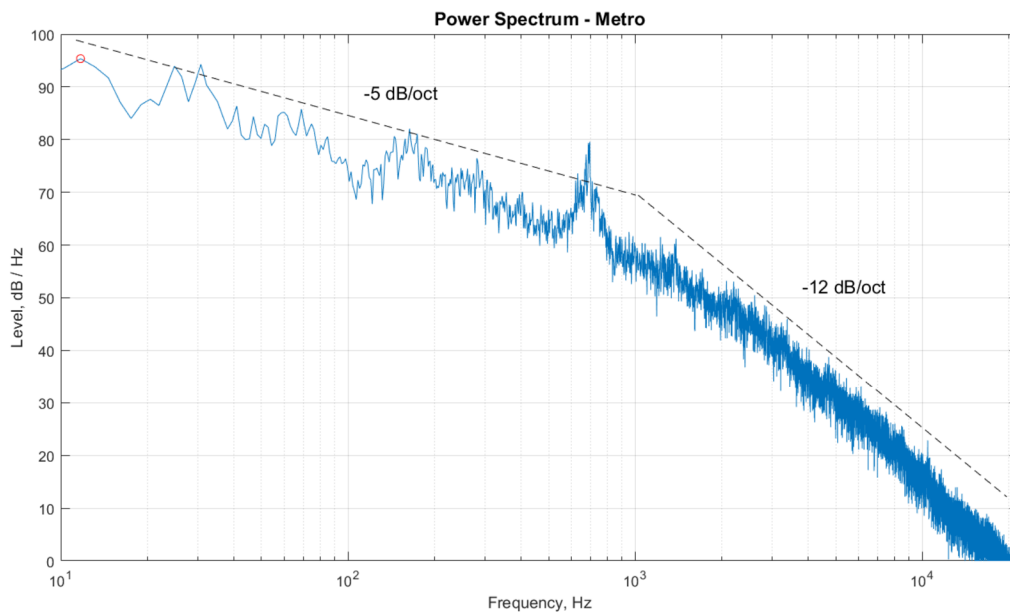
V metru je hlasitost ovlivněna především charakterem okolí, což je tunel. Hluk motoru při vysokých otáčkách se odráží od stěn tunelu a z velké části prochází dovnitř do vozu. Naměřil jsem hlasitost v průměru 79 sonů, což odpovídá hladině hlasitosti cca 103 fónů, v maximum dosáhla hlasitost 96 sonů.



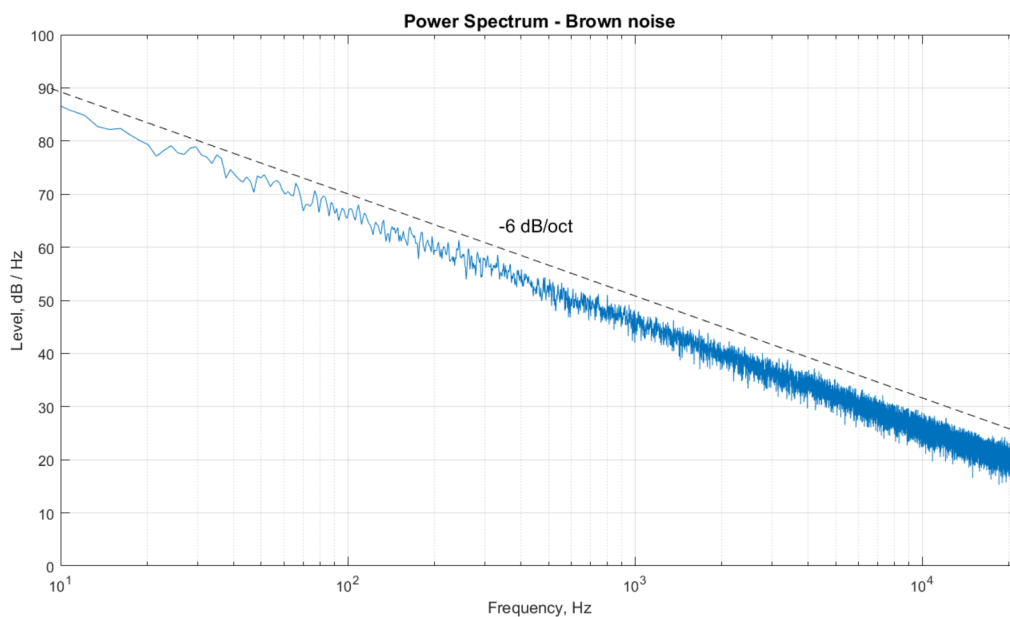
Obr. 3.5: Metro typ 81-71M

Na grafu na obr.3.6 lze pozorovat, že v pásmu do 1 kHz se metro podobá svým energetickým spektrem hnědému šumu. Jak je vidět na obr.3.7 hnědý šum vykazuje velké množství energie v nejnižších frekvencích, s rostoucím kmitočtem pak klesá se sklonem 6 dB/okt. V pásmu do 1 kHz u metra energie klesá 5 dB/okt, cca okolo 1 kHz nastává zlom a sklon je 12 dB/okt. Maskování nastává v celém pásmu řeči. Což způsobuje úplné zamaskování řeči v nezvýšené dynamice.





Obr. 3.6: Zvukové energetické spektrum uvnitř metra

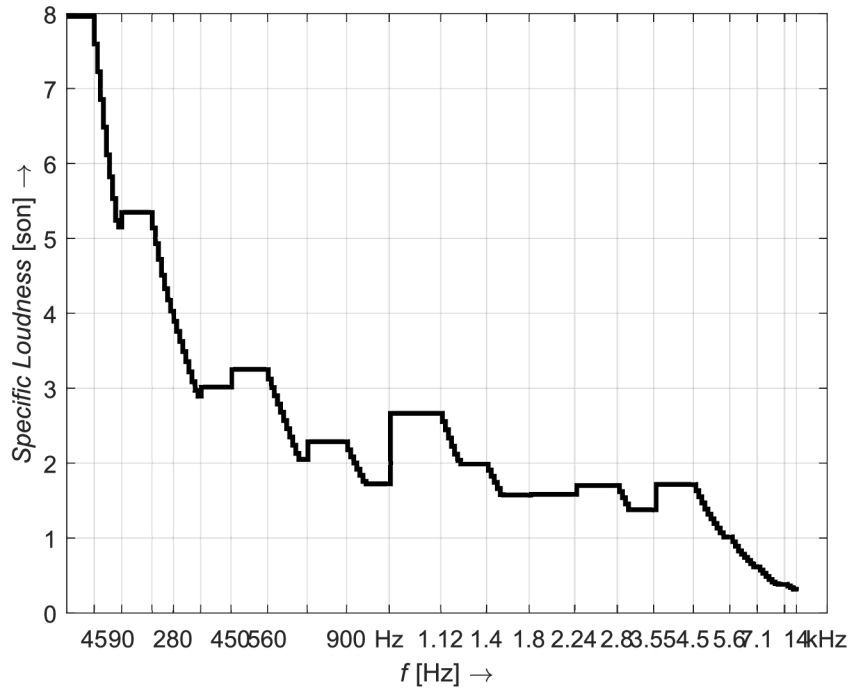


Obr. 3.7: Energetické spektrum hnědého šumu

### 3.2.3 Autobus

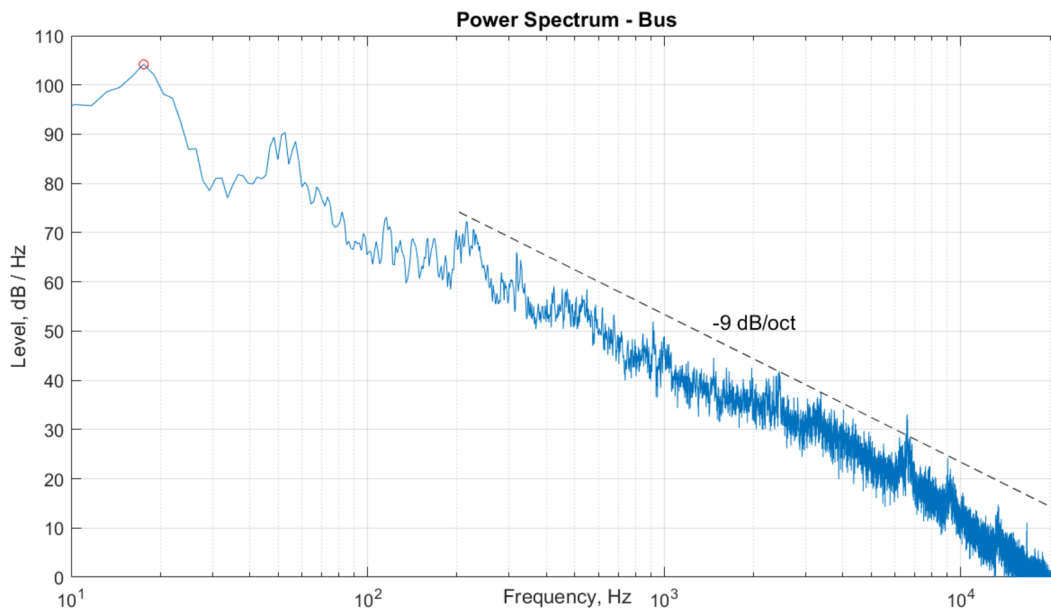
Hlasitost uvnitř autobusu je dána jeho technickým stavem, ale také výkonem motoru. Menší BN12 se používají v úsecích úžších silnic, ať už v centru města (jízda

přes dlažební kostky), nebo v kopcovitých oblastech (kde je třeba vysokých otáček). Větší NB15 se většinou používá pro vytížené linky, které jezdí na hlavních tazích. Nejvyšší hlasitost jsem naměřil 89 sonů, konkrétně to bylo při zavírání dveří. Jelikož jde o zvuk s impulzní charakteristikou, tak může při řeči maskovat nanejvíš dvě slova.



Obr. 3.8: Bus BN12

Při jízdě se hlasitost pohybovala v rozmezí 50 - 60 sonů. Podobně jako u tramvaje, největší množství energie vykazují první dvě kritická pásma. Na obr.3.9 je vidět maxima okolo 18 Hz a 50 Hz, to jsou kmitočtové oblasti motoru autobusu. Od 200 Hz s rostoucím kmitočtem energie klesá se sklonem 9 dB/okt. V autobuse tedy nastává efekt maskování v pásmu řeči minimálně a jen při nižších dynamikách.



Obr. 3.9: Energetické spektrum autobusu

### 3.3 Souhrn

Z hlediska hlasitosti je metro dopravním prostředkem, ve kterém bychom si měli chránit sluch nejvíce. Maximální hlasitosti dosahují řádově stejných hodnot. Z grafů podle Zwickera i ze spektrálních obrazů jednotlivých prostředků jasně vyplývá, že efekt maskování by se měl projevovat nejvíce v metru, a to v celém pásmu řeči. Tramvaj a autobus jsou na tom lépe, k maskování by mělo docházet při běžné jízdě pouze v prvních třech kritických pásmech tzn. v pásmu od 0 do 300 Hz.

Maximální hlasitost		
	N[son]	$L_N$ [fón]
Tramvaj	78,07	102,85
Metro	96,79	105,85
Autobus	89,38	104,76

Obr. 3.10: Tabulka naměřených maximálních hodnot

Průměrné hlasitosti			
Dopravní prostředek	Model	N[son]	L <sub>N</sub> [fón]
Tramvaj	T3R.P Tatra	59,11	98,82
Tramvaj	15T Škoda	64,04	100
Tramvaj	15T Škoda (bezžlábkové koleje)	47,5	95,55
Metro	81-71M	78,85	102,85
Autobus	SOR BN12	61,84	99,31
Autobus	SOR NB15	50,59	96,44

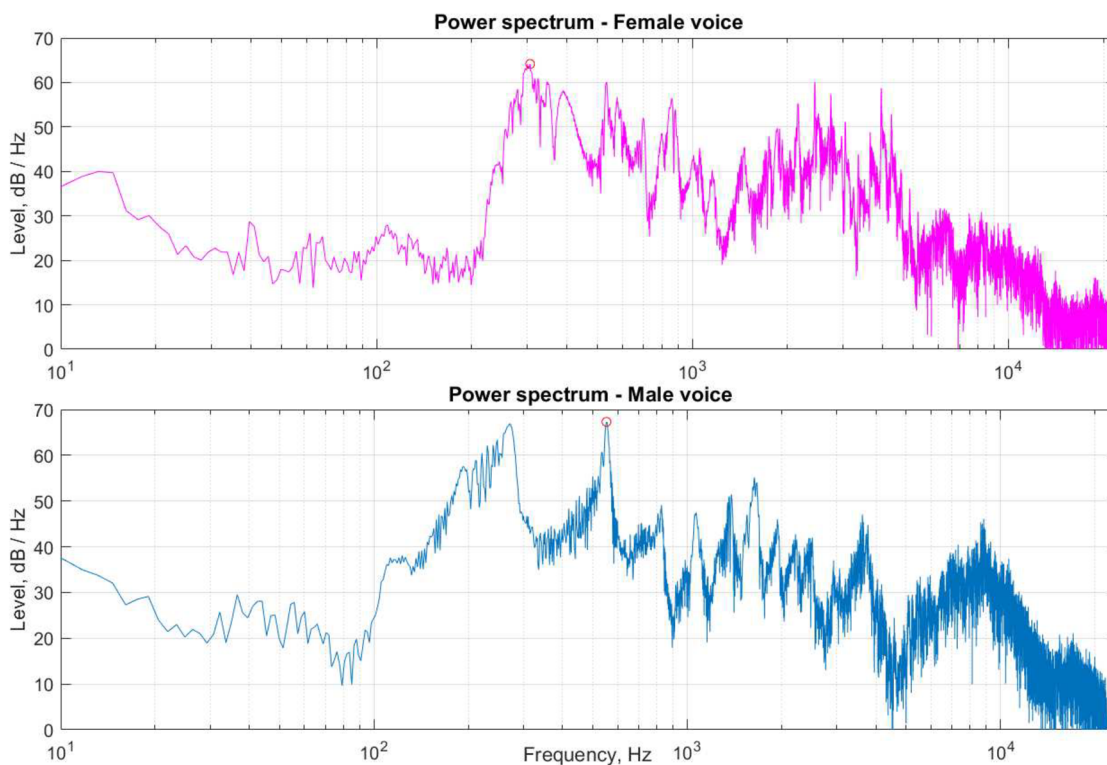
Obr. 3.11: Tabulka naměřených průměrných hodnot

### 3.4 Nahrávání řeči

Probíhalo v bezdrazové komoře na VUT. Přičemž se použil stejný nahrávací řetězec, jako při nahrávání hlukového pozadí, viz obr.3.1. Mikrofon byl vzdálen od hlavy mluvčího na délku 60 cm (délka paže). Tuto vzdálenost jsem nastavil z kulturního důvodu, pokud to není nutné, v naší kultuře nezasahujeme do intimní zóny jedince. Jako vzorek byli vybráni 4 lidé, 3 muži 1 žena.

Testovací věty byly sestaveny tak, aby dohromady zachytily všechny fonémy řeči. Kompletní test jedné metody hodnocení srozumitelnosti by vyžadoval nahrání celé baterie testů, což by svým rozsahem vydalo na samostatnou práci. Jelikož jde spíše o předtest, rozhodl jsem se pro kombinaci metod SPIN a HINT. Věty byly postavené podle metody SPIN, výstup je pak, jako u HINTu procento rozpoznávaných slov ve větě. Nahrálo se postupně 8 vět, z nichž bylo použito 7. Mluvčí byli požádáni aby mluvili přirozeně ve čtyřech dynamikách:

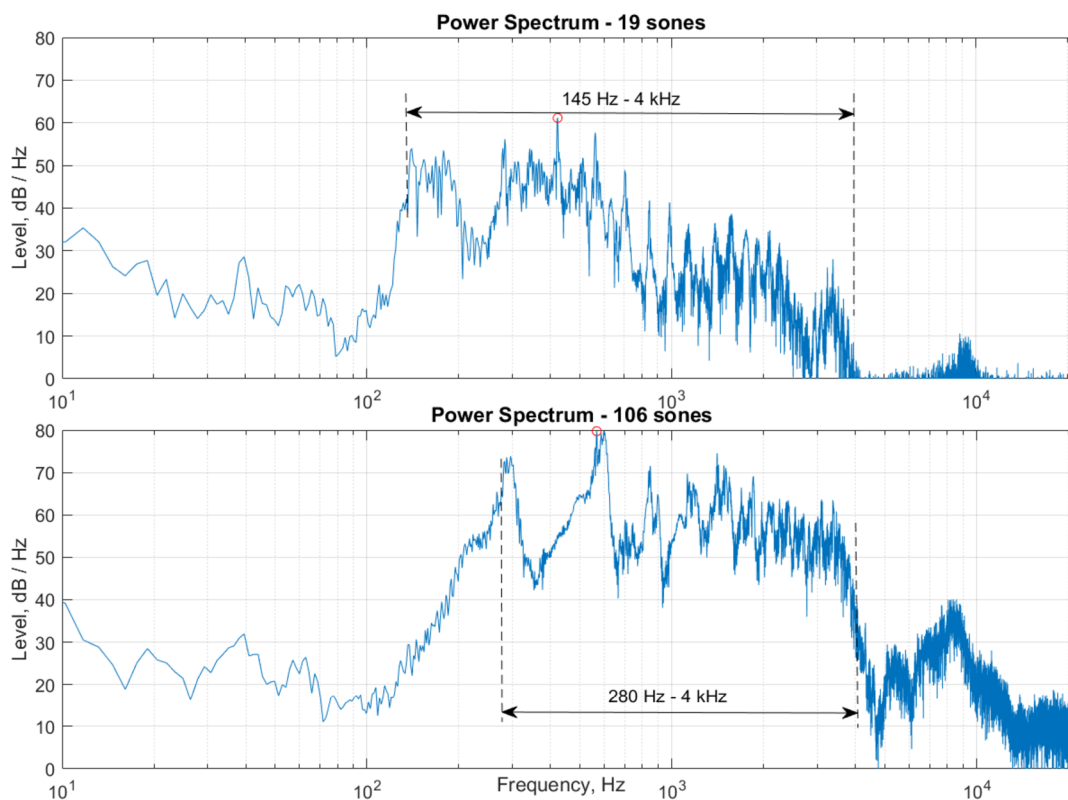
1. dynamika představuje normální mluvu bez jakýchkoli rušivých podnětů.
2. dynamika je mluva zesílená, která odpovídá mluvě v normálním prostředí.
3. dynamika je mluva hlasitá, která je vynucena hlukem prostředí.
4. dynamika je maximum, kterého jsou mluvčí schopni, ať už z kulturního nebo fyzického hlediska.



Obr. 3.12: Spektrum žena (nahore) a muž

Na obr.3.12 vidíme spektrální srovnání mužského a ženského hlasu při stejné rozmluvě a stejné hlasitosti. Spektra jsou si podobná, mužský hlas začíná o několik desítek Hz níže a koncentruje nejvíce energie v pásmech okolo 190 - 275 Hz a 520 - 580 Hz. 1. pásmo ženské řeči, které zároveň obsahuje nejvíce energie leží mezi hodnotami 250 a 400 Hz. Ženský hlas má oproti mužskému také výrazné pásmo 2,2 kHz - 4 kHz, které může být klíčové, z hlediska srozumitelnosti, protože v tomto pásmu je lidský sluch nejcitlivější.

Zajímavý je fakt, že při některých promluvách, bylo pro některé mluvčí přirozenější místo zesílení, změnit barvu hlasu a lehce prodlužovat samohlásky. To se může projevit vyšší srozumitelností v prostředích s vyšším maskovacím prahem v prvních kritických pásmech, jako je třeba autobus nebo tramvaj. Na obr.3.13 můžeme vidět srovnání spekter nejtíší a nejhlasitější dynamiky jednoho mluvčího, při čtení stejné věty. Vidíme, že hlasitost je více než 5 krát větší a začátek pásma mluvy se posouvá výše ve spektru ze 145 Hz na 280 Hz.



Obr. 3.13: Srovnání spekter při rozdílných dynamikách

Databáze nahrávek řeči i prostředí a také tabulka s naměřenými hlasitostmi nahraných promluv jsou na příloženém CD. Nahrávky prostředí se dají využít jako ruchy při filmové postprodukci.

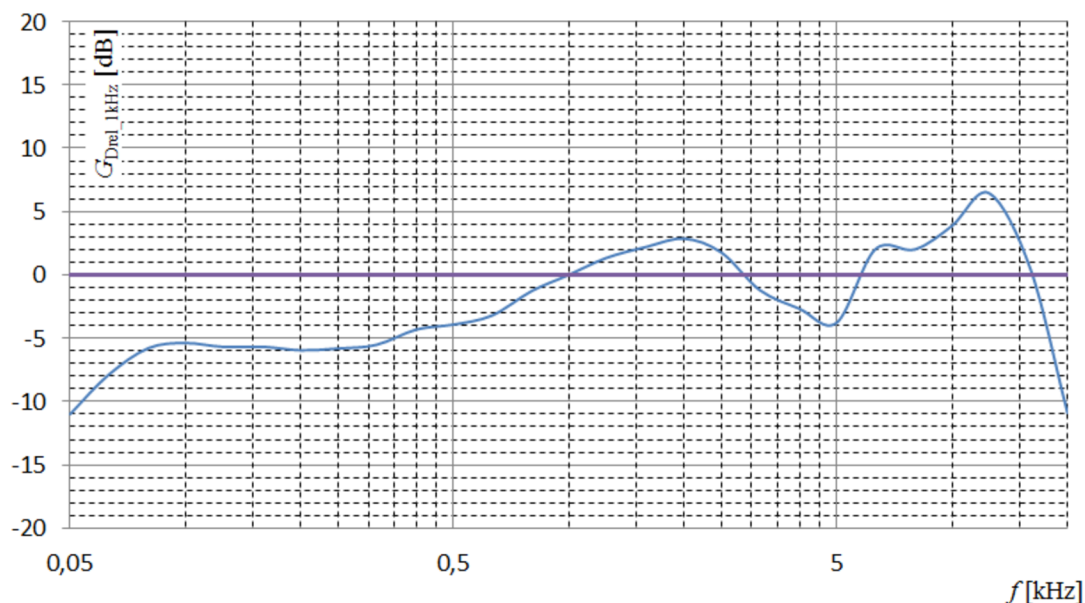
## 3.5 Poslechové testy

Abych určil vzorky, které použiji v testování, změřil jsem hlasitost všech promluv. Poté jsem pro každou dynamiku spočítal průměr hlasitosti všech vět. Vybrané pak byly 4 promluvy, které se nejvíce blížily svou hlasitostí průměru dané dynamiky. Každá promluva měla jiného mluvčího a jiný obsah. Po smíchání s nahrávkami prostředí 3 dopravních prostředků tak vzniklo 12 testovacích nahrávek ve formátu wav.

### 3.5.1 Měření sluchátek

Poslechový test by měl být proveden pokud možno za stejných poslechových podmínek. Použita byla sluchátka, u kterých jsem nejprve proměřil kmitočtovou charakteristiku. Tato sluchátka pak byla použita při všech testech. Měření bylo provedeno

v laboratoři VUT s použitím simulátoru umělé hlavy a torza. Použil jsem vlastní circum-aurální sluchátka Superlux HD 681B s polootevřenou konstrukcí.



Obr. 3.14: Kmitočtová odezva sluchátek k poslechovému testu

### 3.5.2 Výsledky testů

Samotný test se sestával ze 12 částí, ve kterých byla dotazovaným pouštěna testovací nahrávka. Respondenti neměli možnost ovlivnit hlasitost, ta byla pevně dána pro všechny testy. Úkolem respondentů bylo, napsat všechna slova, která slyšeli. Mezi nahrávkami byla vždy pauza 10 sekund. Výsledek testu pak byla průměrná procentuální hodnota správně rozpoznaných slov, jako u metody HINT. Vyhodnocení proběhlo pro všechny části testu.

	Procento rozpoznaných slov			
	1	2	3	4
Dynamika				
Autobus	11,7%	93,3%	75,0%	83,3%
Tramvaj	11,7%	15,4%	20,8%	96,7%
Metro	0,0%	0,0%	6,3%	23,9%

Obr. 3.15: Tabulka průměrů rozpoznaných slov

Jak vyplývá z tabulky na obr.3.15. Efekt maskování se nejvíce projevuje v prostředí metra, kde je řeč úplně maskována v nižších dynamikách a při vyšších dynamikách je maskována do té míry, že obsah sdělení je nesrozumitelný. V tramvaji

maskování zapříčinilo v nižších dynamikách nesrozumitelnost, když respondenti byli schopni správně určit většinou jedno slovo. V nejvyšší dynamice pak byla srozumitelnost nejlepší. Autobus vyšel ze všech testů nejlépe, když pouze na nejnižší dynamice byla srozumitelnost nízká. Ve všech dalších dynamikách byla úspěšnost nad 75 procent.

Ve výsledcích autobusu se vyskytla zajímavost, kdy nahrávka v druhé dynamice měla nejvyšší úspěšnost rozpoznání slov. To může být zapříčiněno tím, že nahrávka obsahovala ženský hlas, který je posazen kmitočtově výš a je tak větší pásmo mezi hlasem a hlukem. Je také možné, že obsah sdělení, což byla věta "Dávejte pozor na levou stranu", je svým obsahem alarmující a z psychologického hlediska bychom jí právě v dopravním prostředí čekali.



## 4 ZÁVĚR

Úspěšně byla nahrána prostředí tří dopravních prostředků: tramvaje, metra a autobusu. Na základě těchto nahrávek se pak podařilo určit hlasitost a maskovací práh u všech prostředků, z nichž nejvýše posazený z hlediska řeči je u metra. Hluk v metru se v pásmu do 1 kHz podobá hnědému šumu a je nejhlasitější. Na rozdíl od toho v autobusu i v tramvaji je nejhlasitější oblast do 150 Hz, což ovlivňuje pásmo řeči jen minimálně. Maximální hlasitosti jsou řádově stejné, v průměru je výrazně nejhlasitější metro. U autobusu i tramvaje je hlasitost zhruba na stejné úrovni. Obecně lze říci, že rychlost dopravního prostředku je vykoupena hlukem jenž prostředek způsobuje.

Úspěšně se podařilo nahrát a smíchat s nahrávkami prostředí dopravních prostředků testovací promluvy navrhnuté na základě testovací metody SPIN. Testy pak byly formou dotazníku provedeny na vzorku otologicky normálních osob. Pro snazší vyhodnocení se pak použil systém vyhodnocení procentuálního rozpoznání všech slov ve větě. Z výsledků poslechových testů vyplývá, že v tramvaji při nižších dynamikách řeči dochází ke snížení srozumitelnosti, ale nedochází k úplnému maskování. V autobusu pak k maskování docházelo pouze v nejnižší dynamice řeči, zatímco při normální a hlasitější mluvě byli respondenti bezproblémů schopni určit obsah sdělení. V metru dochází k úplnému maskování pásma řeči v nižší a střední dynamice. A ani v nejvyšší dynamice nebyli schopni respondenti určit obsah sdělení s úspěšností alespoň 50 procent.

Přesnější měření v oblasti hluku v dopravních prostředcích, naráží na skutečnost, že podmínky při měření jsou nestálé. Metro je v tomto ohledu asi nejlépe změřitelné, protože není ovlivněné atmosferickými podmínkami a v ČR jsou v provozu pouze dva typy souprav. Naproti tomu autobusy a tramvaje jsou prostředky, u kterých je každé měření originální, ať už jde o typ, technický stav vozu, vozovky nebo u tramvaje kolejí. Technický stav a stav povrchu nejvíce ovlivňuje hluk uvnitř tramvajů a autobusů. Jelikož se efekt maskování a obecně vyšší hlasitost hluku projevuje nejvíce v metru, další měření a možnou studii z hlediska potlačení hluku a maskování bych doporučil v prostředí metra. Tato studie by měla zahrnout oba typy metra, změření impulsní odezvy uvnitř vagonu, pro pozdější vytvoření dozvuku pomocí konvoluce. Také by bylo třeba určit zdroje hluku a možné prostředky k jeho snížení, což by zahrnovalo měření průzvučnosti samotného vagonu a odrazivosti stěn v tunelu. Pro relevantnější výsledky poslechových testů by bylo vhodné počítat s adaptací na prostředí, která vede k selektivní pozornosti a tím k rozdílným výsledkům.

# LITERATURA

- [1] MELKA, A. *Základy experimentální psychoakustiky*. Praha: Akademie múzických umění, 2005. ISBN 80-7331-043-0
- [2] LOIZOU, Philipos C. *Speech enhancement: theory and practice*. CRC Press, Boca Raton, Canada, 2007, ISBN 0-8493-5032-8.
- [3] HOVORKA, Jaroslav. *Kombinované vícepásmové adaptivní zvýraznění řeči*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2015.
- [4] KADLEC, František. *Zpracování akustických signálů*. ČVUT v Praze. Fakulta elektrotechnická, 2005, ISBN 80-01-02588-8.
- [5] SIGMUND, Milan. *Rozpoznání řečových signálů*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2007, ISBN 978-80-214-3526-1.
- [6] SCHIMMEL, Jiří. *Elektroakustika*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2016. ISBN 978-80-214-4716-5.
- [7] ČSN 01 1600: Akustika - Terminologie. Česká technická norma, Český normalizační institut, únor 2003.
- [8] ČSN ISO 532-1: Metody výpočtu hlasitosti – Zwickerova metoda, Česká agentura pro standardizaci, červenec 2018.
- [9] KULDANOVÁ, Pavlína. *Zvuková stránka češtiny*. Ostravská univerzita v Ostravě, září 2013. ISBN 978-80-7464-469-6.
- [10] LEPA, Ondřej. *Psychoakustický model lidského sluchu*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2011.
- [11] KOSTELNÍK, Pavel. *Index srozumitelnosti řeči*. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2013.
- [12] HASTINGS, Aaron. *ISO 532 B / DIN 45631 Loudness*. School of Mechanical Engineering, Purdue University, 2000.
- [13] A-weighting. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001. Dostupné z: <http://en.wikipedia.org/wiki/A-weighting>.

- [14] BLAHA, Petr. *Hodnocení kvality řečových signálů na bázi virtuální instrumentace*. VŠB – Technická univerzita Ostrava. Fakulta elektrotechniky a informatiky, 2017.

# A OBSAH PŘILOŽENÉHO CD

Všechny nahrávky jsou ve formátu wav se vzorkovací frekvencí 48 kHz.

/	.....	kořenový adresář přiloženého CD	
	LaTeX	..... balíčky pro sazbu práce v LaTeX a zdrojové soubory práce	
		loga	..... loga školy a fakulty
		obrazky	..... ostatní obrázky
		pdf	..... pdf stránky generované informačním systémem
		text	..... zdrojové textové soubory
	Matlab	..... zdrojové soubory v prostředí Matlab	
	Nahrávky DP	..... databáze nahrávek prostředí dopravních prostředků	
	Nahrávky Hlas	..... databáze nahrávek řeči	
	Tabulky	..... tabulky s výsledky	
	Testy	..... nahrávky k poslechovým testům	