

1. Zvuk

zvuk – kmitání molekul v oboru slyšitelných frekvencí, tlakové vzruchy – zhuštěniny, zředěniny – podélné vlnění

sinusová vlna – nejpřirozenější kmitání, perioda T [s], frekvence f [Hz], amplituda, fáze, matematicky = amplituda * $\sin(2 * \pi * \text{frekvence} * \text{časový_okamžik} + \text{fáze})$

rychlost zvuku c závisí na médiu, při 20°C 343 m/s, ve vokálním traktu 350 m/s (vyšší vlhkost)

vlnová délka λ - zastavíme-li čas, tak vzdálenost, kdy se začíná vlna opakovat (délka periody)

$$\lambda = c / f$$

(jde jednoduše odvodit ze vztahu pro rychlost: $v = s / t \rightarrow c = \lambda / T \rightarrow c = \lambda / (1/f) \rightarrow c = \lambda * f$,

$$\lambda = c / f)$$

šíření zvuku – Huygensův princip (šíření vlnoplochy, ohyb a odraz, delší (pomalejší) vlny se lépe ohýbají), Dopplerův efekt

druhy vlnění – příčné (voda, lano), podélné (pružina, zvuk). Postupné – klasické, postupně se šíří. Dojde-li k odrazu, sčítá se dopředná vlna s odraženou a vzniká stojaté vlnění (dojem ustáleného stavu) – některé body jsou v klidu (uzly), jiné kmitají nejvíce (kmitny).

periodické vlny – pravidelně se opakující či kvaziperiodické vlny

složené vlny – Fourierův teorém, harmonické složky, základní tón a vyšší harmonické (aliquóty, částkové tóny), f_0 , H_1 , H_2 , H_3 ...

spektrum – frekvenční oblast, pro periodické signály diskrétní, pro impulsy spojité, spektrum krátkých vs. delších impulsů, spektrální sklon

rázy – blízké frekvence, vznikají záněže. Frekvence kolísání dána rozdílem frekvencí, vnímaná frekvence (subjektivní pocit) je průměr obou frekvencí.

Dopplerův jev – vzniká, pokud se zdroj či příjemce zvuku pohybuje.

spektrogram – spektra ze segmentovaného signálu. Čím delší segment (hrubší časové rozlišení), tím jemnější frekvenční rozlišení. Ve fonetické praxi často 5 ms – harmonické složky se slévají dohromady, vidíme oblasti vyšší energie (formanty), zároveň díky jemnému časovému rozlišení můžeme sledovat vliv hlasivkových pulsů. Nebo 15 ms – pro zobrazení harmonických složek.

kochleogram – modeluje reakci bazilární membrány. Podobný spektrogramu, ale jiná osa y (např. barky) – nižší frekvence zabírají více místa, navíc není tolik „rozmazaný“.

základní klasifikace hlásek – vokály a konsonanty (sonory či obstruenty)

základní formy řečového zvuku – frikce, kvaziperiodický zvuk, nemá forma, exploze. Jak je poznáme v časové oblasti (oscilogram) či ve frekvenční (spektrogram)? Viz Milánek.

2. Rezonance

rezonanční křivka - materiály rády kmitají, díky rozměrům a vlastnostem se některým frekvencím daří lépe (vzniká silné stojaté vlnění) než jiným (zatlumí se). Rezonuje i vokální trakt. Rezananční křivka – rezonanční (střední) frekvence f_s , šířka pásma Δf (pokles amplitudy o 3 dB vůči rezonanční frekvenci, nebo-li na 70,7 % amplitudy = 50 % výkonu), strmost úpadku dB / oktávu (2x) či dB / dekádu (10x).

filtry – pásmový filtr (pásmová propust band pass), blokující pásmový filtr (pásmová zadrž band stop), dolní propust (low pass), horní propust (high pass).

vokální trakt – kombinace více filtrů (rezonančních křivek), proto ve výsledku více kopců (frekvence vrcholů nazýváme formanty).

3. Psychoakustika

lidské ucho – vnější ucho (zevní zvukovod, bubínek), střední ucho (převod zvukové energie ze vzduchu do kapaliny) – kladívko, kovadlinka, třmínek – pákový mechanismus (reflex – při velkém hluku (asi nad 90 dB) se svaly stáhnou a přenos frekvencí pod 1 kHz se sníží), vnitřní ucho (hlemýžď – cochlea, rozdělena 2 membránami, na bazilární membráně nervové buňky, propojeny s mozkem). Membrána v různých částech různě tuhá a různě tlustá – naladěna na různé rezonanční frekvence. Začátek naladěn na vysoké frekvence, konec na hluboké. Dochází k frekvenčnímu rozkladu (podobné Fourierovi).

vnímání výšky – zjednodušeně f_0 je základní kmitočet, vyšší harmonické určují barvu. F_0 si můžeme ale domyslet, pokud není přítomna – stačí, když ucho rozpozná periodu opakujících se úseků (podle konstantní vzdálenosti frekvenčních čar vyšších harmonických složek ve spektru). Vnímáme přibližně 16 Hz – 20 kHz. Pod je infrazvuk, nad ultrazvuk.

diference limen (just noticeable difference) – minimální vnímatelný rozdíl (ať už výšky či hlasitosti). DL výšky – přibližně 0,5 % frekvence, DL hlasitosti – zhruba 1 dB, záleží na povaze zvuku.

hudební stupnice – stejný poměr frekvencí vnímáme jako stejnou intervalovou vzdálenost.

subjektivní vnímání výšky – cca od 800 Hz výše je vnímání výšky zkreslené oproti matematicky ideální hudební stupnici – na základně experimentů vznikla stupnice mel – vylepšené Hz, kde opravdu platí, že stejný poměr vede na stejnou vnímanou vzdálenost. Stupnicí mel je více variant (bylo provedeno více experimentů s různými výsledky).

kritické pásmo - v rámci něj dochází k maskování, např. tón může být zamaskován silným šumem obklopujícím frekvenci tónu. Stupnice Bark (starší) a Erb (novější, přesnější). 1 bark i erb odpovídají šířce 1 kritického pásma. Do 500 Hz je šířka kritického pásma přibližně 100 Hz, pak se zvyšuje. 0 bark odpovídá 0 Hz, 1 bark 100 Hz atd. Přibližné pravidlo nad 500 Hz – šířka pásma je cca 17 % střední frekvence.

objektivní popis síly zvuku

- rms amplituda (root-mean-square) – všechny hodnoty amplitudy na druhou, pak průměr,

z výsledku odmocnina – pro různé tvary vln přepočítání na obdélník, který by způsobil stejnou energii – dobré pro porovnávání. Min. vnímaný tlak je asi 10 μPa , maximální asi 64 Pa, tedy 64 milionů krát větší.

- výkon W [watt] = energie za jednotku času (energie je úměrná amplitudě na druhou)
- intenzita I [watt / m^2] = výkon vztažený na plochu (při směrovém zvuku může být v každém směru jiná hodnota). Klesá s druhou mocninou vzdálenosti od bodového zdroje.
- akustický tlak p [Pa] – tlak, kterým zvuk působí např. na mikrofon či bubínek v uchu
- decibely [dB] – nejčastěji používané, nesmírně praktické. Z výkonů: $\text{dB} = 10\log(\text{poměr})$, z amplitud $\text{dB} = 20\log(\text{poměr})$. Výsledek jsou pak již ty samé dB, které vždy odpovídají logaritmu poměru výkonů. Základní dB odpovídají vždy poměru. dB sčítáme a odčítáme. Zesilovač zesiluje 20 dB – přičteme k hladině 20 dB, přestože amplitudu ve skutečnosti zesílí 10x (násobilo by se). 0 dB neznamená nulovou hlasitost – znamená to stejnou hlasitost, poměr 1.

0 dB = 100%

-3 dB = 70,7% amplitudy či 50% výkonu.

10 dB = 10x výkon, subjektivně vnímáno jako cca 2x větší hlasitost (platí pro 1 kHz čistý tón)

20 dB = 100x výkon, 10x amplituda, subjektivně cca 4x větší hlasitost

opačné vzorec – poměr výkonů = $10^{\text{dB}/10}$, poměr amplitud = $10^{\text{dB}/20}$

Absolutní hladina hlasitosti v dB – opět poměr, ale vztažen k referenční hodnotě 20 μPa při 1000 Hz. Označení dB SPL (sound pressure level). Tzv. **hladina zvuku**.

Počítačové dB (FS), zkratka FS se často vynechává. 0 dB FS je vztaženo k sinusovce o maximální amplitudě, kterou dovolí bitová hloubka zvukového souboru. Proto většina zvuků má dB záporné = jsou slabší. Pokud bychom sinusovku 0 dB FS zesilovali, došlo by ke zkreslení – k ořezu vysokých hodnot amplitudy mimo povolený rozsah. Tím se dostáváme do kladných dB, dochází k nevratnému a slyšitelnému poškození signálu, jev nazýváme clipping.

Subjektivní vnímání hlasitosti

Na různých frekvencích vnímáme čisté tóny různě nahlas. Izofona – křivka vnímání stejné hlasitosti (equal loudness countour). Vynášíme do sluchového pole.

Vnímaná hlasitost závisí na šířce pásma zvuku a je ovlivněna i kritickým pásmem (pokud šířka zvuku překročí šířku kritického pásma, vnímáme více nahlas), dále závisí na délce trvání (delší jsou více nahlas).

Práh slyšitelnosti – cca 0 dB (SPL), práh bolesti – cca 120 dB, práh nepříjemných počitků – cca 100 dB.

Diference limen – přibližně 1 dB, závisí na povaze zvuku.

Málo používané jednotky – hladina hlasitosti fon (Ph) – pro 1 kHz odpovídá hodnotě v dB, pro ostatní pak hodnotám dle izofon. Tedy stejné Ph – stejná vnímaná hlasitost pro různé frekvence. Hlasitost son – 1 son je při 1 kHz a 40 dB. Zvýšení intenzity o 10 dB vnímáme jako zdvojnásobení hlasitosti (při 1 kHz), sony jsou tedy také dvojnásobné (50 dB = 2 sony, 60 dB = 4 sony, 30 dB = 0,5 sonu). Tyto jednotky se týkají pouze čistých tónů. Vzhledem ke složitosti subjektivního vnímání ale raději zůstáváme u objektivních dB a k analýze psychoakustického vjemu používáme známé diagramy (izofony, sluchové pole apod.).

Kolísání frekvence a amplitudy – vibrato (pravidelné kolísání frekvence), třas = jitter (kolísání frekvence F_0), tremolo = shimmer (kolísání amplitudy (síly) F_0), glissando (klouzavé zvýšení či snížení F_0)

4. Filtrová teorie produkce řeči

model – zdroj, filtr, radiační impedance. Hlasivkové pulsy, sklon -12 dB/okt, filtr vokálního traktu (rezonance – formanty), rty – radiační impedance, sklon +6 dB/okt (horní propust)

hlasivkové pulsy – poměr otevření OQ, normálně 0,5, při hlasitějším např. 0,43 (též se více narovnává spektrální sklon), záleží též na způsobu fonace

fonace – normální, třepená (uvolněná a tlačaná), dyšná fonace.