

KUI342 Úvod do neurovied

Týždeň 9: Sluch, časť 1: metódy štúdia, funkcie, lineárny systém a Fourierova transformácia, binaurálne a priestorové vnímanie

Organizačné záležitosti

- Zajtra prednáška 10
- Zadanie 4 zverejnene – máte 2 týždne (do 7.12.; 12.týždeň)
- Projekt do 11. týždňa
- Záverečná písomka v 13. týždni v utorok
- Uzatváranie známok v 13. týždni (streda)
- Hodnotenia zadaní/písomiek – sledujte web

Minulý týždeň

Zrak 1 – doplním v 11. týždni, viac v 5. ročníku

Dnes

Vnemové systémy človeka – Časť 2 - sluch:

Čo je zvuk a sluch

Prehľad tém sluchového štúdia

Frekvenčná selektivita

Binaurálny a priestorový sluch

Zvuk a sluch - definície

Zvuk:

mechanické vlnenie hmotného prostredia vnímateľné sluchom (Krátky slovník slovenského jazyka)

ale tiež:

psychologický vnem vygenerovaný vnútorne v mozgu ako odozva na fyzikálny podnet nazývaný „zvuková vlna“

Sluchové vnímanie:

- subjektívny zážitok
- to, čo počujeme sa líši od toho, aký zvuk sa **fyzicky** dostal do našich uší

Tri nevyhnutné zložky:

- zdroj zvuku; médium, ktorým sa zvuk šíri; poslucháč

Témy sluchového štúdia

Sluchový podnet:

- fyzika vzniku, šírenia, skladania, filtrácie, odrazov zvukov

Spracovanie v periférnom sluchovom systéme:

- anatómia a fyziológia sluchového nervu
- neurálny kód – reprezentácia základných charakteristík zvuku (frekvencia, intenzita, časový priebeh) v sluchovom nerve

Vnímanie jednoduchých podnetov:

- absolútne prahy počuteľnosti
- prahy rozlíšiteľnosti frekvencie, intenzity, času, modulácia
- maskovanie jedného zvuku iným
- hlasitosť a výška tónu

Vnímanie zložitých podnetov a centrálny sluchový systém:

- priestorový sluch a binaurálne počutie, plasticita sluchového systému
- anatómia a fyziológia podkôrových a kôrových sluchových štruktúr
- analýza sluchovej scény, „grouping a streaming“, informačné maskovanie
- hudobné vnímanie
- spracovanie reči a jazyka
- poruchy sluchu a prostetika (kochleárne implantáty)

Funkcie sluchu

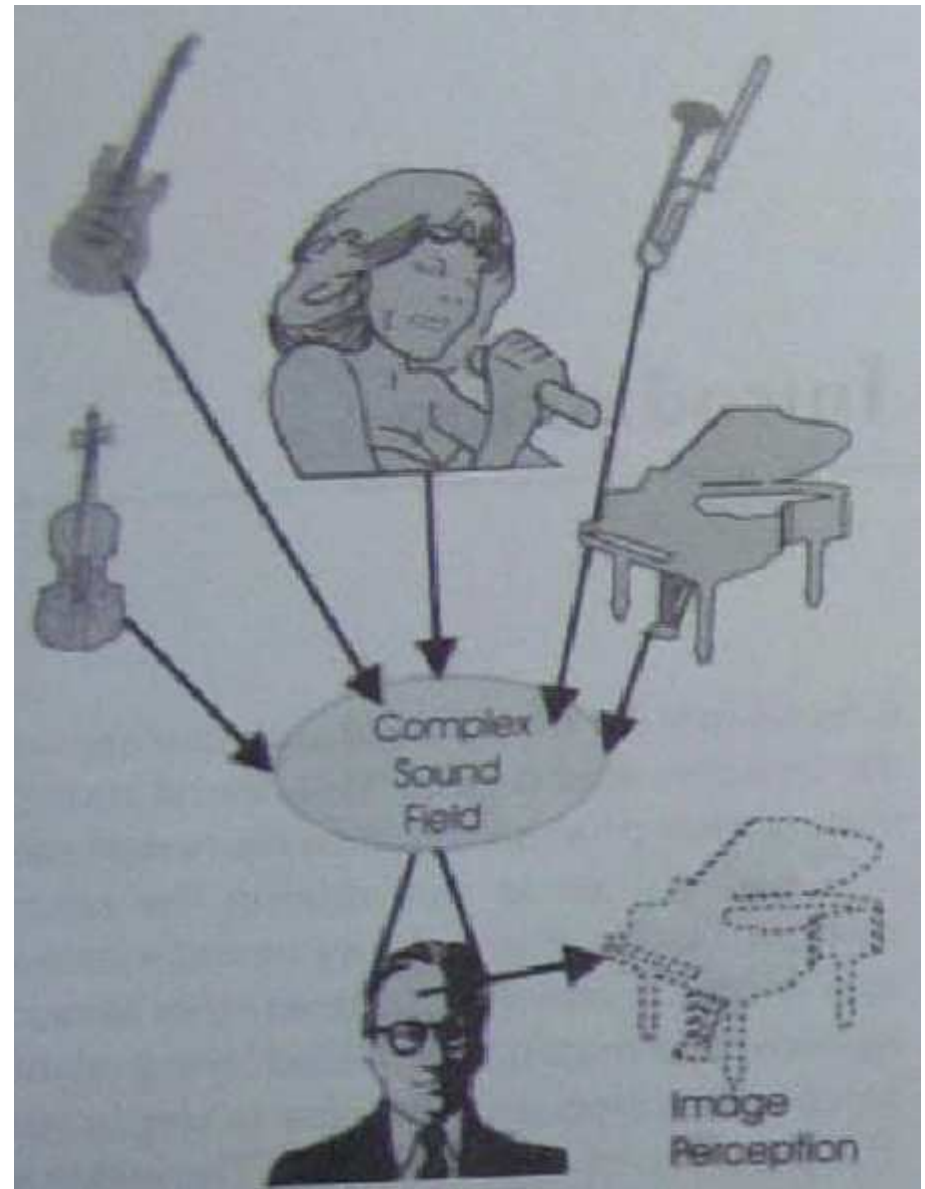
Človek je schopný zvuky:

- Identifikovať
- Lokalizovať
- Porozumieť

Ale ako vieme, ktorý zvuk pochádza z ktorého zdroja, ak počujeme viac zvukov súčasne?

Predstavte si úlohu identifikovať všetky lode v prístave na základe pohybu dvoch bóji...

Dnešná téma: Ako analyzujeme zvuk, a aké neurálne mechanizmy pri tom používame



Aké metódy môžeme použiť pri štúdiu sluchu?

Fyzika (Akustika)

Anatómia a fyziológia

Psychológia

Psychofyzika: Slúži na kvantifikáciu vzťahu medzi fyzikálnymi a psychologickými veličinami (Fechner, 1801-1887). Používa sa na štúdium ľudských zmyslov.

Fyzikálne

Zvuk

Intenzita

Frekvencia

Spektrum a časové
charakteristiky

Psychologické

Sluchový vnem

Hlasitosť

Výška tónu

Farba zvuku (timber)

Psychofyzika: Subjektívne a objektívne metódy

Subjektívne metódy

- odhadovanie hlasitosti, hľadanie tónu zhodnej výšky, hodnotenie „kvality zvuku“, atď
- vyžaduje sa introspekcia
- meranie ovplyvnené nezmyslovými faktormi (systematická chyba, odklon, bias)

Objektívne metódy

- detekcia stimulu alebo diskriminácia
- meria sa presnosť, takže musí existovať správna odpoveď
- vplyv systematickej chyby (odklonu/biasu) sa dá oddeliť od citlivosti použitím *Teórie detekcie signálov (Signal Detection Theory)*.

Sluch a schopnosť separovať frekvencie

Každá periodická funkcia sa dá popísať ako suma sínusoíd (Fourier).

Sluchový systém tiež funguje na tejto báze: rozkladá zvuk na jednotlivé frekvenčné zložky.

Frekvenčná selektivita

- základný organizačný princíp v sluchovom systéme
- mapovanie frekvencia → poloha sa deje už na úrovni kochley
- táto „tonotopická“ organizácia sa udržiava v celom sluchovom systéme
- pozn.: napriek zložitým transformáciám ďalej v sluchovej dráhe, hranice frekvenčnej selektivity sú určené na úrovni kochley

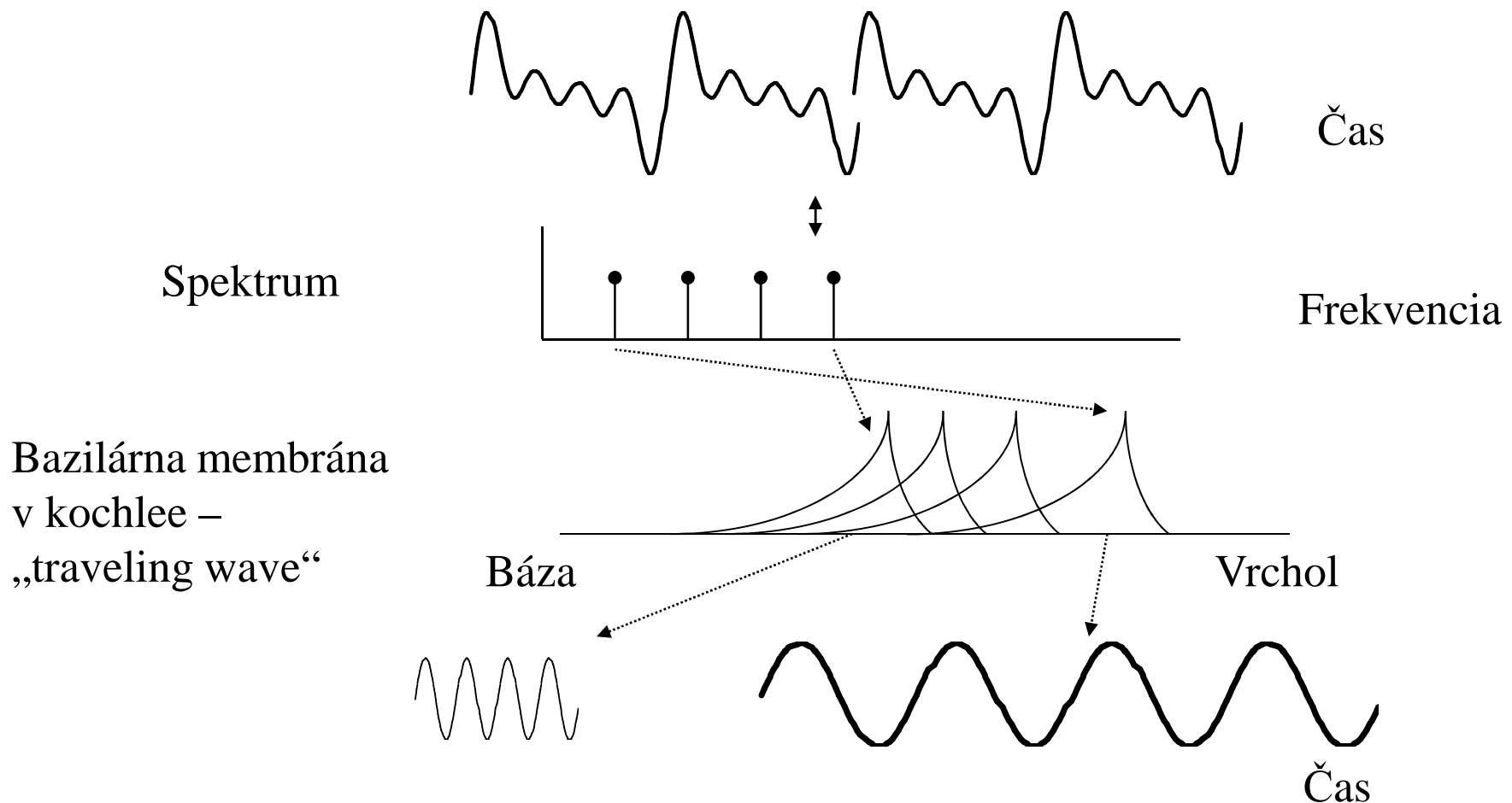
V roku 1863 Helmholtz ukázal, že je možné rozoznať jednotlivé sínusové zložky zloženého tónu [je to tak aj pri videní?]

Demo: Vynechané harmonické zložky 

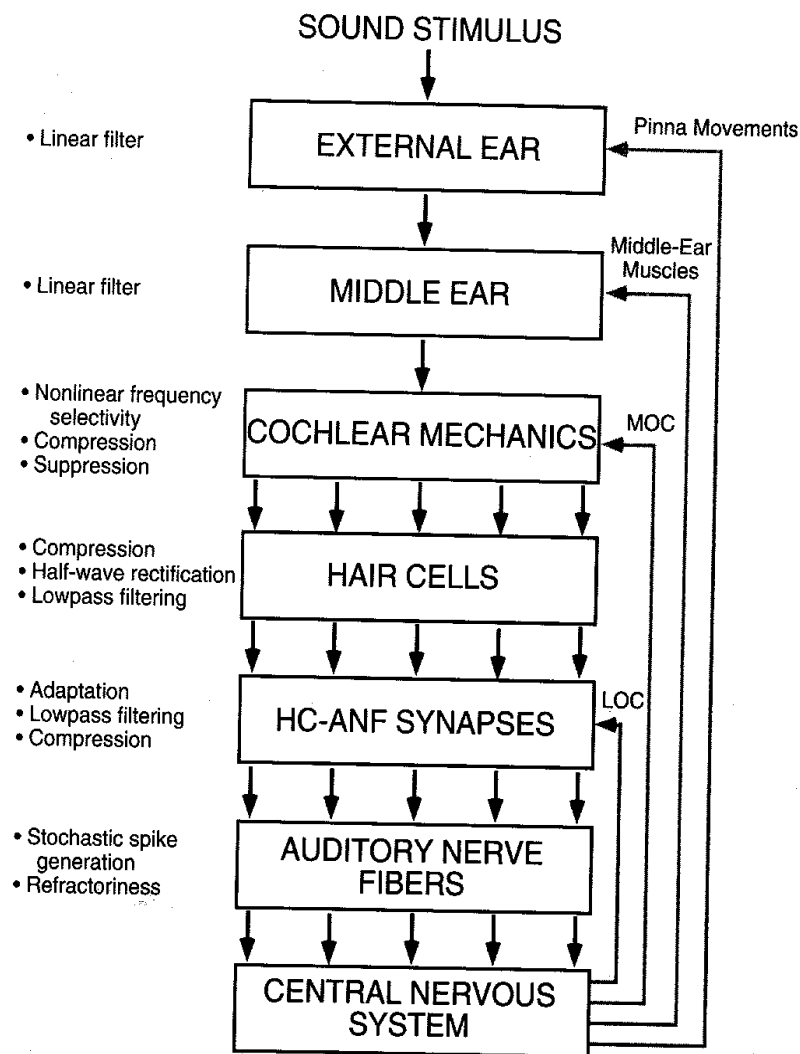
- *príklad syntetického a analytického počúvania: najprv počujeme **jeden** zvuk s jednou výškou tónu. Potom sa z neho **oddelia** jednotlivé harmonické zložky tým, že ich pridávame a uberáme (ASA Auditory Demonstrations CD)*

Ako funguje frekvenčná selektivita

Konverzia frekvencia → poloha v kochlee umožňuje oddeliť jednotlivé frekvenčné zložky zvuku [Cochlear traveling wave – von Békésy (1960)]



Proces vnímania zvuku



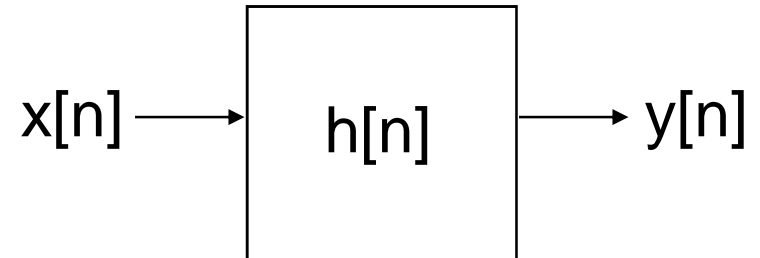
Akustická vlna sa doširi k hlave a vchádza do sluchového kanálu, kde spôsobí:
mechanické vibrácie v strednom uchu
šírenie vln pozdĺž kochley
pohyb vláskových buniek (senzorických neurónov) na bazilárnej membráne
depolarizáciu vláskových buniek
vláskové bunky vygenerujú AP
nervovo kódovaná informácia sa spracováva na rôznych úrovniach (často spätná väzba na predošlé úrovne)

mnoho úrovní sa dá popísať použitím teórie lineárnych systémov

Lineárne časovo invariantné (LTI) systémy

Matematický aparát vhodný na popis fungovania sluchového systému

Pre LTI systémy platí, že jedinou funkciou $h[n]$ vieme popísať všetko, čo potrebujeme k určeniu odozvy systému na ľubovoľný vstup



$h[n]$ sa nazýva prenosová funkcia alebo impulzná odozva (je to odozva, ak sa na vstup privedie impulz v čase $n = 0$)

Odozvu na ľubovoľný vstup získame konvolúciou:

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$

$$y[n] = x[n] * h[n]$$

Fourierova transformácia

Funkcia $X(e^{j\omega})$ je fourierovou transformáciou funkcie $x[n]$ (závislou na ω).

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} = \mathfrak{F}(x[n])$$

Fourierova transformácia $x[n]$:

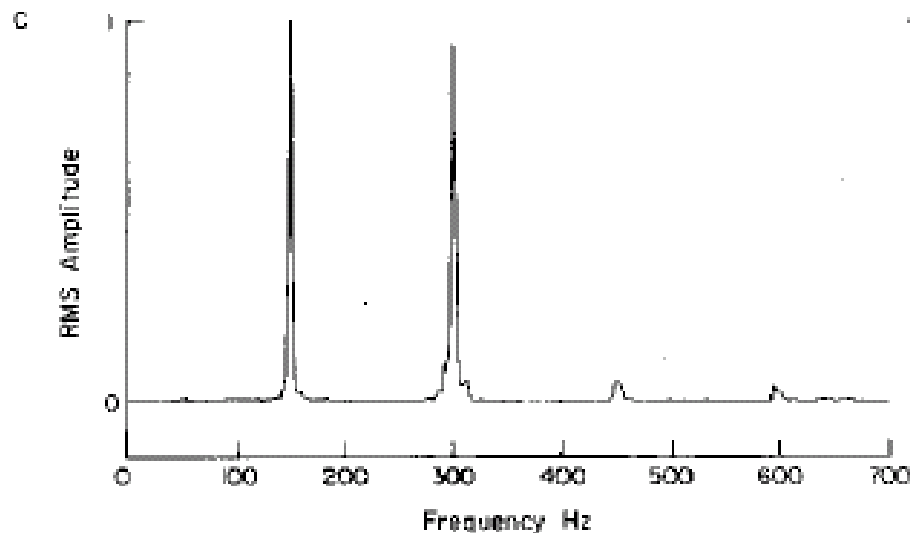
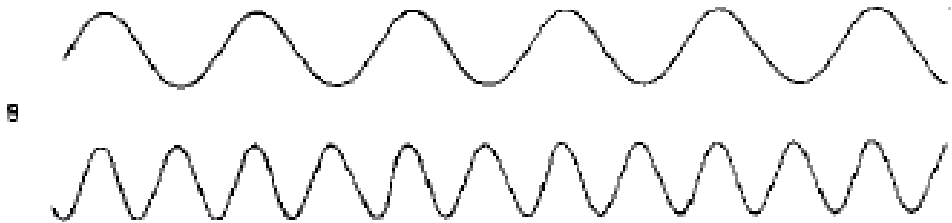
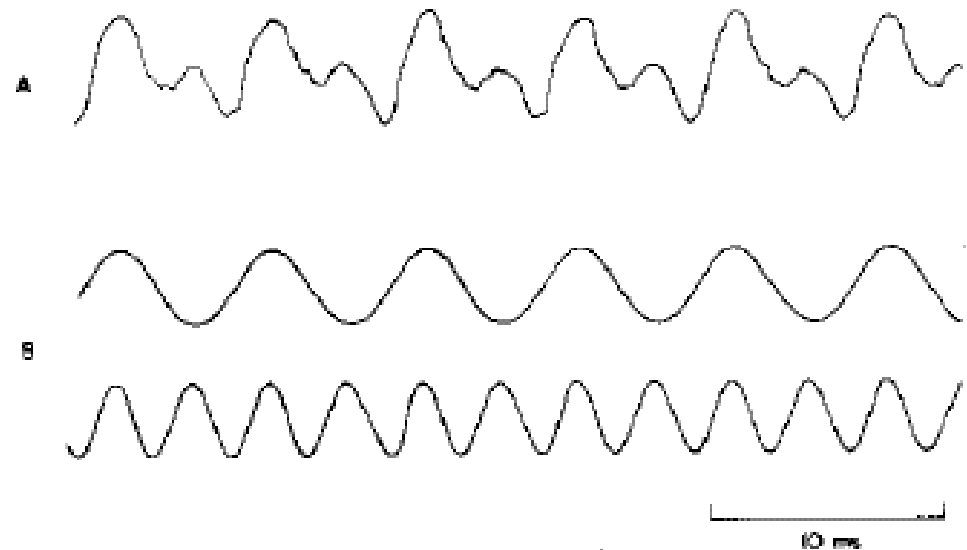
- je alternatívnou reprezentáciou funkcie $x[n]$ s použitím ortonormálnych báзовých funkcií $\{e^{j\omega}\}$
- rozkladá $x[n]$ na jej frekvenčné zložky
- $X(e^{j\omega})$ je komplexná funkcia - má reálnu a imaginárnu časť
- $X(e^{j\omega})$ sa dá popísať amplitúdou a fázou (táto forma je častejšia než komplex.)
- amplitúda určuje veľkosť sínusovej zložky signálu s frekvenciou ω
- fáza popisuje časový posun, pri ktorom dosahuje sínusoida maximum
- ak poznáme $X(e^{j\omega})$, inverznou Fourierovou transformáciou môžeme získať pôvodnú časovú reprezentáciu signálu $x[n]$
- poznať len amplitúdu alebo len fázou na výpočet inverznej transformácie nestačí

Fourierova transformácia popisuje frekvenčné zložky signálu

A: Suma dvoch sinusoíd

B: Zložky signálu

C: Amplitúdové
spektrum Fourierovej
transformácie signálu
z A

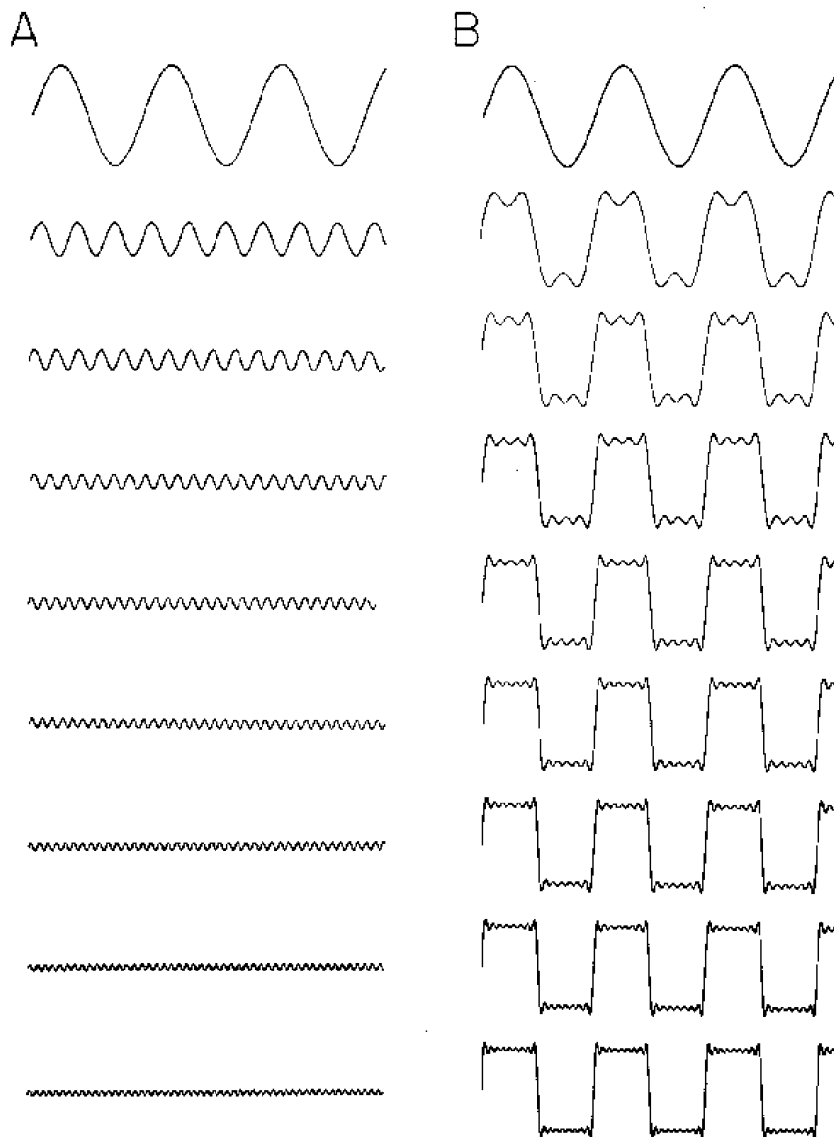


Príklad – pravouhlý priebeh

Sčítaním sínusoid s
nepárnym násobkom
frekvencie získame
pravouhlý priebeh:

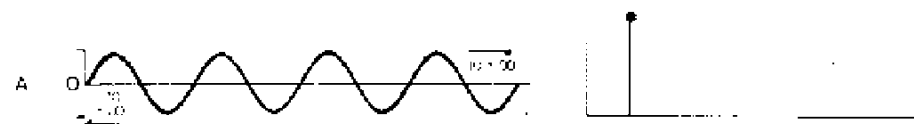
A: sínusové zložky

B: Postupne pričítavané
zložky z A

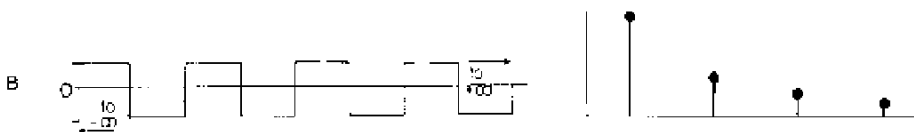


Ďalšie príklady Fourierovej transformácie

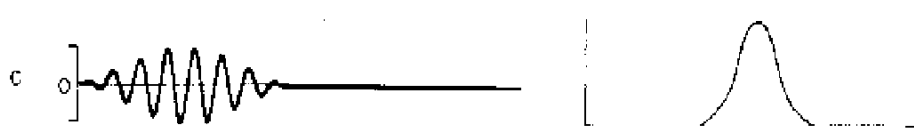
A: Čistá sínusoida



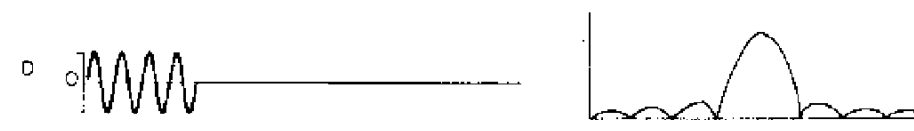
B. Pravoúhlý priebeh



C. sínusoida s postupným nábehom a ukončením



D. sínusoida s pravouhlým nábehom a ukončením



E. Impulz



F. Biely šum



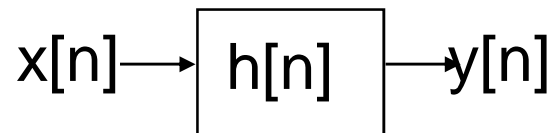
Time

Frequency

Základné operácie Fourierovej transformácie

súčet $\mathfrak{F}(x[n] + y[n]) = X(e^{j\omega}) + Y(e^{j\omega})$

impulz $\delta[n] = \begin{cases} 1 & \text{pre } n=0 \\ 0 & \text{inak} \end{cases} \quad \mathfrak{F}(\delta[n]) = 1$



konvolúcia $y[n] = x[n] * h[n] \quad Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega})$

odozva na impulz je prenosovou funkciou systému $x[n] = \delta[n] \Rightarrow \begin{cases} Y(e^{j\omega}) = 1 \cdot H(e^{j\omega}) = H(e^{j\omega}) \\ y[n] = h[n] \end{cases}$

Fyzikálny popis zvuku: efektívny výkon, intenzita

Akustický tlak zvukovej vlny určuje akustický výkon (hlasitosť)

Efektívny výkon (RMS, Root-mean-square) pre tlakovú vlnu $p(t)$ udáva priemernú hodnotu amplitúdy akustickej vlny, a je priamo úmerný intenzite zvuku

$$P_{\text{RMS}} = \sqrt{\frac{1}{T} \int_{t_0}^{t_0+T} p^2(t) dt}$$

Ak $p(t) = A \sin \omega t$, $P_{\text{RMS}} = 0.707A$

Intenzita (výkon/plocha) je približne priamo úmerná hlasitosti

$$I = \frac{P_{\text{RMS}}^2}{\rho_0 c}$$

ρ_0 = hustota média

c = rýchlosť zvuku

Decibelová mierka

Používa sa, pretože rozsah hlasitosti, ktorú môže človek vnímať, je veľký

Rozsah:

0.00002 Pa (0 dB SPL) – Prah citlivosti

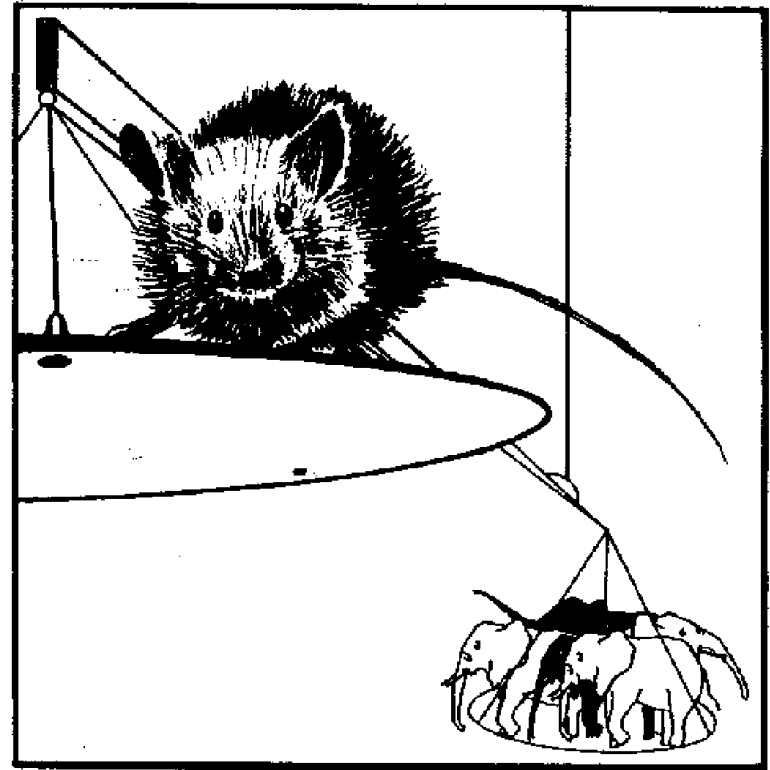
20 Pa (120 dB SPL) – Prah bolesti

(zodpovedá pomeru hmotnosti myši k 5 slonom)

Ďalší dôvod: Citlivosť človeka na zmenu intenzity je charakterizovaná *Weberovým pravidlom*: najmenšia vnímateľná zmena je priamo úmerná intenzite

Decibel:

$$10 \log_{10} \frac{I}{I_{\text{ref}}} = 10 \log_{10} \frac{P_{\text{RMS}}^2}{P_{\text{ref.RMS}}^2} = 20 \log_{10} \frac{|P_{\text{RMS}}|}{|P_{\text{ref.RMS}}|}$$



Decibelová mierka

Referenčná hodnota je ľubovoľná

Dve často používané referenčné hodnoty sú

- SPL (Sound Pressure Level)
 - $I_{\text{ref}} = 10^{-12} \text{ W/m}^2$
 - $P_{\text{ref}} = 20 \mu\text{Pa}$
 - približne prahová hodnota pre človeka počúvajúceho 1000-Hz tón
- SL (Sensation Level)
 - definuje sa pre každého individuálne
 - predstavuje prah citlivosti konkrétneho človeka

Šírenie zvuku

Ak je zdroj zvuku bodový (ideálny), zvuk sa z neho šíri kruhovými vlnami

Pre zjednodušenie, neuvažujeme pri šírení žiadne straty

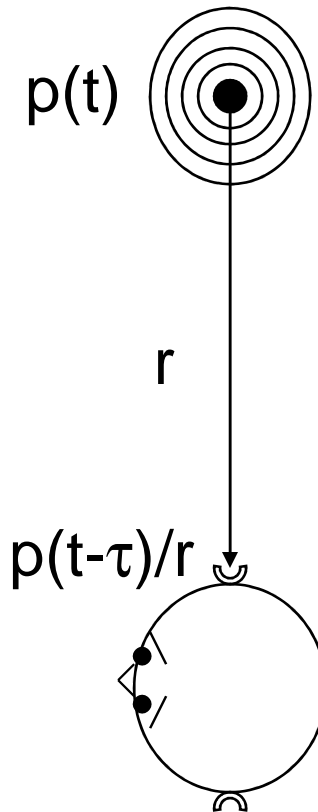
Ak zdroj generuje zvuk s energiou E

(a kruhová vlna má plochu πr^2)

=> intenzita zvuku vo vzdialenosti r je $E/(\pi r^2)$

Energia je priamo úmerná P_{RMS}^2 => takže amplitúda akustického tlaku je úmerná $1/r$

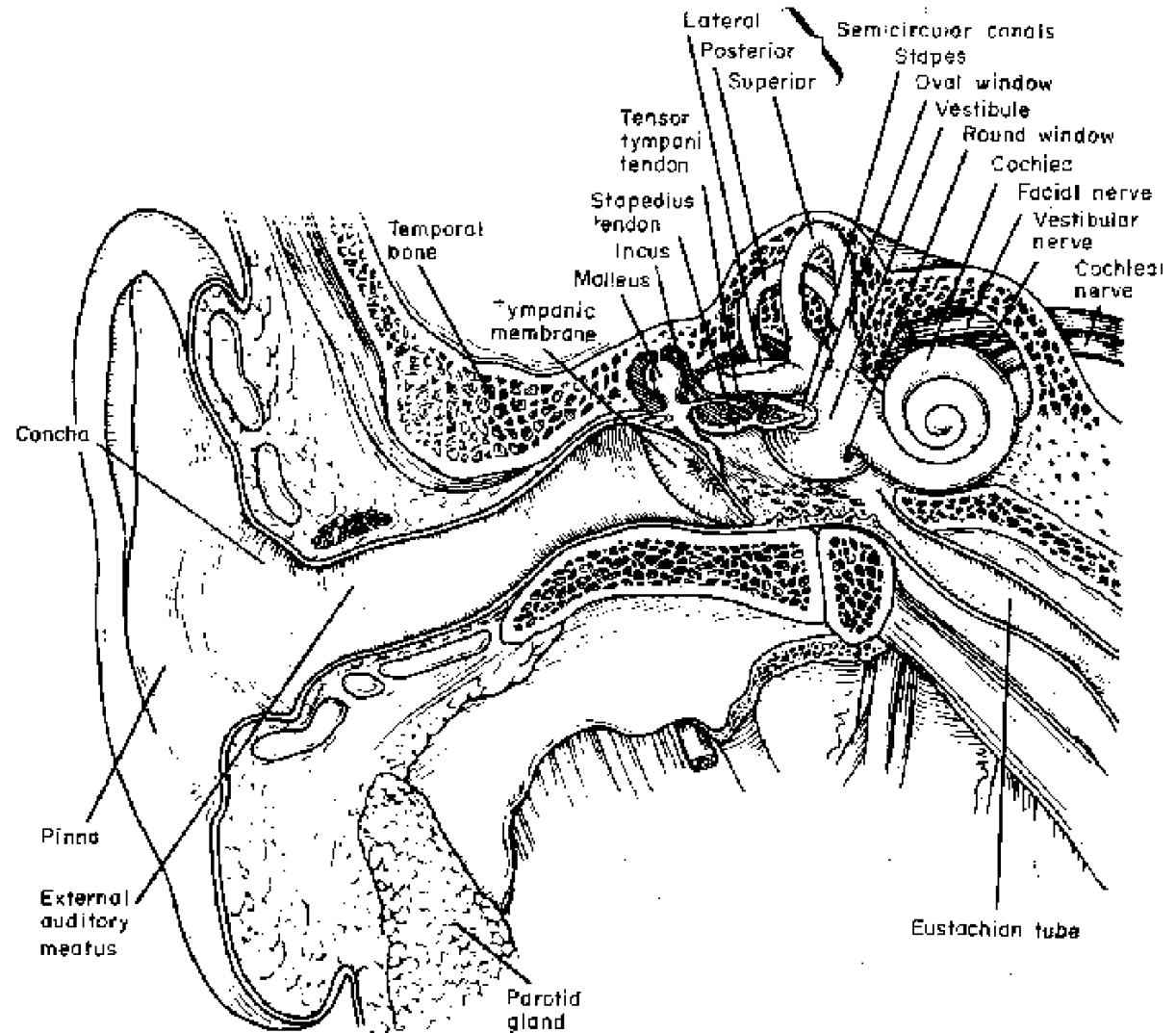
Akustická vlna dorazí do ucha v čase $\tau = r/c$ sekúnd (c je rýchlosť zvuku)



Od akustických vln po kochleu

Pri šírení z prostredia sa zvuk postupne mení rôznymi interakciami:

- telo a hlava
- vonkajšie ucho: ušnica
- sluchový kanál až bubienok
- stredné ucho
- kochlea (slimák)
- vlasové bunky (tiež vestibulárny systém)
- sluchový nerv



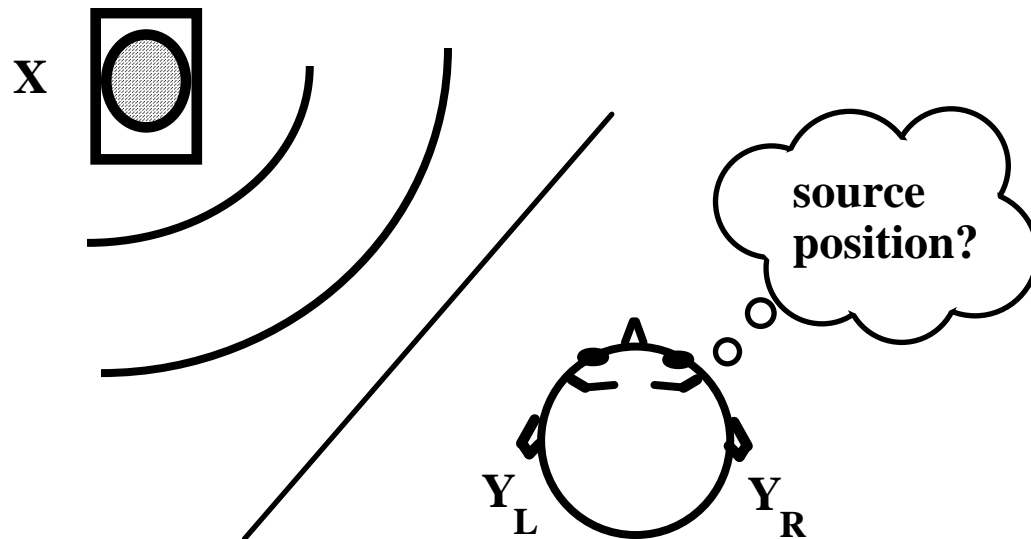
Interakcia zvuku s hlavou a HRTF

To, ako sa zvuk zmení po ceste od zdroja po bubienok, sa dá charakterizovať nájdením prenosovej funkcie:

$$y_L[t] = \frac{1}{r} h_L[t, \theta, \phi] * x_L[t] \quad H_L[\omega, \theta, \phi] = \mathfrak{S}(h_L[t, \theta, \phi])$$

$$y_R[t] = \frac{1}{r} h_R[t, \theta, \phi] * x_R[t] \quad H_R[\omega, \theta, \phi] = \mathfrak{S}(h_R[t, \theta, \phi])$$

Táto prenosová funkcia sa nazýva Head-Related Transfer Function (HRTF)



Pre zdroje aspoň 1 meter od hlavy sa HRTF mení len s azimutom a eleváciou

Pre blízke zdroje sa HRTF mení aj so vzdialenosťou

Ak poznáme HRTFs, môžeme:
- analyzovať ako sa zvuk mení v závislosti na polohe
- simulovať rôzne zdroje zvuku

Načo máme dve uši?

Lokalizovanie zvukov v priestore: v porovnaní so zrakom je sluch podstatne horší (dva rády), ale pokrýva celý priestor. Užitočné napr. ako výstražný mechanizmus.

Separácia zvukov: ak počúvame viac vecí naraz, sluchový systém môže porovnať zvuk zaznamenaný ušami a na základe tohoto porovnania zlepšiť porozumenie jednému z počutých stimulov. Používame napr. ak hovorí viac ľudí naraz (cocktail party effect), ale aj pri potláčaní odrazov zvuku od stien.

Sluchová lokalizácia: Vnímanie azimutu a elevácie

Základné binaurálne parametre, ktoré sluchový systém používa na určenie polohy zdroja zvuku sú:

interaурálne časové rozdiely – rozdiel medzi časmi, keď zvuk dorazí k ľavému a k pravému uchu

interaурálne rozdiely v intenzite – rozdiel medzi intenzitami, s akými zvuk dorazí k ľavému a k pravému uchu

Interaurálne časové rozdiely – Interaural Time Differences (ITDs)

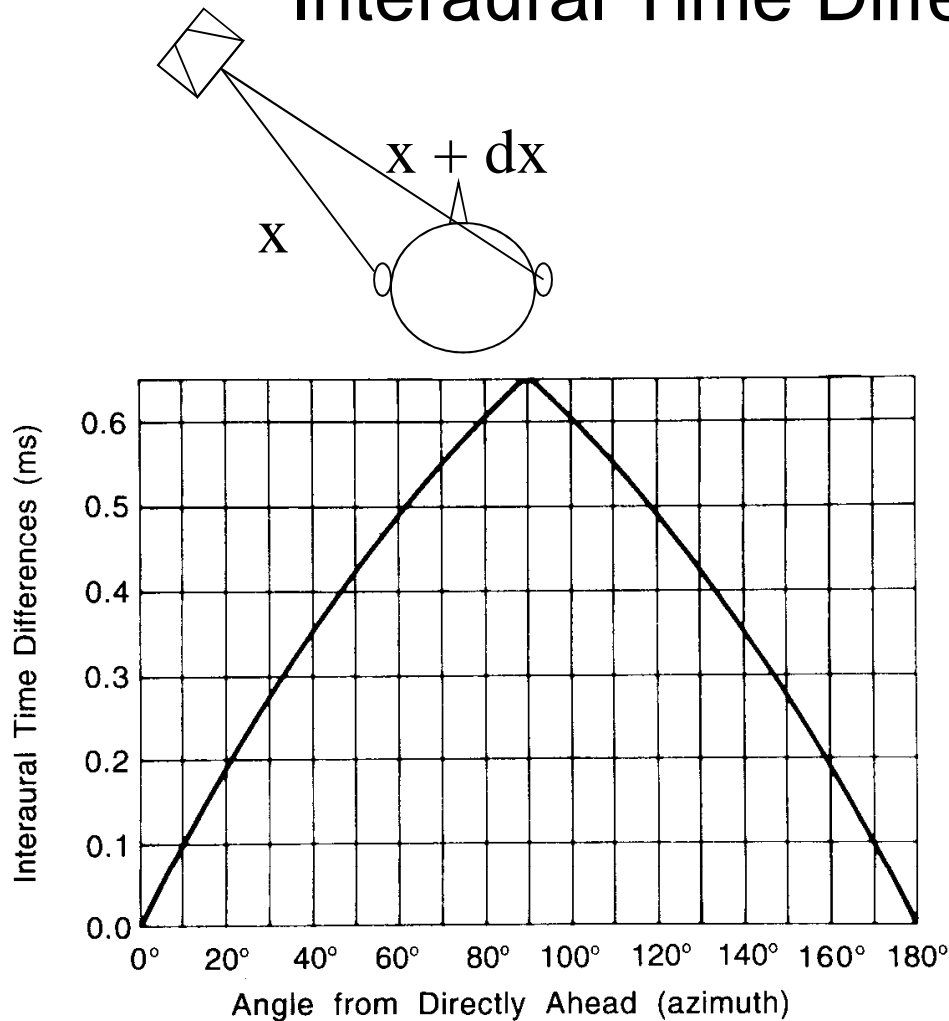


FIG. 6.4 Interaural time differences (ITDs) plotted as a function of azimuth. Adapted from Feddersen *et al.* (1957).

Rozdiel vo vzdialenosti od zdroja po ľavé a pravé ucho spôsobuje rozdiel v čase kedy zvuk dorazí k ušiam

Väčšina druhov zvierat je schopná rozpoznať veľmi malé interaurálne časové rozdiely (človek 10-50 μ s, len nízke frekvencie!!!)

ITD závisia hlavne na relatívnom rozdiel vo vzdialenosti, sú približne nemenné pre daný azimut

ITD nezávisí na frekvencii

Interaurálne rozdiely v intenzite – Interaural Level Differences (ILDs)

Rozdiely v intenzite s akou zvuk dorazí k ľavému a k pravému uchu

Ľudia sú citliví na rozdiely okolo 0.5 dB

ILDs sú **pre vzdialené zdroje** spôsobené tým, že hlava vytvára akustický „tieň“

ILD je frekvenčne závislé – rastie s frekvenciou

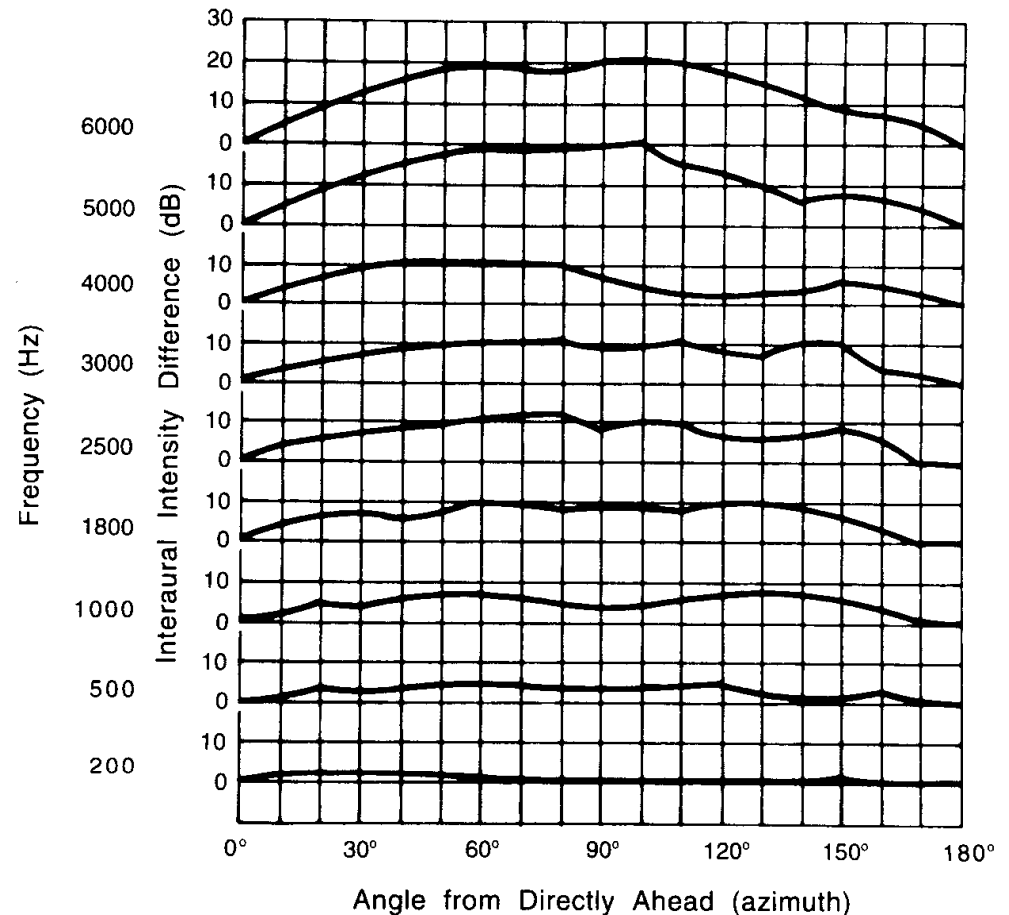
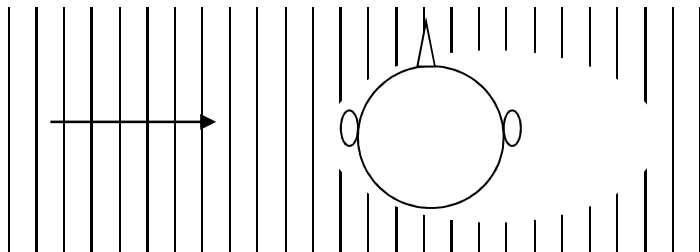


FIG. 6.2 Interaural intensity differences (ILDs) for sinusoidal stimuli plotted as a function of azimuth; each curve is for a different frequency. Adapted from Feddersen *et al.* (1957).

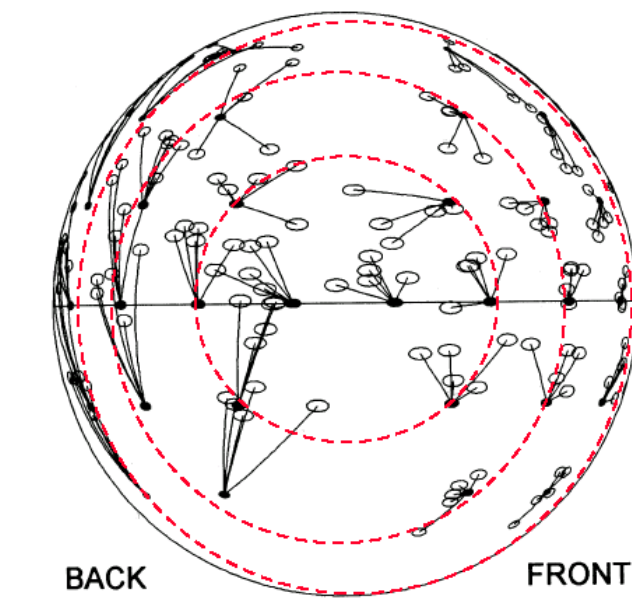
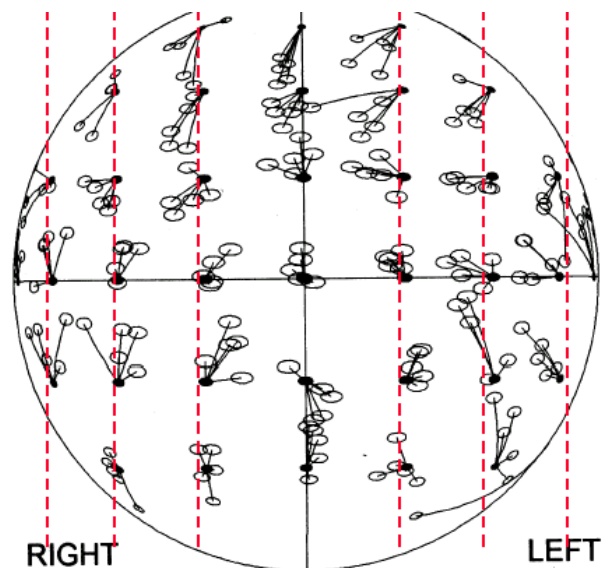
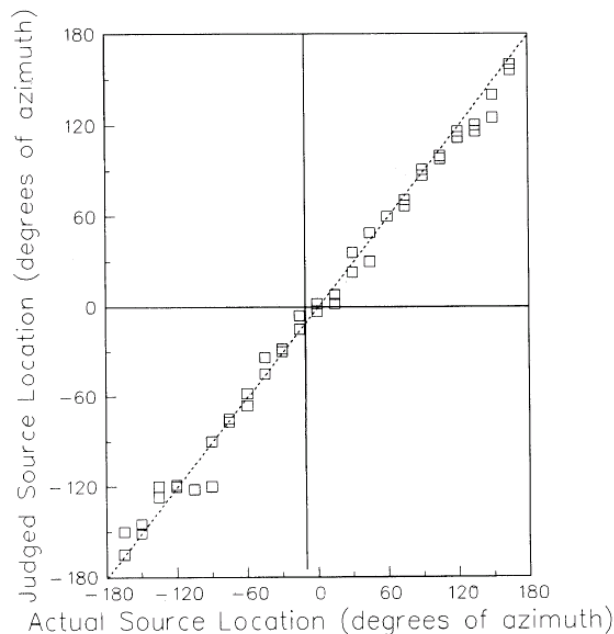
Lokalizovanie širokospektrálnych zvukov

Schopnosť lokalizovať, keď sa mení len azimut je vysoká (obr. vľavo dole)

Obr. vpravo ukazujú, aké chyby robíme, keď sa mení azimut aj elevácia (hore: pohľad spredu, dole: pohľad sprava, body zobrazujú skutočnú polohu zdroja, krúžky počutú polohu)

Pri vnímaní vľavo/vpravo chyby nerobíme, robíme ich ale v smere hore/dolu a vpredu/vzadu

Prečo?



Cone of confusion – kužel nejednoznačnosti

Chyby v smere hore/dole a
vpredu/vzadu robíme lebo:

ITD a ILD neurčujú polohu zdroja
zvuku jednoznačne

Ak by bola ľudská hlava úplne
guľatá:

- ITD a ILD na povrchu kužela by
sa nemenilo
- jediným parametrom by bol uhol
od osi spájajúcej uši

V reále je kužel trocha
deformovaný, čo je určené
skutočným tvarom hlavy (obr.
dole)

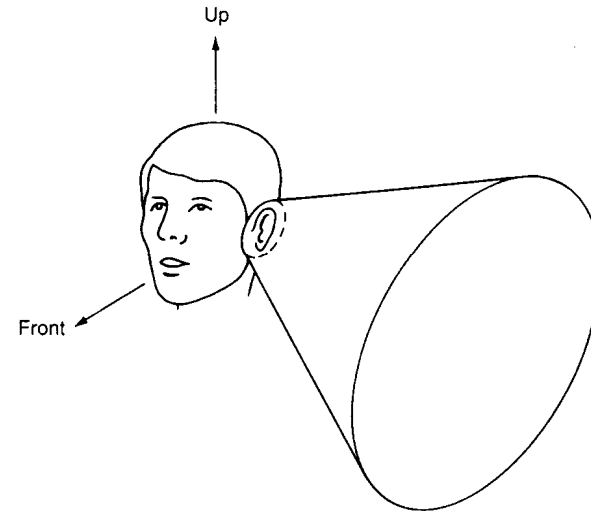
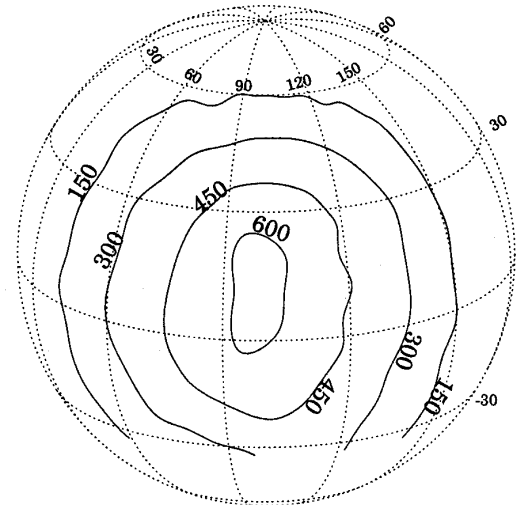


FIG. 6.7 A cone of confusion for a spherical head and a particular ITD. All sound sources on the surface of the cone would produce that interaural time delay. For details of how to calculate the cone of confusion see Mills (1972).



Čo s tým?

Lokalizácia: Vnímanie vzdialenosti

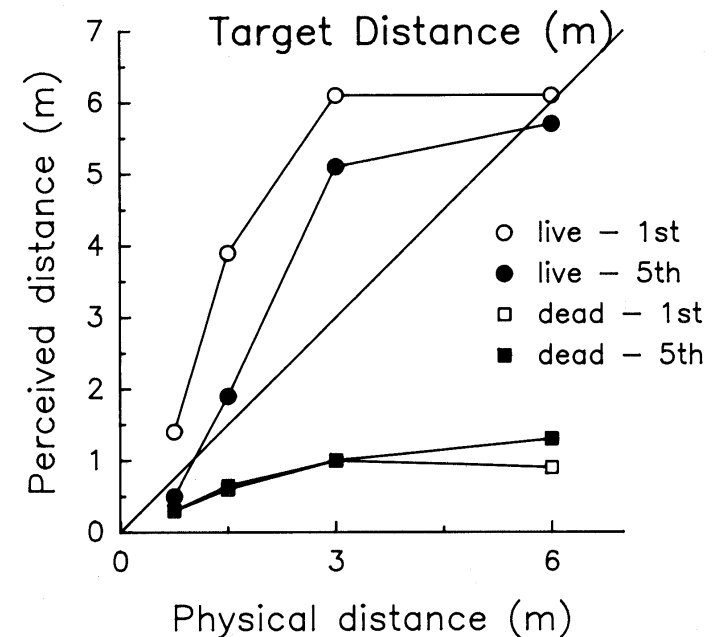
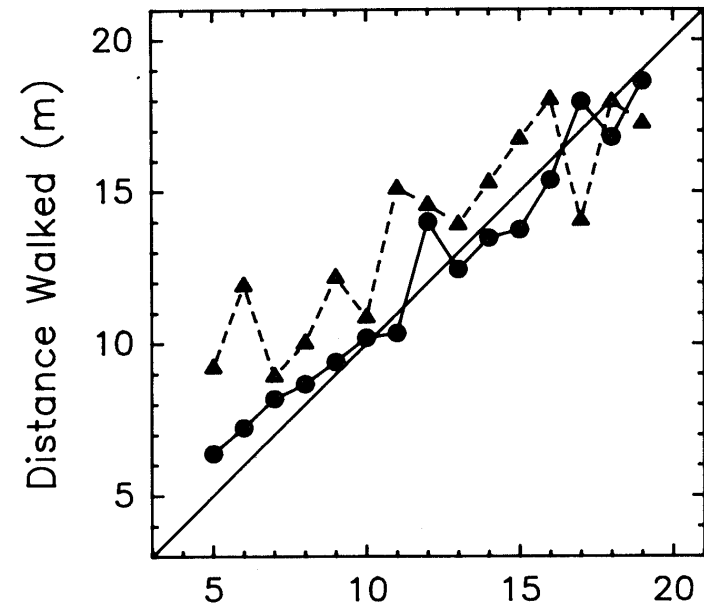
Vie sa o ňom málo

- pohyb (priblíženie sa) pomáha
- hlasitosť a „zvuk“ známych zvukov sú tiež informatívne
- intenzita odrazov od stien (je dobrým zdrojom informácií, ale je v každej miestnosti iná)

Príklad z dvoch rôznych miestností

- poslucháči v neznámej miestnosti so šatkou na očiach
- prvá odozva veľmi zlá
- už piata odpoveď je oveľa presnejšia

Odrazy od stien sú dôležitou informáciou (live = miestnosť s mnohými odrazmi, dead = miestnosť bez odrazov)



Ako je to v reálnej miestnosti?

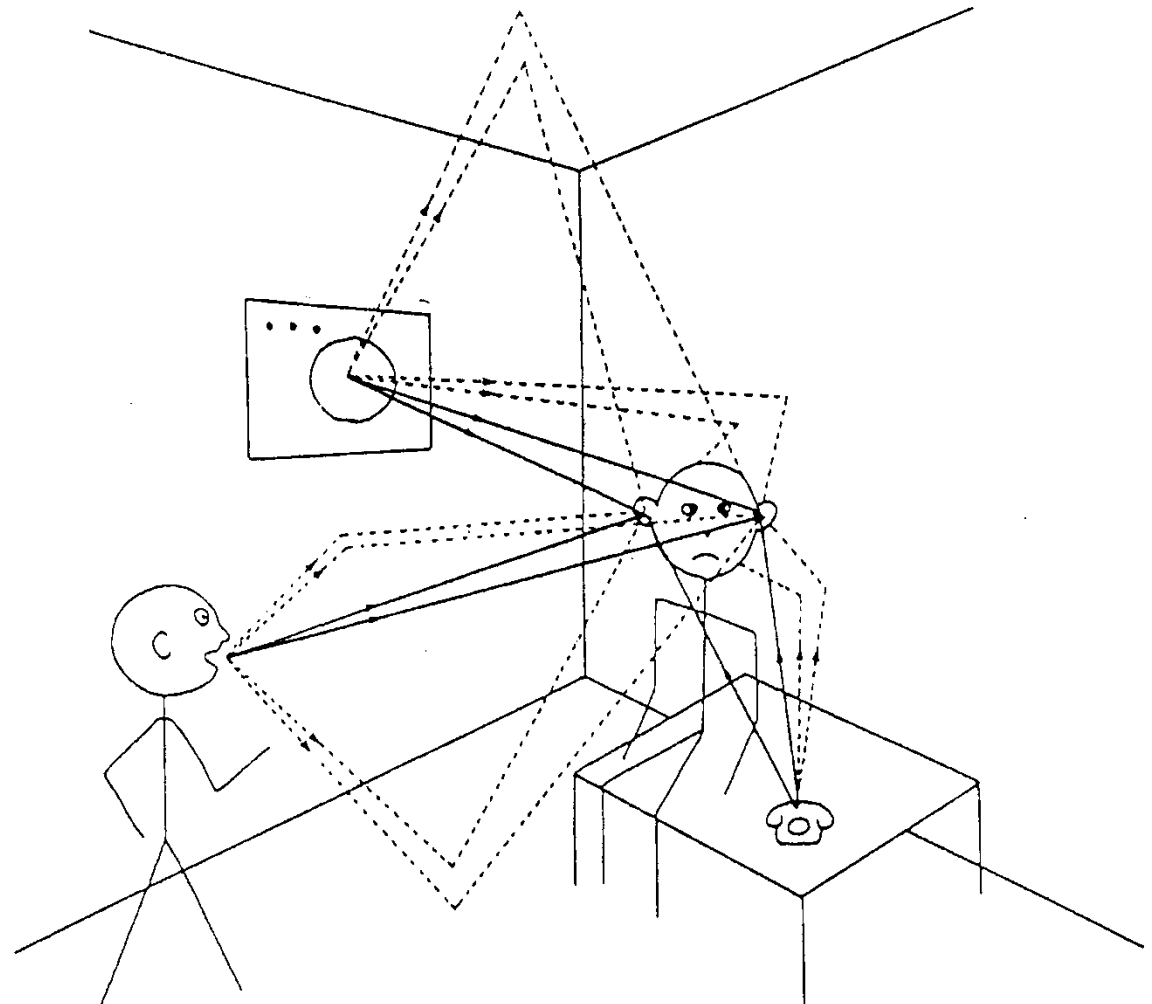
Situácia je oveľa komplikovanejšia:

Zvuk sa odráža od stien, podlahy, stropu atď.

Tento odrazený zvuk sa kombinuje so zvukom, ktorý sa šíril do uší priamo

Sluchový systém je schopný tieto odrazy potlačiť (neuvedomujeme si ich)

Naviac informácie z týchto odrazov používame na odhad vzdialenosti zdroja

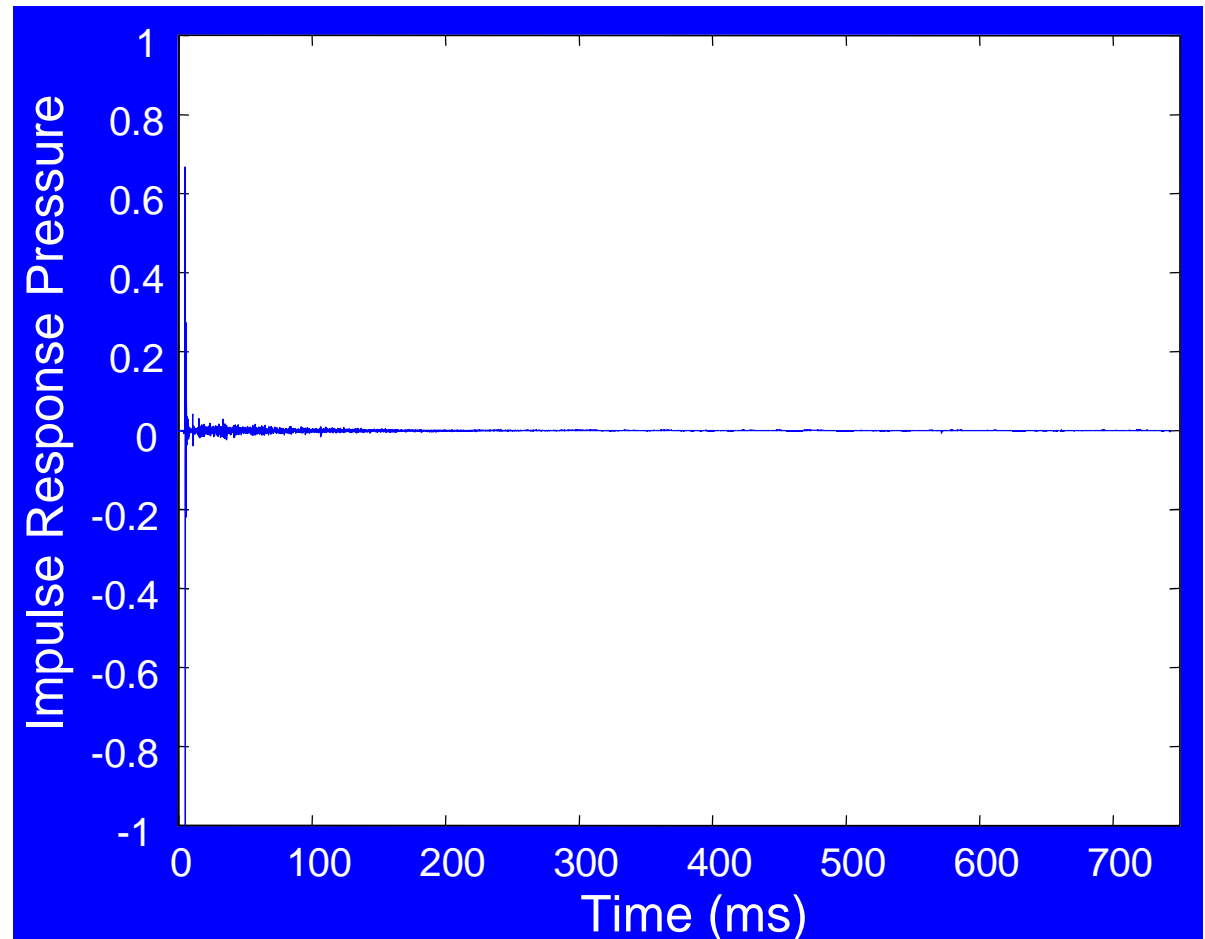


Durlach & Colburn (1978)

HRTFs nahrané v skutočnej miestnosti

Impulzná odozva (HRTF)
pre zdroj vzdialený 1m
od pravého ucha pre
človeka posadeného
v miestnosti 5x9 metrov

Odrazy od stien sa ustália
až po relatívne dlhom
čase

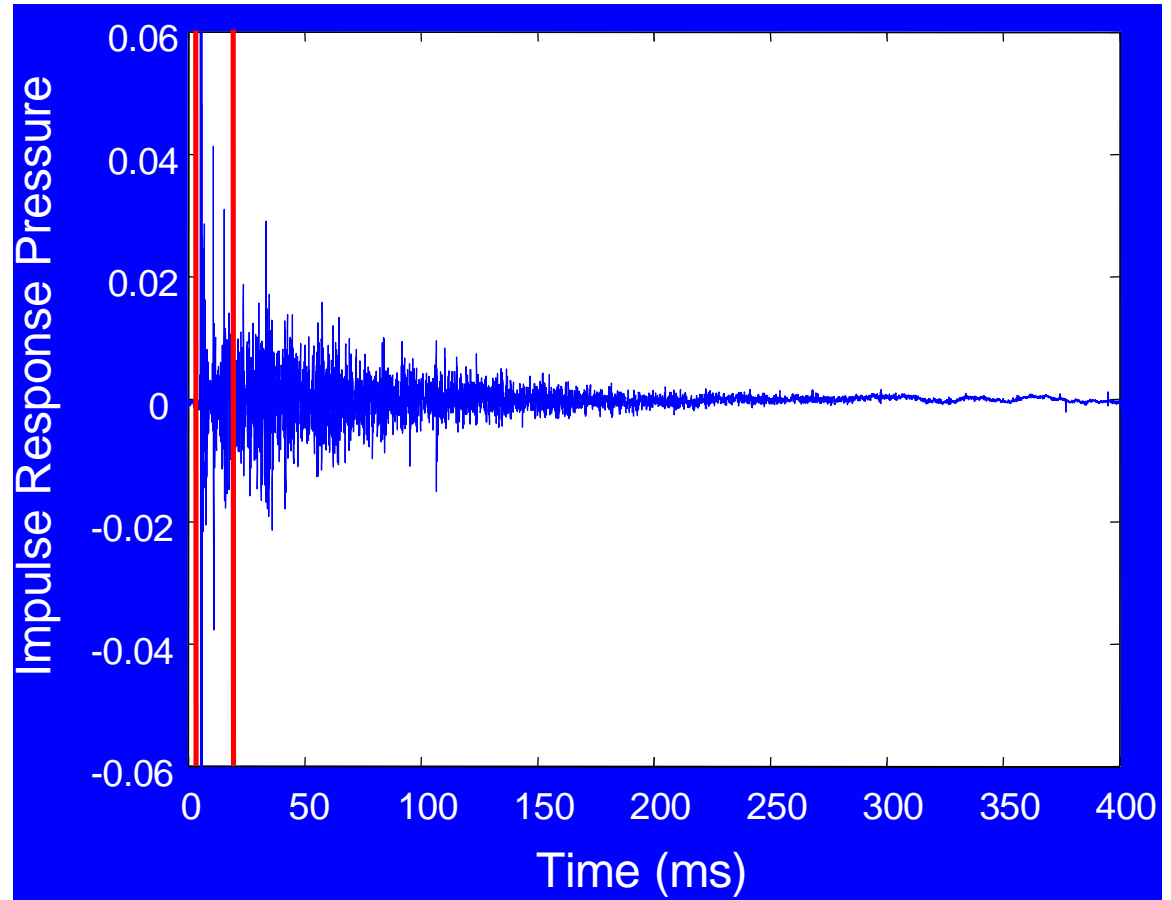


HRTFs nahrané v skutočnej miestnosti

Impulzná odozva pre zdroj vzdialený 1m od pravého ucha pre človeka posadeného v miestnosti 5x9 metrov

Odrazy od stien sa ustália až po relatívne dlhom čase

Môžeme sa pozrieť čo sa deje na začiatku



HRTFs nahrané v skutočnej miestnosti

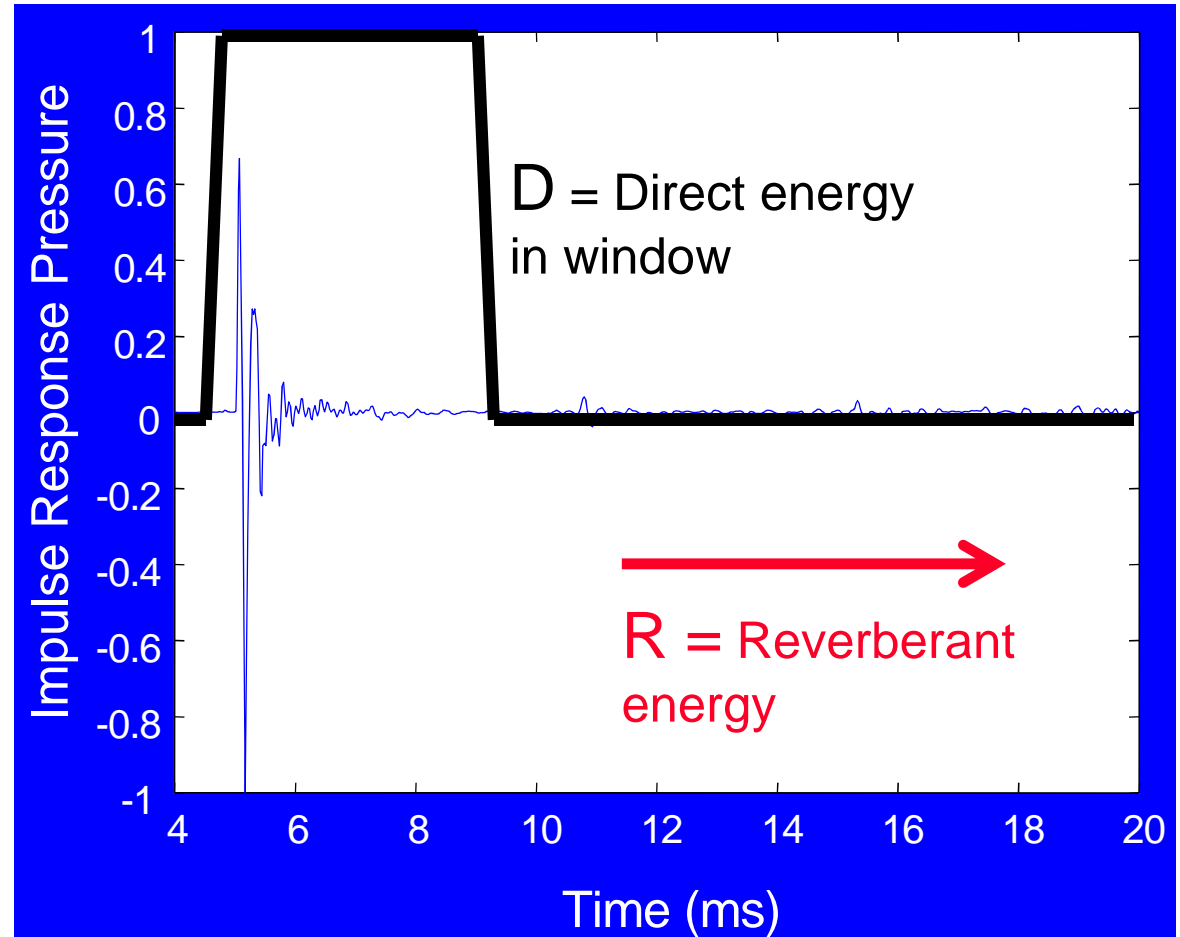
Impulzná odozva pre zdroj vzdialený 1m od pravého ucha pre človeka posadeného v miestnosti 5x9 metrov

Odrazy od stien sa ustália až po relatívne dlhom čase

Môžeme sa pozrieť čo sa deje na začiatku

Môžeme oddeliť priamy signál od odrazov

Pomer energie v D/R – používame na odhad vzdialenosti



Binaurálne rozdiely a citlivosť detekcie

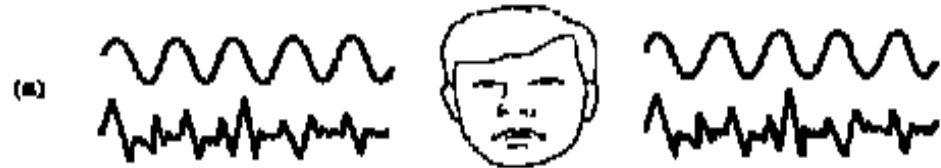
Typická paradigma pre štúdium binaurálnej detekcie:

- používame slúchadlá
- úlohou poslucháča je detekovať prítomnosť čistého tónu maskovaného šumom
- ak sa interaurálne parametre tónu a šumu líšia, tón je detekovateľný oveľa ľahšie (t.j., aj pri tichšom prehrávaní)
- nie vždy ale počujeme samotný tón, niekedy vnímame „zmenu priestorových charakteristík šumu“

Tento fenomén sa nazýva „Binaural Masking Level Difference“ (BMLD), t.j., niečo ako Binaurálny rozdiel v hladine maskovania

Binaural Masking Level Difference

N_0S_0 - tón aj šum diotický



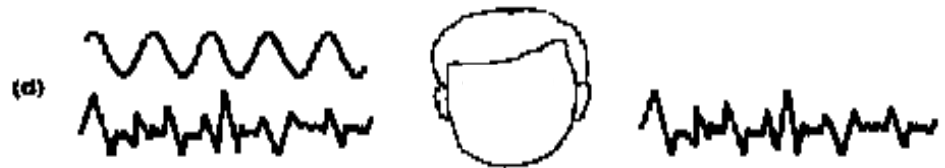
N_0S_π - šum diotický, tón dichotický



$N_M S_M$ - tón aj šum monaurálny



$N_0 S_M$ - šum diotický, tón monaurálny

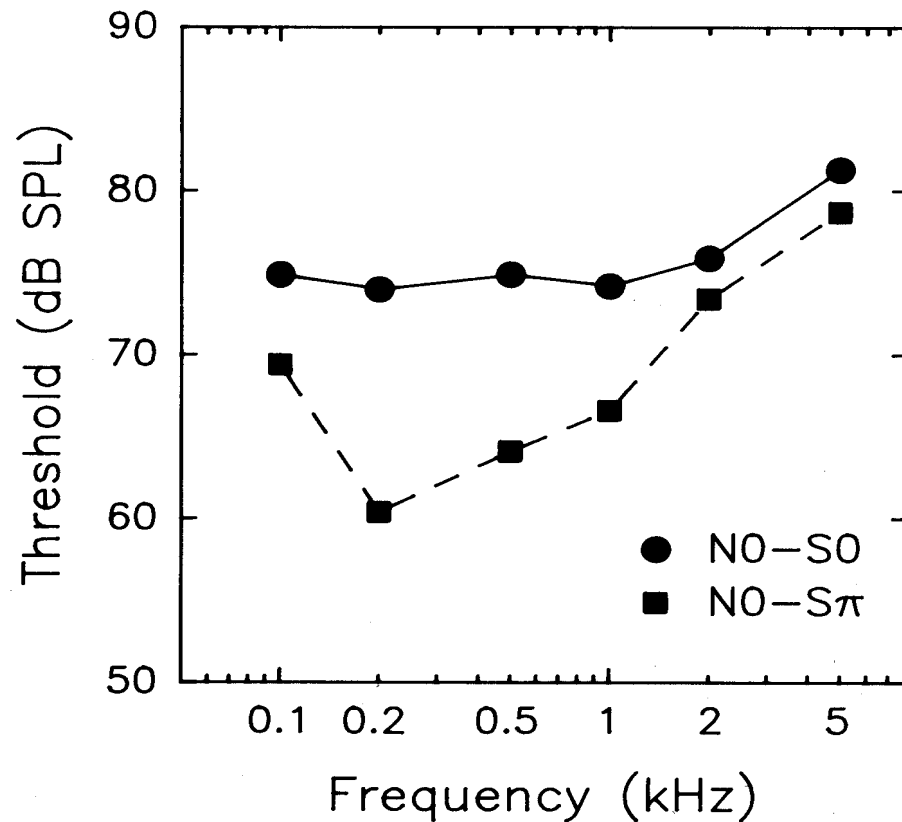


Detekcia NoSo a NoS π

BMLD je definovaná ako rozdiel medzi prahom detekcie tónu v móde N_0S_0 a N_0S_π

BMLD závisí na frekvencii tónu (?prečo?)

V reálnom prostredí (nie pri počúvaní cez slúchadlá) sa tento efekt prejavuje ako „priestorové odmaskovanie“



Priestorové odmaskovanie a koktailové večierky

Cocktail party effect:

Človek je schopný zamerať sa na počúvanie jedného konkrétneho zdroja zvuku z množstva zvukov vo svojom okolí (stroje a ľudia s poruchami priestorového sluchu to zatiaľ nevedia).

Ako to robíme?

- sčasti je tento efekt spôsobený „výhodou dvoch uší“, t.j., tým, že často je jedno z našich dvoch uší bližšie k zdroju zvuku, ktorý chceme počuť (nám potom stačí monaurálne zapnúť to správne ucho a počúvať ním)
- sčasti je spôsobený binaurálnym počúvaním (sústredenie sa na smer z ktorého zvuk prichádza)
- sčasti je to dôsledok schopnosti sluchového systému segregovať zvuky (nie na základe polohy) napr. na základe ich spektra, časovania, atď.

Model párov Sterna, Trahiotisa a Colburna

- vychádza z Jeffresovho modelu z roku 1949
- je schopný vysvetliť mnohé javy binaurálneho sluchu
- 2-rozmerné pole modelových neurónov, ktoré simulujú pole „detektorov koincidencie“ nachádzajúce sa v olivárnom komplexe mozgového kmeňa (nazývanom MSO)
- dve osi tohoto poľa sú:
 - τ (ITD, na ktoré je daný neurón citlivý)
 - f (frekvencia, na ktorú je daný neurón citlivý)
- každý neurón odpovedá približne periodicky s periódou $1/f$
- maximá v (τ, f) displeji určujú ITD počutého zvuku na danej frekvencii

Správanie sa modelu

Displej zobrazuje aktiváciu MSO neurónov

Pri úzkospektrálnom šume

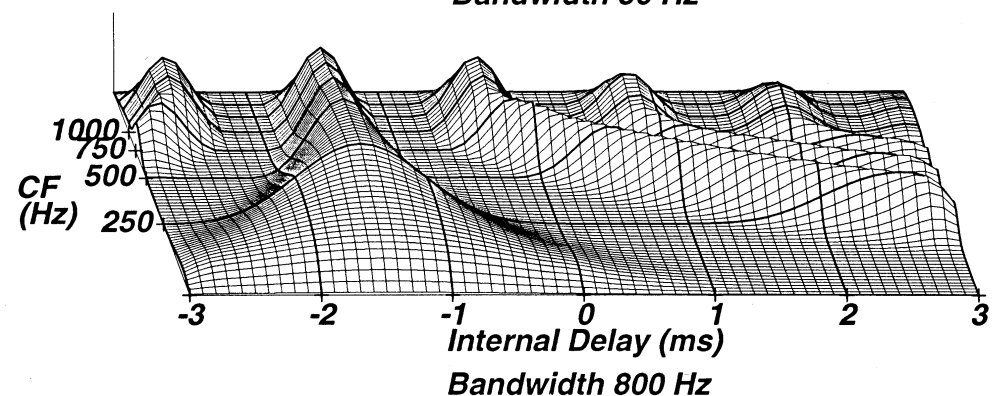
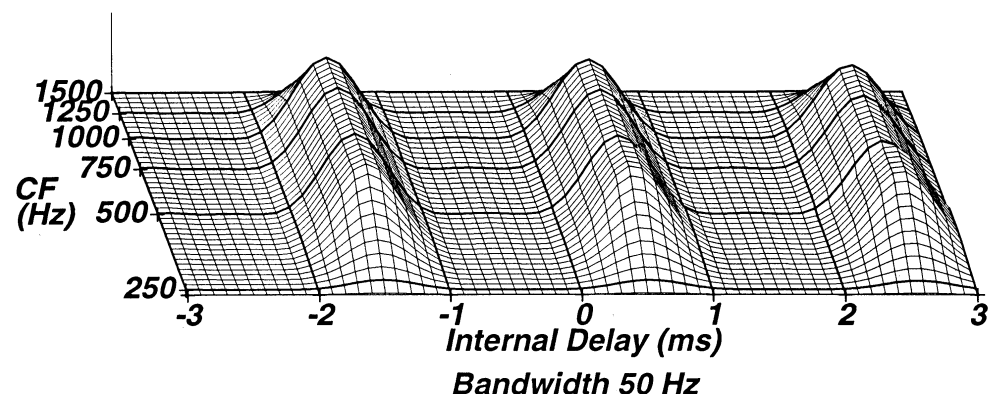
- maximá sa nachádzajú na rovnakých ITD pre všetky frekvencie
- subjekty vnímajú zvuk sprava (+ITD), nie zľava (-1.5 ms)
- princíp “centrality”

Pre širokospektrálny zvuk

- existuje len jedno ITD, ktoré je konzistentné pre všetky frekvencie,
- zvuk vnímaný zľava (správne)
- hypotéza o „vážení priamosti“ (straightness weighting)

hore: šum so šírkou pásma 50 Hz
dole: širokospektrálny šum

interaaurálny časový posun ITD=-1.5 ms



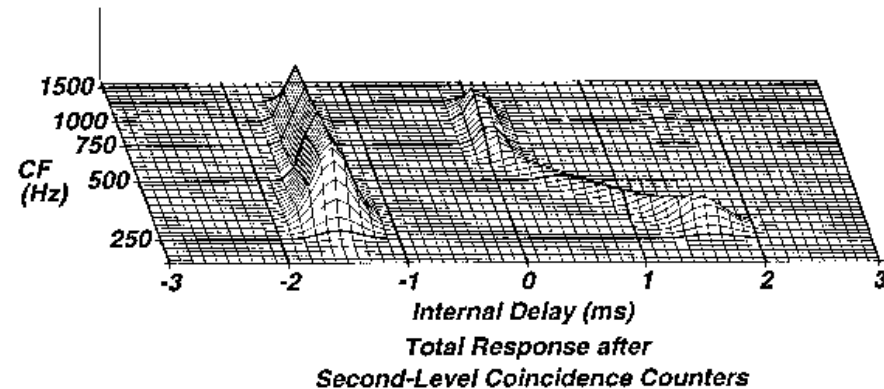
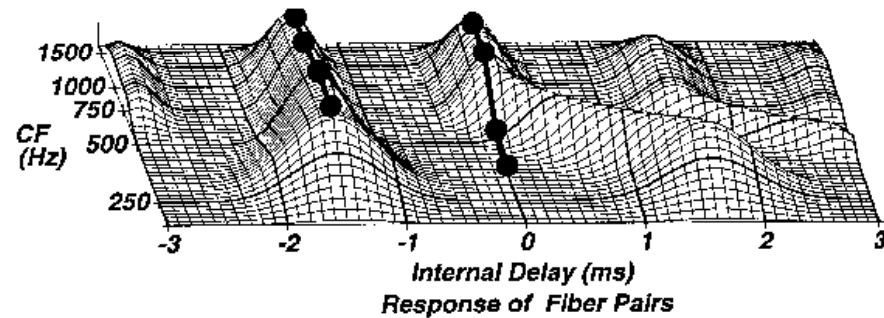
Váženie priamosti

Hypotéza:

Existuje pozitívne laterálne prepojenie medzi neurónmi s rôznymi frekvenciami ale konštantným ITD

Prepojenie medzi neurónmi s rôznym ITD je negatívne

Výsledok: Váženie priamosti je silnejšie než centralita.



Zhrnutie

- Sluchová veda
- funkcie
- lineárny systém a Fourierova transformácia
- binaurálne a priestorové vnímanie

Nabudúce

Jazyk a reč

Na budúci rok

Sluchová dráha:
anatómia, fyziológia

Sluchové vnímanie a kognícia