

Objektívne meranie zrozumiteľnosti reči metódou STIPA

Aktuálny štandard STN EN 60268-16 (rev. 4 : 2011) presne definuje a upresňuje metodológiu na objektívne meranie zrozumiteľnosti reči pomocou Indexu prenosu reči STI (Speech Transmission Index). STI vyjadruje fyzikálne vlastnosti prenosového kanála (akustické vlastnosti miestnosti a elektroakustické vlastnosti zariadenia) a vyjadruje schopnosť kanála prenášať fyzikálne parametre rečového signálu. STI je osvedčený objektívny ukazovateľ na zistenie, ako vlastnosti prenosového kanála ovplyvnia zrozumiteľnosť reči.

Vplyv, ktorý má prenosový kanál na zrozumiteľnosť reči, závisí od:

- Úrovne rečového signálu
- Frekvenčnej prenosovej charakteristiky kanála
- Nelineárnych skreslení
- Odstupu signál/šum
- Kvality zvukového reprodukčného zariadenia
- Ozvien (odrazov s oneskorením > 100 ms)
- Doby dozvuku miestnosti
- Psychoakustické vlastnosti sluchu (maskovanie signálu)

STI bola vynájdená a publikovaná T. Houtgastom a H. Steenekenom prvý raz v roku 1971 v časopise *Acustica* a neskôr v roku 1980 v časopise *J. Acoustical Society of America*. Houtgast a Steeneken, ako výskumní pracovníci holandskej organizácie aplikovaného výskumu TNO, vynášali objektívnu metódu merania STI ako náhradu za dlhotrvajúce a únavné subjektívne merania zrozumiteľnosti. Ich výskumný tím v TNO ďalej vyvíjal a zdokonaľoval metódu merania, hardvér a softvér na meranie STI až do roku 2010. V tomto roku výskumná skupina TNO, zaoberajúca sa meraním STI, opustila TNO a pokračuje v práci ako súkromná firma s názvom *Embedded Acoustics*. H. Steeneken pôsobí v súčasnosti v tejto spoločnosti ako senior konzultant.

Okolo roku 2000 pracovníci TNO (J. Verhave a H. Steeneken) začali výskumné práce na novej STI metóde, ktorá sa neskôr stala známa ako **STIPA** (**STI** pre **P**ublic **A**dress systems).

V r. 2003 bola metóda STIPA štandardizovaná v norme EN 60 268-16 (rev. 3) ako metóda na rýchle meranie zrozumiteľnosti reči v systémoch verejného ozvučenia (PA) a systémoch evakuačného rozhlasu (systémoch hlasovej signalizácie požiaru HSP).

Opis STI metódy

STI je číselné vyjadrenie vlastností komunikačného kanála a nadobúda hodnoty od 0 = zlá kvalita do 1 = vynikajúca kvalita. Pre väčšinu aplikácií je vyžadovaná hodnota STI najmenej 0,5.

Barnett (1995, 1999) navrhol použitie referenčnej stupnice všeobecnej zrozumiteľnosti reči **CIS** (**C**ommon **I**ntelligibility **S**cale), ktorá je v nasledovnom matematickom vzťahu k STI:

$$\text{CIS} = 1 + \log(\text{STI})$$



Zrozumiteľnosť **ZLÁ** **SLABÁ USPOKOJIVÁ** **DOBRÁ** **VÝBORNÁ**

Zrozumiteľnosť reči sa dá vyjadriť jednoducho číselnou hodnotou. Najčastejšie používané stupnice sú STI a CIS.

Vzťah objektivnej metódy merania STI a subjektívnej metódy merania zrozumiteľnosti možno vyjadriť prostredníctvom pravdepodobnosti zrozumiteľnosti slabík, slov a viet. V nasledovnej tabuľke je uvedená táto pravdepodobnosť pre poslucháčov počúvajúcich rodný jazyk:

Hodnota STI	Zrozumiteľnosť podľa EN 60268-16	Zrozumiteľnosť slabík v %	Zrozumiteľnosť slov v %	Zrozumiteľnosť viet v %
0 - 0,3	zlá	0 - 34	0 - 67	0 - 89
0,3 – 0,45	slabá	34 - 48	67 - 78	89 - 92
0,45 – 0,6	uspokojivá	48 - 67	78 - 87	92 - 95
0,6 – 0,75	dobrá	67 - 90	87 - 94	95 - 96
0,75 - 1	výborná	90 - 96	94 - 96	96 - 100

V prípade poslucháčov počúvajúcich iný ako rodný jazyk, u nedoslýchavých ľudí alebo ľudí s inými poruchami sluchu sú pravdepodobnosti zrozumiteľnosti iné.

Je zaujímavé ale nie prekvapujúce, že hodnota STI nezávisí od hovoreného jazyka, pretože sa merajú fyzikálne vlastnosti prenosového kanála.

EN 60268-16 ver. 4 definuje kvalifikačné pásma STI (označené písmenami U až A+) pre typické aplikácie:

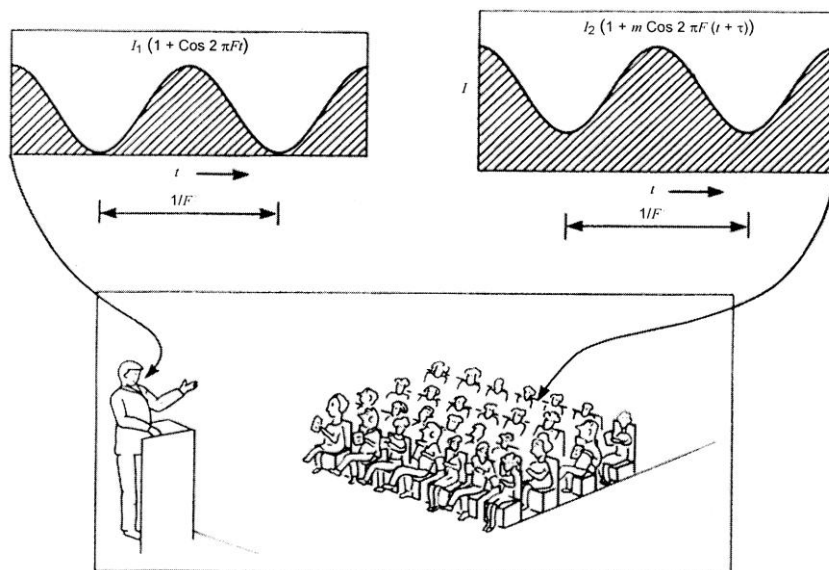
Kategória	Menovitá hodnota STI	Typ správy	Typický príklad použitia	Komentár
A+	> 0,76		Záznamové štúdio	Výborná zrozumiteľnosť, zriedka dosiahnuteľná vo väčšine prostredí
A	0,74	Komplexné správy, neznáme slová	Divadlá, posluchárne, parlamenty, súdy, systémy pre sluchovo postihnutých	Vysoká zrozumiteľnosť reči
B	0,7	Komplexné správy, neznáme slová	Divadlá, posluchárne, parlamenty, súdy, systémy pre sluchovo postihnutých	Vysoká zrozumiteľnosť reči
C	0,66	Komplexné správy, neznáme slová	Divadlá, posluchárne, telekonferencie, parlamenty, súdy	Vysoká zrozumiteľnosť reči
D	0,62	Komplexné správy, známe slová	Prednáškové sály, učebne, koncertné haly	Dobrá zrozumiteľnosť reči
E	0,58	Komplexné správy, známy kontext	Koncertné haly, moderné kostoly	Vysoko kvalitné PA ozvučovacie systémy
F	0,54	Komplexné správy, známy kontext	PA systémy v nákupných centrách, kancelárie, hlasové evakuačné systémy, katedrály	PA ozvučovacie systémy dobrej kvality
G	0,5	Komplexné správy, známy kontext	Nákupné centrá, kancelárie, hlasové evakuačné systémy,	Cieľová hodnota pre hlasové evakuačné systémy
H	0,46	Jednoduché správy, známe slová	Ozvučovanie PA a hlasové evakuačné systémy v obtiažnych akustických prostrediach	Za normálnych podmienok najnižšia hranica pre hlasové evakuačné systémy
I	0,42	Jednoduché správy, známy kontext	Ozvučovanie PA a hlasové evakuačné systémy vo veľmi obtiažnych miestach	
J	0,38		Nevhodné pre PA ozvučovanie systémy	
U	< 0,36		Nevhodné pre PA ozvučovanie systémy	

Pozn. 1: Hodnoty STI uvedené v tabuľke sú minimálne cieľové hodnoty.

Pozn. 2: Dosiagnutá zrozumiteľnosť závisí aj od frekvenčnej charakteristiky v každom mieste posluchu.

Pozn. 3: STI hodnota predstavuje nameranú hodnotu vo vybraných miestach posluchu alebo hodnotu požadovanú v špecifických aplikačných štandardoch.

Koncepcia STI metódy je založená na empirickom zistení, že rýchle zmeny časového priebehu amplitúdy rečového signálu nesú najdôležitejšie informácie o zrozumiteľnosti reči. Rýchle časové zmeny priebehu signálu sú spôsobené akustickým oddelením viet, slov a fonémov (základných prvkov reči). Rýchle časové zmeny amplitúdy signálu, tzv. modulácie, sa môžu vyjadriť ako funkcie modulačnej frekvencie F , ktorá vytvára modulačné spektrum. Modulačné frekvencie sa vyskytujú v reči typicky v rozsahu 0,5 Hz až 16 Hz s maximom okolo 3 Hz. Akékoľvek narušenie modulačného spektra spôsobeného prenosovým kanálom má za následok zhoršenie zrozumiteľnosti reči. Stupeň narušenia modulačného spektra korešponduje so zmenšením modulačnej hĺbky pri jednej alebo viacerých modulačných frekvenciách a vypočíta sa ako modulačný prenosový parameter pre každé frekvenčné oktávové pásmo celého prenášaného pásma reči. Obr. 1 zobrazuje koncepciu redukcie hĺbky modulácie, ktorá vznikne medzi rečníkom a poslucháčom.



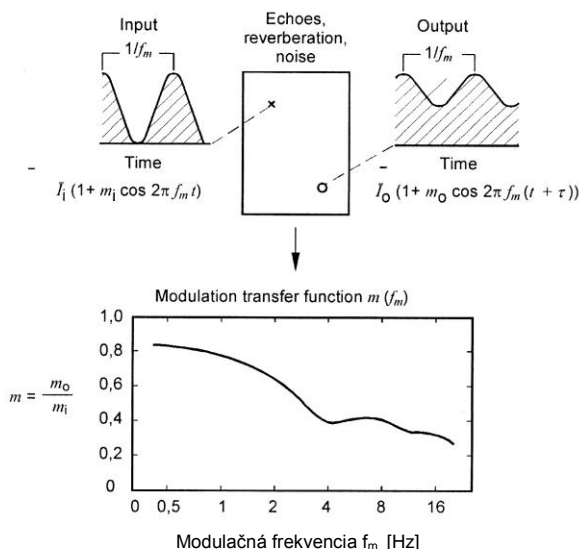
Obr. 1 Koncepcia redukcie hĺbky modulácie spôsobenej prenosovým kanálom

STI metóda bola optimalizovaná a overená subjektívnymi meraniami zrozumiteľnosti s pridaním rozmanitých rušivých vplyvov do prenosového kanála (šum, odrazy signálov, ozveny, nelineárne skreslenie).

STI testovací signál je umelý signál, odvodený z vlastností rečového signálu. Je tvorený šumovým signálom v 7 frekvenčných oktávových pásmach od 125 Hz do 8 kHz. Každý nosný signál v oktávovom pásme je modulovaný 14 modulačnými frekvenciami 0,63 Hz až 12,5 Hz s odstupom 1/3 oktávy.

Modulačný index m_i testovacieho signálu je prehrávaný v miestnosti alebo cez komunikačný kanál a prijímaný na mieste poslucháča ako modulačný index m_o . Pri meraní STI podľa obr. 1 by sa testovací signál vysielal zvukovým zdrojom (reproduktorom) simulujúcim živého rečníka (umiestneným na mieste rečníka) a signál by sa snímal na mieste poslucháča meracím mikrofónom.

Typický testovací signál pozostáva zo signálu nosnej frekvencie s frekvenčným spektrom vychádzajúcim zo spektra ľudskej reči, ktorého veľkosť je modulovaná sínusovým signálom s frekvenciou f_m (obr. 2).



Obr. 2 Modulačná prenosová funkcia – porovnanie vstupu a výstupu

Zmenšenie hĺbky modulácie pri frekvencii f_m sa kvantifikuje **modulačnou prenosovou funkciou $m(f_m)$** , ktorá sa určí ako $m(F) = m_o(F)/m_i(F)$ a interpretuje sa ako **efektívny pomer signál/šum SNR_{eff}** (bez ohľadu na príčinu zmenšenia hĺbky modulácie, ktorá môže byť spôsobená odrazmi, ozvenou, nelineárnym skreslením alebo rušivým hlukom):

$$SNR_{eff} = 10 \log [m(F)/1-m(F)]$$

Hodnoty efektívneho pomeru signál/šum sú ohraničené v rozsahu -15 dB až +15dB. Hodnotám menším ako -15dB sú priradené hodnoty -15dB a hodnotám väčším ako +15dB hodnoty +15dB.

Modulačná prenosová funkcia $m(F)$, niekedy označovaná ako **MTF**, sa používa aj na posúdenie akustiky miestnosti. Pre tento účel sa vyjadruje matematickým vzťahom:

$$m(F) = \frac{1}{\sqrt{1 + (2\pi F \cdot RT / 13,8)^2}} \cdot \frac{1}{1 + 10^{-\left(\frac{S/N}{10dB}\right)}}$$

F – modulačná frekvencia [Hz]

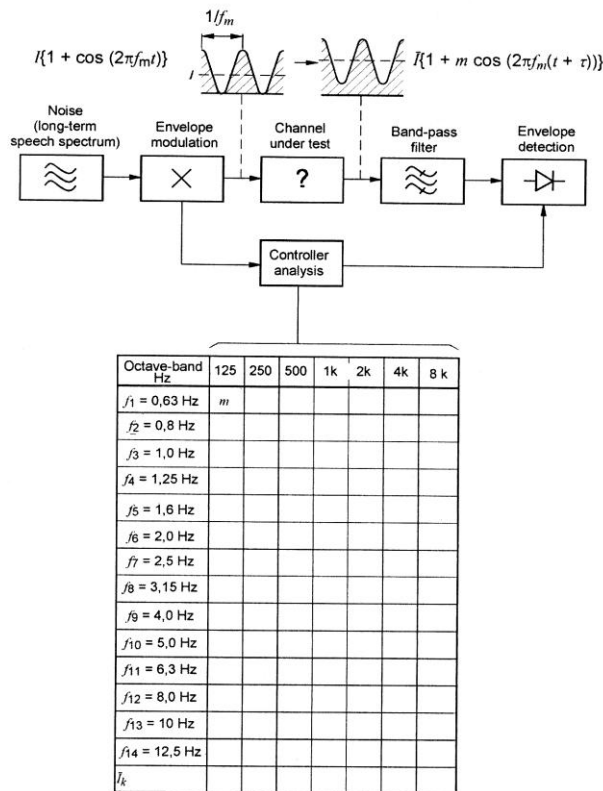
RT – doba dozvuku [s]

S/N – pomer signál/šum [dB]

Prvý činiteľ predstavuje len vplyv doby dozvuku (s exponenciálnym priebehom), druhý činiteľ predstavuje len vplyv rušivého hluku.

Meranie STI (niekedy označované ako FULL STI) vyžaduje uskutočniť 98 meraní (14 modulačných frekvencií v 7 oktávových pásmach), čo je časovo náročné (každé meranie sa uskutočňuje samostatne) a v praxi sa používa len v špeciálnych prípadoch. Princíp merania STI je uvedený na obr. 3.

Vo výpočte STI sa berú do úvahy aj vlastnosti sluchu, ktoré negatívne ovplyvňujú merané hodnoty zrozumiteľnosti. Je to efekt maskovania zvuku (silnejší zvuk nižšej frekvencie maskuje zvuk vyššej frekvencie) a efekt absolútneho prahového vnímania zvuku. Vlastnosti sluchu redukujú efektívny pomer signál/šum v rôznych oktávových pásmach a ich vplyv možno vyjadriť ako redukciu modulačnej prenosovej funkcie a všeobecne nižšiu hodnotu STI.



Obr. 3 Princíp merania STI a meracie frekvencie

Modulačná prenosová funkcia sa zahrnutím vlastností sluchu koriguje nasledovne:

$$m'_{k,fm} = m_{k,fm} \times I_k / (I_k + I_{am,k} + I_{rt,k}),$$

kde

$m_{k,fm}$ je hodnota modulačnej prenosovej funkcie pre oktávovové pásmo k a modulačnú frekvenciu f_m

I_k je intenzita akustického signálu v oktávovom pásme k

$I_{am,k}$ je intenzita maskovacieho signálu v oktávovom pásme k

$I_{rt,k}$ je prahová intenzita akustického signálu pre oktávové pásmo k

Princíp akustického maskovania je v STI metóde modelovaný tak, že nižšie oktávové pásmo $k-1$ má maskovací efekt na susedné vyššie oktávové pásmo k (125 Hz oktávové pásmo maskuje 250 Hz oktávové pásmo, 250 Hz oktávové pásmo maskuje 500 Hz oktávové pásmo, atď.). Oktávové pásmo 125 Hz nie je maskované vôbec.

Intenzita maskovacieho signálu $I_{am,k}$ v oktávovom pásme k je daná vzťahom:

$$I_{am,k} = I_{k-1} \times amf,$$

kde

I_{k-1} je intenzita akustického signálu v susednom nižšom oktávovom pásme $k-1$

amf je sluchový maskovací faktor (závislý od intenzity akustického signálu), ktorý je funkciou intenzity akustického signálu I_{k-1} v nižšom oktávovom pásme.

Intenzita akustického signálu I_{k-1} v oktávovom pásme $k-1$ je daná vzťahom:

$$I_{k-1} = 10^{(L_{k-1}/10)}$$

kde L_{k-1} je úroveň akustického tlaku [dB] v oktávovom pásme $k-1$.

Sluchový maskovací faktor **amf** pre oktávové pásmo **k** závisí na intenzite akustického signálu v oktávovom pásme **k-1**. V tabuľke je uvedená hladina sluchového maskovania **amdB** [dB] pre oktávové pásmo **k** ako funkcia úrovne hladiny akustického tlaku **L** [dB] v oktávovom pásme **k-1**. Úroveň sluchového maskovania je funkciou celkovej hladiny akustického tlaku v danom oktávovom pásme (úroveň testovacieho signálu a úrovne okolitého hluku).

Hladina akustického tlaku L v oktávovom pásme k-1 [dB]	< 63	≥ 63 a < 67	≥ 67 a < 100	≥ 100
Hladina sluchového maskovania amdB [dB]	0,5 x L - 65	1,8 x L - 146,9	0,5 x L - 59,8	-10

Sluchový maskovací faktor **amf** pre oktávové pásmo je daný vzťahom:

$$amf = 10^{(amdB/10)},$$

kde **amdB** je hladina sluchového maskovania [dB] v oktávovom pásme.

Intenzita prahového vnímania signálu $I_{rt,k}$ v oktávovom pásme **k** je daná vzťahom:

$$I_{rt,k} = 10^{(ART_k/10)},$$

kde

ART_k je absolútna úroveň prahového vnímania [dB] reči v oktávovom pásme **k** (pozri tab.).

Oktávové pásmo [Hz]	125	250	500	1 000	2 000	4 000	8 000
Absolútna úroveň prahového vnímania reči ART [dB]	46	27	12	6,5	7,5	8	12

Korigovaná úroveň modulačnej prenosovej funkcie sa použije vo výpočte efektívneho pomeru signál/šum **SNR_{eff k,fm}** nasledovne:

$$SNR_{eff\ k,fm} = 10 \times \log [m'_{k,fm} / (1 - m'_{k,fm})]$$

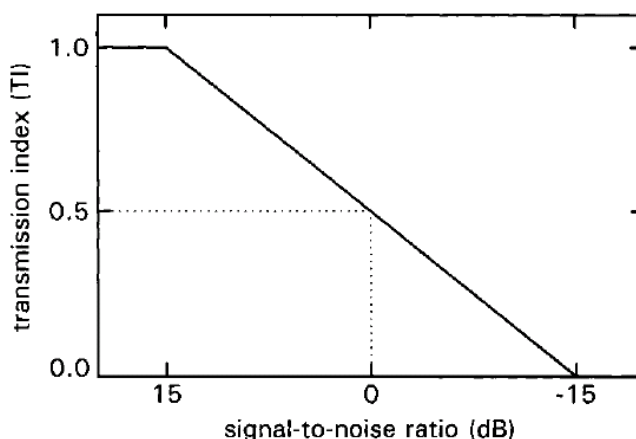
Pretože výsledná hodnota pomeru signál/šum môže nadobudnúť nekonečnú hodnotu, je potrebné hodnoty ohraničiť v rozsahu -15 dB až +15 dB.

Z hodnôt **SNR_{eff}** pre každé oktávové pásmo **k** a modulačnú frekvenciu **f_m** sa vypočíta **prenosový index TI** podľa vzťahu:

$$TI_{k,fm} = (SNR_{eff\ k,fm} + 15) / 30,$$

kde **SNR_{eff k,fm}** je efektívny pomer signál/šum pre každé oktávové pásmo **k** a **f_m** je modulačná frekvencia.

Podľa koncepcie STI je hodnota **TI_{k,fm}** v rozsahu 0 < **TI_{k,fm}** < 1 (pozri nasledovný obr.).



Vzťah medzi efektívnym pomerom signál/šum $SNR_{eff\ k,fm}$ a prenosovým indexom $TI_{k,fm}$ (rozsah 30 dB, posun +15 dB).

Hodnota **modulačného prenosového indexu** MTI_k v oktávovom pásme k sa vypočíta ako priemerná hodnota prenosových indexov $TI_{k,fm}$:

$$MTI_k = \frac{1}{n} \sum_{m=1}^n TI_{k, f_m}$$

kde

$TI_{k,fm}$ je prenosový index pre oktávové pásmo k a modulačnú frekvenciu f_m

m je index modulačnej frekvencie

n je počet modulačných frekvencií v oktávovom pásme

STI sa vypočíta zo známych hodnôt **modulačných prenosových indexov** MTI_k v každom oktávovom pásme k nasledovne:

$$STI = \sum_{k=1}^7 \alpha_k \times MTI_k - \sum_{k=1}^6 \beta_k \times \sqrt{MTI_k \times MTI_{k+1}}$$

kde

MTI_k je modulačný prenosový index pre oktávové pásmo k

α_k je váhový oktávový faktor pre pre oktávové pásmo k

β_k je redundančný faktor medzi oktávovým pásmom k a oktávovým pásmom $k+1$

Váhové faktory α_k a β_k pre mužskú reč ako funkcie oktávových pásiem sú uvedené v nasledovnej tabuľke:

Oktávové pásmo [Hz]	125	250	500	1 000	2 000	4 000	8 000
α	0,085	0,127	0,230	0,233	0,309	0,224	0,173
β	0,085	0,078	0,065	0,011	0,047	0,095	-

Zhrnutie postupu výpočtu zrozumiteľnosti reči priamou metódou STI

Namerané hodnoty modulačnej prenosovej funkcie pre oktávové pásmo k a modulačnú frekvenciu f

Namerané úrovne akustického tlaku v oktávach $L_{eq} \Rightarrow$ intenzita akustického signálu I_k v oktávach

Výpočet intenzity maskovacieho signálu I_{amk} v oktávach

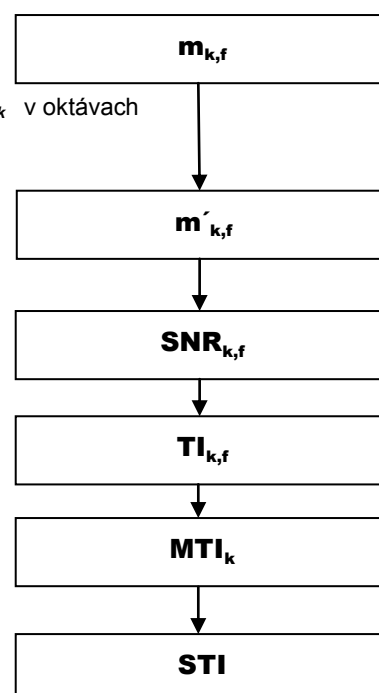
Výpočet intenzity prahového vnímania signálu reči $I_{rt,k}$ v oktávach

Výpočet efektívnej hodnoty pomeru signál/šum $SNR_{k,f}$ pre oktávové pásmo k a modulačnú frekvenciu f

Výpočet prenosového indexu $TI_{k,f}$ pre oktávové pásmo k a modulačnú frekvenciu f

Výpočet modulačného prenosového indexu MTI_k pre oktávové pásmo k

Výpočet hodnoty prenosového indexu reči **STI** použitím váhových faktorov α, β



Meracia metóda STIPA

Meracia metóda STIPA je zjednodušená metóda na rýchle a praktické meranie zrozumiteľnosti reči. Používa špeciálne vyvinutý testovací signál, vyvinutý podľa vlastností mužskej reči.

Namiesto 14 modulačných frekvencií, ktoré sa aplikujú samostatne za sebou vo všetkých 7 oktávových pásmach (FULL STI), STIPA používa pri meraní všetky modulačné frekvencie súčasne (pre každé oktávové pásmo 2 rôzne modulačné frekvencie). Celkovo sa používa $2 \times 7 = 14$ rôznych modulačných frekvencií.

Oktávové pásmo [Hz]	125	250	500	1 000	2 000	4 000	8 000
Prvá modulačná frekvencia [Hz]	1,60	1,00	0,63	2,00	1,25	0,80	2,50
Druhá modulačná frekvencia [Hz]	8,00	5,00	3,15	10,0	6,25	4,00	12,5

Meracia metóda STIPA je odporúčaná (STN EN 60286-16 rev. 4) na meranie zrozumiteľnosti reči v týchto aplikáciách:

- Meranie akustiky miestnosti (bez elektronického zosilnenia zvukového signálu). Živého rečníka možno jednoducho simulovať reproduktorom (s max. priemerom 10 cm a vyrovnanou frekvenčnou charakteristikou v rozsahu 100 Hz – 10 kHz) umiestneným na mieste rečníka vo výške úst. Na zosilnenie testovacieho signálu STIPA sa použije zosilňovač s výkonom cca 10W.
- Meranie ozvučovacích systémov PA
- Meranie a posúdenie hlasových evakuačných systémov (systémov HSP)
- Meranie prenosového kanála s amplitúdovou kompresiou signálu

STIPA umožňuje merať zrozumiteľnosť reči za prítomnosti rušivého hluku prostredia, odrazov, ozvien a nelineárneho skreslenia (spôsobeného napr. orezaním signálu pri prebudení).

Nízke namerané hodnoty zrozumiteľnosti STI môžu byť spôsobené:

- Veľkou dobou dozvuku miestnosti, ozvenou a odrazmi
- Nedostatočnými smerovými vlastnosťami reproduktorov alebo nedostatočným pokrytím miestnosti signálom z reproduktorov
- Nevhodnou úrovňou signálu z reproduktorov, napr. nízkou úrovňou (malý pomer signál/šum)