

TECHNICKÁ UNIVERZITA V KOŠICIACH  
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Unimodálne a krosmodálne riadenie pozornosti

Daniel HORVÁTH

DIPLOMOVÁ PRÁCA

2008

TECHNICKÁ UNIVERZITA V KOŠICIACH  
FAKULTA ELEKTROTECHNIKY A INFORMATIKY  
Katedra kybernetiky a umelej inteligencie

**Unimodálne a krosmodálne riadenie pozornosti**

DIPLOMOVÁ PRÁCA

Daniel Horváth

Vedúci diplomovej práce:

Ing. Norbert Kopčo, PhD.

Konzultant diplomovej práce:

Ing. Norbert Kopčo, PhD.

Košice 2008

## Analytický list

---

Autor:	Daniel Horváth
Názov práce:	Unimodálne a krosmodálne riadenie pozornosti
Jazyk práce:	slovenský
Typ práce:	Diplomová práca
Počet strán:	57
Akademický titul:	Inžinier
Univerzita:	Technická univerzita v Košiciach
Fakulta:	Fakulta elektrotechniky a informatiky (FEI)
Katedra:	Katedra kybernetiky a umelej inteligencie (KKUI)
Študijný odbor:	Umelá inteligencia
Študijný program:	
Mesto:	Košice
Vedúci DP:	Ing. Norbert Kopčo, PhD.
Konzultanti DP:	Ing. Norbert Kopčo, PhD.
Dátum odovzdania:	5. máj 2008
Dátum obhajoby:	
Kľúčové slová:	unimodálne a krosmodálne riadenie pozornosti, kokteil párty efekt
Kategória Konspekt:	Výpočtová technika; Umelá inteligencia
Citovanie práce:	Horváth Daniel: Unimodálne a krosmodálne riadenie pozornosti Diplomová práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2008. 57 s.
Názov práce v AJ:	Unimodal and Crossmodal attentional control
Kľúčové slová v AJ:	unimodal and crossmodal attentional control, cocktail party effect

### **Abstrakt v SJ**

Cieľom tejto práce je skúmanie toho, ako sa naša pozornosť zameriava na počúvanie reči pri viacerých hovoriacich a skúma to, ako vedomosť o polohe hovoriaceho ovplyvňuje schopnosť porozumieť mu. Konkrétne práca popisuje výsledky experimentu ktorý skúmal ako zvukové vodítko ovplyvní našu schopnosť zamerať sa na hlas. Tiež sa práca zameriava na to ako sa zvýši naša pozornosť, ak o polohe hovoriaceho sme informovaní, iba zvukovým stimulom. Pozornosť ovplyvňovala lokalizáciu zvukov a ako sa ukázalo aj porozumenie reči, ak naraz rozprávalo viacej hovoriacich v prostredí. Schopnosť porozumieť reči jedného človeka, ak súčasne hovorilo viacero ľudí, sa výrazne zvýšila, ak sa poloha hovoriaceho nemenila a subjekty vedeli, kde sa nachádza, v porovnaní s tým, ak sa jeho poloha menila. Tiež je z výsledkov vidieť, že schopnosť zamerať sa na polohu hovoriaceho bola vyššia, použitím sluchového stimulu, ak sa čas na zameranie pozornosti zvýšil.

### **Abstrakt v AJ**

The aim of this study is to examine how did your attention influences on speech reception by many speakers. The study of this diploma is focused on speech recognition in multitalker environment and examines how knowledge about talker's position influences ability to understand him/her. Ethier is this diploma focused how can grow us attention if the possition of speeker is know only by sound stimul. To summarize, attention influenced sound localization and as was showned ability to recognize speech what one person was saying when other people were simultaneously speaking in enviroment. The results schowned that ability to recognize what one person was saying when others were sumultaneously speaking significantly improved if position of the talker remained the same and subjects knew the position, contrary to the situation, when talker's position was unexpectedly changing. Also we can in results see, that ability to focus position of speeker vas significantly improved by using sound stimul and if time of that stimul was growing



# Zadanie práce

TECHNICKÁ UNIVERZITA V KOŠICIACH

Fakulta elektrotechniky a informatiky Katedra kybernetiky a umelkej inteligencie

## DIPLOMOVÁ PRÁCA

Študent: Daniel Horváth  
Študijný odbor: Umelá inteligencia  
Akademický rok: 2007/2008  
Názov práce v slovenskom a anglickom jazyku:

Sluchové riadenie pozornosti v dynamickom prostredí s viacerými hovoriacimi  
Auditory attentional control in a dynamic multi-talker environment

### Pokyny na vypracovanie:

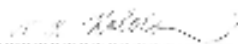
1. Vypracovať prehľad problematiky priestorového sluchuového vnímania a mechanizmov riadenia sluchovej pozornosti, ako aj relevantných modelov s dôrazom na priestorové odhliadnutie a Ušiek: koktailovej párty.
2. Implementovať experimentálnu štúdiu popísanú v článku Torralba B. and Kempe A. (2008) Auditory Spatial Cuing for Speech Perception in a Dynamic Multi-talker Environment. In: Proceedings of the 6th International Symposium on Applied Machine Intelligence and Informatics, Herlany, Slovakia, January 21 - 22, 2008.
3. Upraviť experimentálnu procedúru v prostredí MATLAB tak, aby sa pozorovať riadila sluchová Experimentálne otčíť najvhodnejší typ sluchového podnetu.
4. Vykonať experimentálne meranie na minimálne 8 dobrovoľných ľudských subjektoch.
5. Analyzovať a vyhodnotiť experimentálne dáta.
6. Navrhnuť koncepciu modelu popisujúci dynamické aspekty pozostnostného riadenia pri priestorovom odhliadnutí.
7. Vypracovať dokumentáciu podľa požiadaviek vedeckého diplomovej práce.

Vedúci diplomovej práce: Ing. Norbert Kopács, PhD.  
Konzultant diplomovej práce: Ing. Norbert Kopács, PhD.  
Dátum odovzdania diplomovej práce: 5.5.2008

  
prof. Ing. Ján Samovský, C.Sc.

vedúci zadávejúcemu  
vedecko-pedagogického pracoviska



  
Ing. Liberos Vekotokos, PhD.

dekan

V Košiciach, dňa 5.3.2008

## **Čestné vyhlásenie**

Vyhlasujem, že som celú diplomovú prácu vypracoval samostatne s použitím uvedenej odbornej literatúry.

Košice, 5. máj 2008

.....  
*vlastnoručný podpis*

## **Pod'akovanie**

Ďakujem vedúcemu mojej diplomovej práce a zároveň konzultantovi Ing. Norbertovi Kopčovi, PhD. za usmernenia, trpezivosť, pripomienky a odbornú pomoc pri vypracovaní diplomovej práce a ďakujem tiež všetkým, ktorí sa dobrovoľne zúčastnili na experimentoch.

## Predhovor

Pozornosť sa stala predmetom výskumov pred vyše polstoročím. Od tej doby bolo napísaných množstvo článkov, väčšina bola zameraná na vizuálnu pozornosť, menej z nich na sluchovú. Napríklad, ešte stále nepoznáme jednoznačné odpovede na otázky ako: Ak budeme očakávať, že zvuk príde napr. z ľavej strany, pomôže nám toto očakávanie presnejšie tento zvuk lokalizovať, ak skutočne príde zľava? A bude nám podobné očakávanie nápomocné aj pri porozumení reči v hlučnom prostredí? Názory na to, ako, a či vôbec, pozornosť pomáha pri sluchovom vnímaní, sa rôznia. Cieľom tejto práce je štúdium toho, ako pozornosť pomáha pri vnímaní reči v prostredí, kde súčasne hovorí viacero ľudí (efekt kokteilovej párty – experiment „Cocktail Party Effect“). Toto štúdium sa realizuje formou experimentov a získané dáta sú následne štatisticky vyhodnotené.

# Obsah

<b>Zoznam obrázkov</b> .....	<b>10</b>
<b>Zoznam tabuliek</b> .....	<b>12</b>
<b>Zoznam symbolov a skratiek</b> .....	<b>13</b>
<b>Slovník termínov</b> .....	<b>14</b>
<b>Úvod</b> .....	<b>15</b>
<b>1 Formulácia úlohy</b> .....	<b>17</b>
<b>2 Teoretické poznatky</b> .....	<b>18</b>
2.1 Priestorové sluchové vnímanie .....	18
2.1.1 Lokalizácia zvukov .....	18
2.1.2 Lokalizácia zvukov v horizontálnej rovine: ITD a ILD .....	19
2.1.3 Kužeľ nejednoznačnosti .....	21
2.1.4 Lokalizácia vo vertikálnej rovine .....	23
2.1.5 Smerová prenosová funkcia, tzv. Head-related transfer function (HRTF)	
2.1.6 Vnímanie vzdialenosti .....	24
2.1.7 „Precedence effect“ .....	24
2.2 Pozornosť .....	25
2.3 Multimodálne a krosmodálne vnímanie .....	27
2.4 „Cocktail party effect“ .....	28
2.5 Modely pozornosti .....	30
2.5.1 Broadbentova teória – pozornosť ako filter (1958) .....	30
2.5.2 Anne Triesman – teória zoslabovania – model skorej (alebo strednej) selektie (1964) .....	32
2.5.3 Teória Deutscha & Deutschovej – model neskorej selektie (1963) .....	33
<b>3 Experimentálna časť – vnímanie reči</b> .....	<b>34</b>
3.1 Experiment „Cocktail Party Effect“ .....	34
3.1.1 Motivácia .....	34
3.1.2 Hypotézy .....	34
3.1.3 Metódy .....	35
3.1.4 Tréning .....	38
3.1.5 Experimentálna procedúra .....	38

3.1.6 Spôsoby vyhodnocovania.....	38
3.1.7 Výsledky.....	39
3.1.8 Záver.....	49
<b>4 Záver - zhrnutie experimentu .....</b>	<b>52</b>
<b>Zoznam použitej literatúry.....</b>	<b>54</b>
<b>Prílohy .....</b>	<b>56</b>

## Zoznam obrázkov

Obr. 1	Systém súradníc používaný v sluchových experimentoch; $r$ je vzdialenosť, $\varphi$ je azimut a $\delta$ je elevácia [1].....	19
Obr. 2	ITD v závislosti od horizontálnej polohy zdroja zvuku [8]. .....	20
Obr. 3	Nízkofrekvenčné zvuky nevytvárajú akustický tieň (obrázok vľavo), vysokofrekvenčné áno (obrázok vpravo) [7]. .....	20
Obr. 4	ILD v závislosti od frekvencie zvuku pre rôzne polohy zdroja zvuku [7].....	21
Obr. 5	Kužeľ nejednoznačnosti. Body na jeho povrchu majú rovnaké ITD a ILD [1].....	21
Obr. 6	Pohyby hlavy slúžiace na rozlíšenie polohy zdroja zvuku na kuželi nejednoznačnosti. Zdroj: [21] .....	22
Obr. 7	Vertikálna lokalizácia zvuku založená na odrazoch od ušnice [8]. .....	23
Obr. 8	Precedence effect. Ak umiestnime reproduktory každý z jednej strany a prehráme zvuk v oboch súčasne, zvuk budeme vnímať tak, akoby znel v strede medzi nimi (a). Ak v jednom prehráme zvuk o niečo neskôr (0.1 – 1ms), zvuk budeme vnímať ako bližšie k reproduktoru, z ktorého zaznel prvý zvuk (b). Ak predĺžime opozdenie medzi zvukmi na 1-5 ms, zvuk budeme počuť tak, akoby prichádzal len z vedúceho reproduktora, tu sa prejavuje precedence efekt (c). Ak predĺžime opozdenie ešte viac, vnímame už 2 zvuky, každý z jedného reproduktora (d) [7]. .....	25
Obr. 9	Pozornosť ako filter. Sensorické vstupy rozdelené do jednotlivých kanálov sú najskôr spracovávané paralelne (preatentívne = „predpozornostné“ spracovanie), niektoré sú prepustené na ďalšie podrobnejšie spracovávanie a tento výber realizuje pozornosťný filter [1]. .....	31
Obr. 10	Broadbentov model. Sensorický register drží informáciu iba krátky čas a posúva ju filtru. Filter identifikuje správu, na ktorú sa zameria pozornosť pomocou fyzikálnych charakteristík (hlas, výška tónu, rýchlosť, akcent). Všetky ostatné informácie okrem tej „zameranej“ sú odstránené (odfiltrované). Správa, ktorá prešla filtrom, pokračuje do detektora, kde je ďalej spracovávaná, aby sa určil jej význam. Detektor spracuje všetky informácie, ktoré sa doň dostanú. Výstupom z detektora je krátkodobá pamäť. Informácie z nej sa po krátkom čase buď stratia, alebo prenesú do dlhodobej pamäte [12]. .....	32
Obr. 11	Model teórie zoslabovania. Filter z Broadbentovho modelu je nahradený „attenuatorom“, ktorý zoslabuje informácie, na ktoré sa nesústredíme. Attenuator analyzuje prichádzajúce správy podľa fyzikálnych charakteristík, jazyka, významu. Slovník je ako pamäť. Obsahuje	

	uložené slová, ktoré majú rôzne prahy aktivácie. Nižší prah aktivácie pre dané slovo znamená aj vyššiu citlivosť na dané slovo [12].	33
Obr. 12	Rozdelenie Experimentu Koktail party efekt na jednotlivé kroky.	36
Obr. 13	Percento správnych odpovedí v závislosti od prostredia (vyhodnocovanie podľa spôsobu 1), priemer cez ISI a subjekty, errorbary znázorňujú štandardnú chybu.	39
Obr. 14	Percento správnych odpovedí v závislosti od prostredia pre jednotlivé subjekty a celkový priemer (v pravej časti obrázku), errorbary znázorňujú štandardnú chybu.	40
Obr. 15	Percento správnych odpovedí v závislosti od ISI (priemer cez časové segmenty, pozíciu, subjekty), errorbary znázorňujú štandardnú chybu.	41
Obr. 16	Percento správnych odpovedí v závislosti od pozície (priemer cez časové segmenty, ISI a subjekty), errorbary znázorňujú štandardnú chybu.	42
Obr. 17	Percento správnych odpovedí v závislosti od segmentu prezentácie (priemer cez ISI, pozície a subjekty), errorbary znázorňujú štandardnú chybu.	43
Obr. 18	Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 5ms a subjekty), errorbary znázorňujú štandardnú chybu.	44
Obr. 19	Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 500ms a subjekty), errorbary znázorňujú štandardnú chybu.	45
Obr. 20	Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 2000ms a subjekty), errorbary znázorňujú štandardnú chybu.	46
Obr. 21	Percento správnych odpovedí v závislosti od pozície pre statický typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.	47
Obr. 22	Percento správnych odpovedí v závislosti od pozície pre synchronný typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.	48
Obr. 23	Percento správnych odpovedí v závislosti od pozície pre asynchronný typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.	49



## Zoznam tabuliek

## Zoznam symbolov a skratiek

- HRTF** **H**ead **R**elated **T**ransfer **F**unction, funkcia popisujúca transformáciu, ktorou prejde zvukový signál od zdroja po ucho. Je veľmi individuálna, keďže závisí od tvaru hlavy, uší a ramien, pretože na nich sa zvuk rôznymi odrazmi transformuje.
- ILD** **I**nteraural **L**evel **D**ifference, rozdiel v intenzite, v akej je vnímaný zvuk v jednom a v druhom uchu, spôsobený akustickým tieňom hlavy.
- ISI** **I**nter-stimulus **I**nterval, časový interval medzi dvoma stimulmi.
- ITD** **I**nteraural **T**ime **D**ifference, časový rozdiel medzi tým, kedy dorazí ten istý zvuk k jednému a k druhému uchu.
- SOA** **S**timulus **O**set **A**synchrony, opzdenie cieľového stimulu voči podnetu.

## Slovník termínov

**Cocktail Party effect** je schopnosť vybrať z množstva prichádzajúcich zvukových vstupov jeden a sústrediť sa naň, teda v množstve súčasne prebiehajúcich konverzácií vnímať, čo hovorí jeden človek.

**Unimodálne vnímanie** je ak modalita (zmysel) podvedome ovplyvní samu seba.

**Krosmodálne (intermodálne) vnímanie** je ak jedna modalita (zmysel) podvedome ovplyvňuje vnímanie inou modalitou.

**Modalita** je zmysel (zrak, sluch, atď.).

**Multimodálne vnímanie** znamená, ak je udalosť alebo objekt vnímaná viacerými zmyslami.

**Podnet** je upozorňujúci signál.

**Subjekt** je človek, ktorý sa dobrovoľne zúčastnil experimentov.

**Strategická pozornosť** je pozornosť vedome zameraná, sústredenie sa na niečo.

**Automatická pozornosť** je pozornosť vyvolaná vonkajším podnetom, bez vedomého sústredenia sa.

## Úvod

Pre štyridsiatimi rokmi, Cherry popísal "koktail party problém," [1] a navrhol, aby priestorové počúvanie bola hlavná metóda užívaná systémom na rozdeľovanie zvukových zdrojov pri viaczdrojových akustických podnetoch v priestore. Táto diplomová práca poskytuje prehľad štúdií priestorového počúvania zahrnujúce viac ako jeden zdroj zvuku alebo potenciálnych zvukových zdrojov pri pokusoch determinovať úlohu priestorového počúvania hraných v segregácii zvukových zdrojov. Množstvo dát zahŕňa len dva zvukové zdroje [17], a výsledky indikujú, že priestorové počúvanie nemá byť hlavné vodítko pre segregáciu zvukových zdrojov. Existuje zopár štúdií, ktoré vyšetrujú cocktail party problém pri podmienkach reálneho - prostredia, obzvlášť ak máme viac než dva zvukové zdroje. Pozornosť zohráva dôležitú úlohu pri spracovávaní informácií, ktoré vnímame pomocou zmyslov. Tie podnety, na ktoré bola upriamená pozornosť, sú spracovávané dôkladnejšie ako ostatné [1]. Pri vizuálnom vnímaní je to ľahko predstaviteľné, sledovaný predmet sa dostáva do miesta najostrejšieho videnia [2] a okolie sa vníma len okrajovo. Ako je to však pri sluchovom vnímaní? Ak presunieme pozornosť na nejaký zvukový objekt, pomôže nám to lepšie ho vnímať? Alebo - ak presunieme pozornosť na určité miesto, pomôže nám to lepšie vnímať zvuky odtiaľ prichádzajúce (napríklad tak, že ich budeme vedieť presnejšie lokalizovať)? Aký bude rozdiel v tom, či pozornosť na toto miesto privedieme zrakom alebo sluchom? Vplyv pozornosti na sluchové vnímanie nie je doteraz presne objasnený.

Diplomová práca nadvzuje na výsledky experimentov [19][20] a zaoberá sa tým, ako pozornosť pomáha pri sluchovom vnímaní. Je zameraná na tzv. „cocktail party effect“ – t.j. schopnosť porozumieť reči jedného človeka v prostredí, kde do uší prichádza množstvo zvukových vstupov (konverzácie iných ľudí, hluk z ulice, ...) [3]. Koktejl párty efekt sa skúma už vyše 50 rokov a za ten čas sa prišlo na to, aké faktory ovplyvňujú schopnosť separovať jeden „prúd reči“ od iných, aby sa potom dal dôkladnejšie spracovať a aby bol lepšie zrozumiteľný. Medzi tieto faktory patrí napríklad hlas hovoriaceho, pretože rozdielne hlasy sa ľahšie rozlišujú, rôzna rýchlosť reči hovoriacich, a ďalšie [1]. Jedným z najdôležitejších je aj priestorové oddelenie hlasov. Táto diplomová práca sa zaoberá tým, ako vedomosť o tom, odkiaľ bude prichádzať hlas hovoriaceho, ovplyvní schopnosť porozumieť mu, v porovnaní s tým, ak sa jeho poloha bude počas reči náhodne meniť.

Diplomová práca pozostáva zo 4 kapitol:

- **Kapitola 1** – formulácia úlohy
- **Kapitola 2** – teoretické poznatky - priestorové sluchové vnímanie, pozornosť, unimodálne, multimodálne a krosmodálne vnímanie, cocktail party efekt, modely pozornosti
- **Kapitola 3** – experimentálna časť – vnímanie reči – táto kapitola popisuje experiment „Cocktail Party Effect“
- **Kapitola 4** – záver – zhrnutie všetkých experimentov

## 1 Formulácia úlohy

Experimentálne procedúry ako aj analýzy získaných dát sú naprogramované v prostredí MATLAB 6.5.

Vo formulácii ulohy je zahrnuté vypracovať prehľad problematiky priestorového a sluchového vnímania a mechanizmov riadenia sluchovej pozornosti ,ako aj relevantných modelov s dôrazom na priestorové odmaskovanie a Efekt koktail párty.

Implementovať experimentálnu štúdiu popísanú v [19] .

Upraviť experimentálnu procedúru v prostredí MATLAB tak, aby sa pozornosť riadila sluchom. Experimentálne určiť najvhodnejší typ sluchového podnetu.

Vykonať experimentálne meranie na minimálne 8 dobrovoľných ľudských subjektoch.

Analyzovať a vyhodnotiť experimentálne dáta.

Navrhnuť konceptuálny model popisujúci dynamické aspekty pozornostného riadenia pri priestorovom počúvaní.

Vypracovať dokumentáciu podľa pokynov vedúceho diplomovej práce.

## 2 Teoretické poznatky

### 2.1 Priestorové sluchové vnímanie

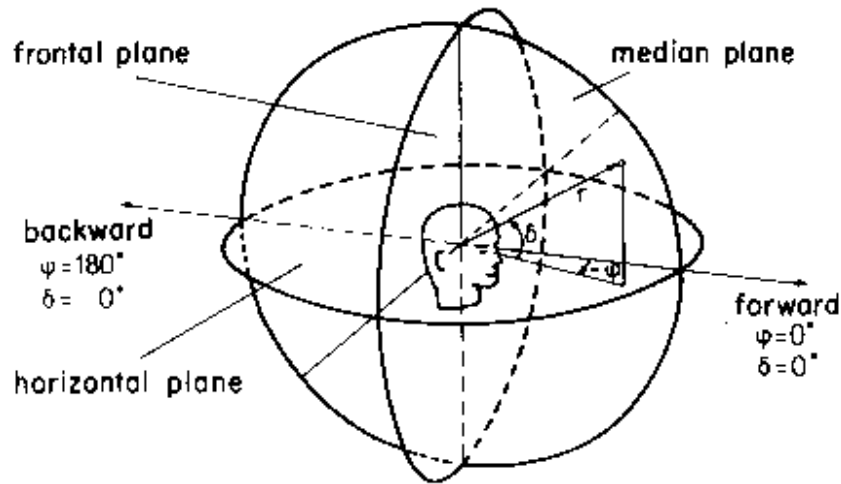
#### 2.1.1 Lokalizácia zvukov

Mechanizmy lokalizácie zvukov sú pre živočíchov životne dôležité. Rýchla a presná lokalizácia predátora a okamžitý útek tým správnym smerom im častokrát zachráni život. Lokalizácia zvukov je preto užitočná ako určitý výstražný mechanizmus [6].

Sluchové vnímanie priestoru je v určitom zmysle zložitejšie ako zrakové, ktoré je na to vhodnejšie vybavené. Rôzne body vo vizuálnom poli sa totiž priamo premietajú na rôzne miesta na sietnici. Sluchový systém však musí polohu zdroja zvuku “vypočítať”, pretože bazilárna membrána v slimákovi (miesto, kde sa vibrácie spôsobené zvukmi menia na nervové impulzy) nemá žiadnu reprezentáciu pozície zvuku [7]. Kokteil party problém sa stal odporúčaným sluchovým javom. V nedávnych rokoch bol koktail party problém delegovaný ako lokalizácia zdroja zvuku. [18]

Poloha zdroja zvuku je často určovaná troma súradnicami:

- Azimut – horizontálna odchýlka od nejakého referenčného bodu, ktorým je zvyčajne predná časť tváre.
- Elevácia – vertikálna odchýlka od nejakého referenčného bodu, väčšinou prednej časti tváre
- Vzdialenosť



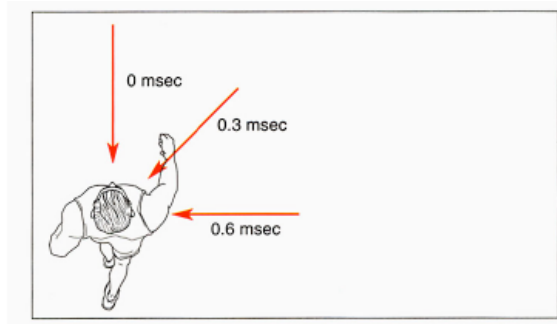
Obr. 1 Systém súradníc používaný v sluchových experimentoch;  $r$  je vzdialenosť,  $\varphi$  je azimut a  $\delta$  je elevácia [1].

Lokalizácia zvukov v horizontálnej a vo vertikálnej rovine je rôzna, na každú totiž využívame rozdielne mechanizmy. Zvuk pohybujúci sa vertikálne vieme ľahko lokalizovať aj pomocou jedného ucha, zvuk pohybujúci sa horizontálne lokalizujeme kombinovaním zvukov prichádzajúcich do oboch uší [8]. Mozog tieto zvuky vyhodnotí a určí rozdiely, pričom podstatný je rozdiel v čase, v ktorom prišiel zvuk do jedného a do druhého ucha (ITD = interaural time difference) a taktiež rozdiel v intenzite (ILD = interaural level difference) medzi zvukom v jednom a v druhom uchu.

### 2.1.2 Lokalizácia zvukov v horizontálnej rovine: ITD a ILD

Zvuk dorazí do ucha, ktoré je bližšie k zdroju zvuku, o niečo skôr ako k druhému, pretože prechádza kratšiu dráhu. Časový rozdiel sa pohybuje v desatinách milisekúnd. Ak by bola šírka hlavy 20 cm, zvuk prichádzajúci kolmo sprava dosiahne ľavé ucho o 0.6 ms neskôr ako pravé [8]. Pre zvuky prichádzajúce z boku je ITD maximálny. Ak zvuk prichádza priamo spredu, ITD je nulový (minimálny) [1].

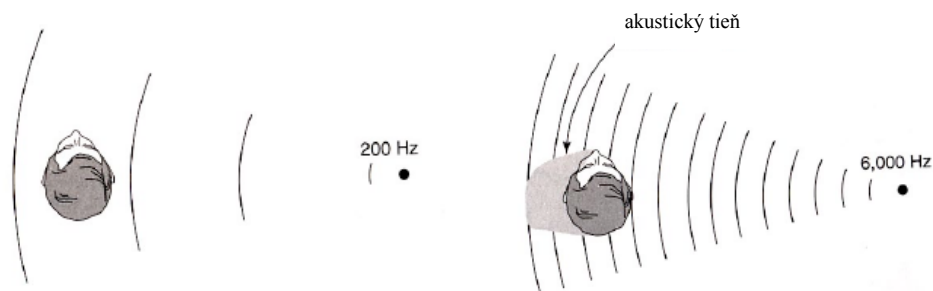




Obr. 2 ITD v závislosti od horizontálnej polohy zdroja zvuku [8].

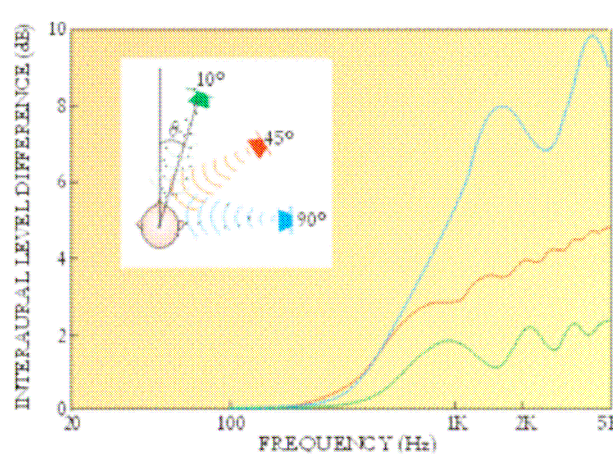
V mozgovom kmeni sú špecializované neuróny citlivé na rôzne ITD a pomocou nich dokážeme určiť, kde bol zdroj daného zvuku. ITD je pre danú polohu zdroja zvuku rovnaký, nezávisí od výšky tónu (frekvencie) [1]. V prípade, že nepočujeme nástup zvuku, je určenie ITD zložitejším problémom, pretože tento zvuk počujeme naraz v oboch ušiach. Môžeme však porovnávať rozdiel v tom, kedy rovnaká fáza zvukovej vlny dôjde do ktorého ucha [8]. V prípade nízkofrekvenčných zvukov, ktoré majú väčšiu vlnovú dĺžku ako je šírka hlavy, je to relatívne jednoduché, keď sa však jedná o vysokofrekvenčné zvuky, situácia je zložitejšia.

Pri lokalizácii zvukov si preto mozog pomáha aj určením ILD, čiže určením rozdielu v intenzitách. Rozdiel medzi intenzitami v akých zvuk počujeme v jednom a v druhom uchu je spôsobený tým, že hlava vrhá zvukový tieň [8]. Ak zvuk znie priamo spredu, rozdiel v intenzitách je nulový, ak kolmo z boku, tak maximálny. Špecializované neuróny týmto spôsobom určia polohu zdroja zvuku. Pre zdroje zvukov vzdialené minimálne 1m má určovanie ILD význam iba pri vysokofrekvenčných zvukoch, pretože zvuky nízkych frekvencií nevytvárajú zvukový tieň (Obr. 3). ILD je teda na rozdiel od ITD frekvenčne závislé, rastie s rastúcou frekvenciou, čo je vidieť aj na Obr. 4.



Obr. 3 Nízkofrekvenčné zvuky nevytvárajú akustický tieň (obrázok vľavo), vysokofrekvenčné áno (obrázok vpravo) [7].

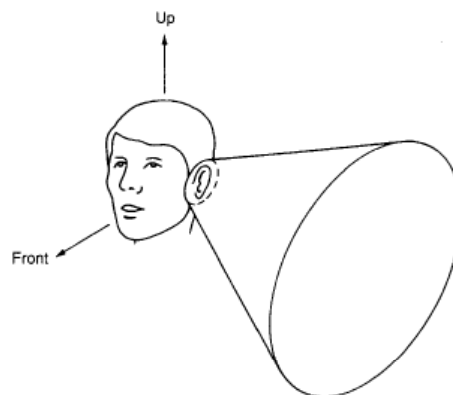
Môžeme zhrnúť, že zvuky v rozsahu približne 20-2000 Hz lokalizujeme pomocou časových rozdielov (ITD) a zvuky z intervalu 2000-20 000 Hz pomocou rozdielov v intenzite (ILD) [8].



Obr. 4 ILD v závislosti od frekvencie zvuku pre rôzne polohy zdroja zvuku [7].

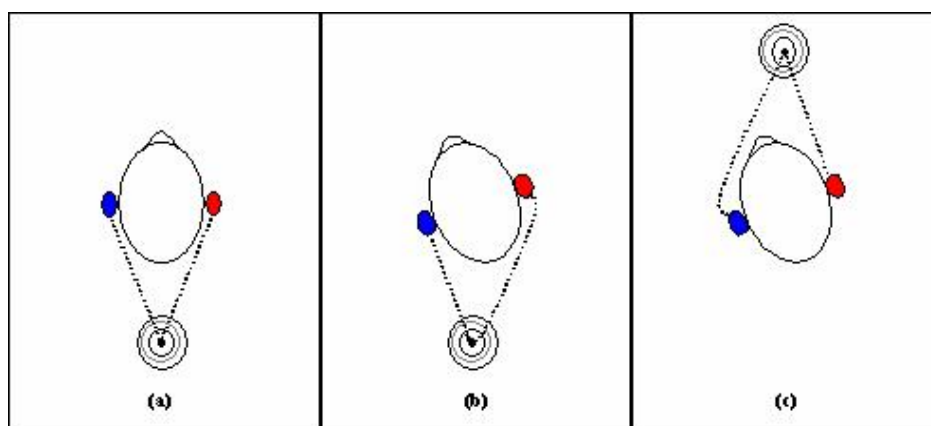
### 2.1.3 Kužel nejednoznačnosti

Aj keď zohrávajú pri lokalizácii zvukov ITD a ILD tú najdôležitejšiu úlohu, nemôžeme sa na ne úplne spoľahnúť. Problém totiž je, že neurčujú polohu zvuku úplne jednoznačne. Jednému ITD aj ILD odpovedá viacero bodov v priestore. Tieto body vytvárajú povrch kužeľa (ilustrovaného na Obr. 5), ktorý sa preto nazýva kuželom nejednoznačnosti. V praxi to znamená, že síce uhádneme uhol, odkiaľ znie zvuk, avšak často neuhádneme to, či prichádza spredu alebo zozadu, zhora alebo zdola [1]. V skutočnosti je tento kužel trochu deformovaný, čo je určené tvarom hlavy.



Obr. 5 Kužel nejednoznačnosti. Body na jeho povrchu majú rovnaké ITD a ILD [1].

Keďže ITD a ILD nie sú na určenie zdroja zvuku postačujúce, využívame aj iné spôsoby. Pomáha napríklad pootočenie alebo naklonenie hlavy, čím sa kužeľ posunie a nejednoznačnosť v smere hore-dolu, vpredu-vzadu, sa vyrieši, čo je ilustrované na Obr. 6 - ak by bol zdroj zvuku v azimute  $180^\circ$  (t.j. za hlavou), bolo by ťažké rozhodnúť, či zvuk ide z  $0^\circ$  alebo zo  $180^\circ$ , kvôli tomu, že obidve pozície majú rovnaké ITD aj ILD. Pootočením hlavy doľava sa ľavé ucho dostalo bližšie k zdroju a preto je jasné, že zvuk šiel zozadu (b). Ak by šiel spredu, bližšie by bolo pravé ucho (c).

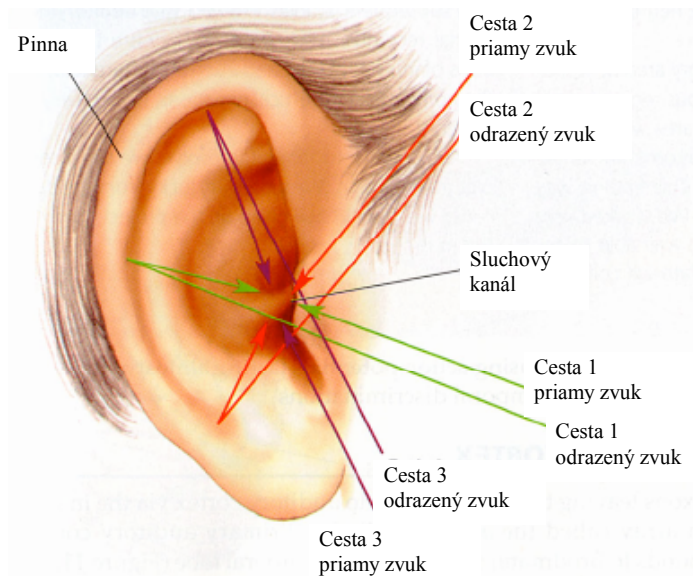


**Obr. 6** Pohyby hlavy slúžiace na rozlíšenie polohy zdroja zvuku na kuželi nejednoznačnosti.

**Zdroj:** [21]

Hlavnou pomôckou pri lokalizácii okrem ITD a ILD je však posúdenie zmien v spektrálnej charakteristike zvuku, ktoré vznikli odrazmi zvukových vln od hlavy a uší.

### 2.1.4 Lokalizácia vo vertikálnej rovine



Obr. 7 Vertikálna lokalizácia zvuku založená na odrazoch od ušnice [8].

Pri lokalizácii zvukov vo vertikálnej rovine nám rozdiely v intenzitách a časoch nepomôžu. Kľúčovú úlohu v tomto prípade zohráva ušnica („pinna“). Zvuky, ktoré k nej prichádzajú, sa na jej záhyboch rôznym spôsobom odrážajú a odrazený zvuk vstupuje do sluchového kanála s určitým omeškaním v závislosti od jeho polohy vo vertikálnej rovine [8]. Priamy zvuk s odrazeným sa skombinuje, a aj keď len nepatrne, je iný, ak zvuk prichádzal zdola, a iný ak zhora. Zvuk je teda určitým spôsobom filtrovaný.

### 2.1.5 Smerová prenosová funkcia, tzv. Head-related transfer function (HRTF)

Zvuky sa odražením netransformujú len na ušnici, ale podobný a nezanedbateľný vplyv má aj tvar pliec a hlavy. Transformáciu, ktorou zvuk prejde od zdroja až po bubienok je možné popísať pomocou prenosovej funkcie, tzv. HRTF (head-related transfer function). Je v nej zahrnuté ITD, ILD a zmeny v spektre spôsobené odrazmi zvukov. Pre zdroje vzdialené viac ako 1m od hlavy HRTF závisí od azimutu a elevácie, pre bližšie zdroje závisí aj od vzdialenosti [1].

Každému bodu v priestore zodpovedá dvojica HRTF, jedna pre pravé, druhá pre ľavé ucho.

Meráním zistené HRTF sa používajú na simuláciu zvukov prichádzajúcich z rôznych pozícií [6]. Keďže HRTF závisí od tvaru hlavy a uší, je individuálna. Ak sa potom pri simulácii používa HRTF nameraná na inom človeku, vnem nemusí byť taký, ako sa očakáva.

### 2.1.6 Vnímanie vzdialenosti

Vzdialenosť zdroja zvuku určujeme viacerými spôsobmi [7]. Pomáha nám pri tom napr. ILD – rozdiely v intenzitách budú väčšie, ak je zdroj zvuku bližšie, a menšie, ak je ďalej. Do vzdialenosti vystretej ruky dáva ILD celkom dobrý odhad vzdialenosti zdroja zvuku, za touto vzdialenosťou však informuje len o tom, že zdroj zvuku nie je na dosah ruky.

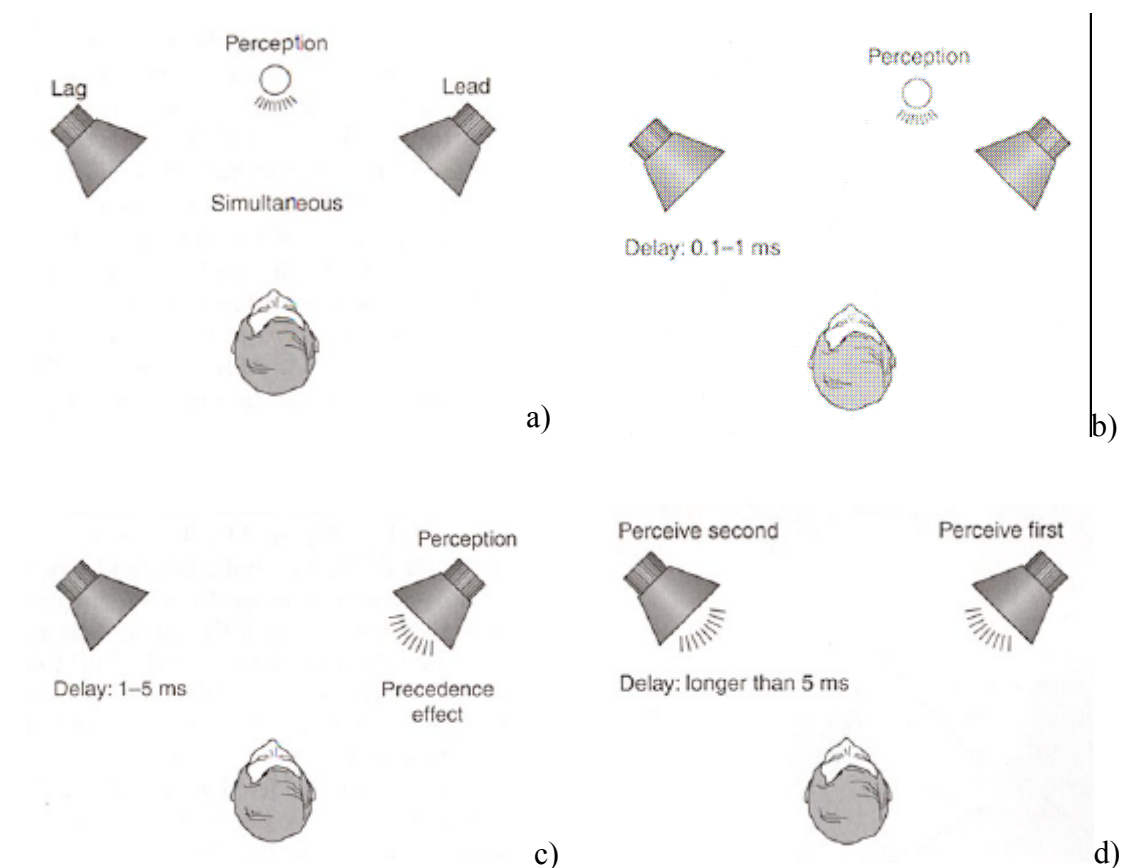
Ďalšou pomôckou je celková hladina zvuku. Niektoré zvuky zvyknú mať charakteristické intenzity – napr. reč sa pohybuje v hodnotách 40-70dB. Ak poznáme bežnú intenzitu nejakého zvuku, vieme na základe aktuálnej intenzity odhadnúť, ako je ďaleko.

Pomáha nám aj pohybový paralax – t.j. ak sa hýbeme smerom do strany, zdroj zvuku, ktorý je bližšie, sa voči nám posúva rýchlejšie ako ten, ktorý je ďalej.

Vzdialenosť zdrojov zvuku určujeme aj na základe jeho odrazov, hlavne v miestnostiach. Odrazený zvuk sa skombinuje s pôvodným, tieto odrazy však potláčame a neuvedomujeme si ich.

### 2.1.7 „Precedence effect“

Odrazy môžu pomôcť pri určení vzdialenosti zdroja zvuku, môžu však spôsobiť aj problémy v lokalizácii zvuku, odrazený zvuk totiž väčšinou prichádza z iného smeru ako priamy. Sluchový systém však dokáže rozlíšiť priamy zvuk od odrazeného na základe toho, kedy dorazí k uchu. Odrazenému zvuku to bude stále trvať dlhšie, pretože prechádza dlhšiu dráhu. Sluchový systém preto pri lokalizácii berie do úvahy len prvý zvuk (priamy) a druhý potlačí. Toto však funguje iba pri určitom opozdení jedného zvuku voči inému, konkrétne pri 1-5ms [7]. Ilustruje to Obr. 8.



**Obr. 8** Precedence effect. Ak umiestnime reproduktory každý z jednej strany a prehráme zvuk v oboch súčasne, zvuk budeme vnímať tak, akoby znel v strede medzi nimi (a). Ak v jednom prehráme zvuk o niečo neskôr (0.1 – 1ms), zvuk budeme vnímať ako bližšie k reproduktoru, z ktorého zaznel prvý zvuk (b). Ak predĺžime opozdenie medzi zvukmi na 1-5 ms, zvuk budeme počuť tak, akoby prichádzal len z vedúceho reproduktora, tu sa prejavuje precedence efekt (c). Ak predĺžime opozdenie ešte viac, vnímame už 2 zvuky, každý z jedného reproduktora (d) [7].

## 2.2 Pozornosť

Náš sensorický systém je vystavený veľkému množstvu podnetov. Zaplavujú ho obrazy, zvuky, vône, ... Nie všetky však vnímame rovnako detailne. Niektoré sú pre nás dôležité, niektoré nie. Keďže ani nie je možné, aby mozog súčasne spracoval také množstvo podnetov, musí existovať mechanizmus, ktorý niektoré z nich vyberie, a týmto mechanizmom je pozornosť [2]. Napriek tomu, že je pojem „pozornosť“ intuitívne jasný, to, ako ho zadefinovať, jasné nie je. Zahŕňa totiž množstvo rôznych aspektov a preto existuje aj množstvo definícií pozornosti, z ktorých spomeniem dve.

Pozornosť je [1]:

- mentálne koncentrovanie (sústredenie sa) na senzorické (vnemové) alebo mentálne udalosti (Dana Murphy)
- aktivácia mechanizmov, ktoré umožňujú neprerušenu kognitívnu aktivitu zameranú na objekt pozornosti (David Sommers)

Nie je jasné ani to, či pozornosť sama o sebe vôbec existuje, je možné, že je len súčasťou tých mozgových funkcií, v ktorých ju pozorujeme [1]. Má teda zmysel sa pozornosťou ako takou vôbec zaoberať?

Tí, ktorí sa rozhodli, že áno, vo svojich doterajších výskumoch objasnili niektoré jej vlastnosti:

Pozornosť je **obmedzená**, to znamená, že ju nie je možné venovať všetkým podnetom naraz, je **selektívna**, to znamená, že z množstva podnetov sa vyberú niektoré, ktoré budú ďalej spracovávané, pričom nemusí byť zameraná len na senzorické podnety, ale aj na informácie v pamäti a motorické odozvy [1]. Veľakrát sa v literatúre uvádza metafora pozornosti ako pódiového reflektora. To, čo je osvetlené, detailne spracujeme, to, čo nie je, ignorujeme.

Ak hovoríme, že niečomu venujeme pozornosť, zvyčajne tým myslíme to, že ju niekam vedome zameriavame. Je to cielený proces, preto hovoríme o tzv. **strategickej (endogénnej) pozornosti**. Niekedy sa však stane, že nás nejaký podnet zrazu vyruší (napríklad náhly hlasný zvuk) a bez vedomej kontroly pritiahne na seba pozornosť. To je tzv. **automatická (exogénna) pozornosť**.

Pozornosť môžeme rozdeliť aj na základe jej 3 komponentov: selekcie, sledovania a riadenia [1].

Pri **selekcii** sa vynára otázka, na základe čoho pozornosť vyberá. Či sa pozornosť zameriava len na určitú oblasť v priestore, alebo na objekty, nezávisle od toho, aké je ich priestorové umiestnenie. Výskumy ukazujú, že selekcia môže byť založená na priestorovej oblasti, teda umiestnení objektu („location based“), na samotnom objekte („object oriented“) alebo môže byť objektovo-znakovo založená („object-token based“) [9].

Druhý komponent, **sledovanie**, znamená udržiavanie pozornosti na cieľovom objekte napriek prítomnosti distraktorov.

Tretí komponent pozornosti, **riadenie**, je proces, pri ktorom je systém schopný zmeniť stálu pozornosť z jedného objektu na iný objekt, ktorý sa stáva novým cieľom interakcie [1]. Pozornosť môžeme riadiť viacerými spôsobmi, napr. zrakovo, sluchovo a pod. S tým súvisí otázka, aký je rozdiel medzi tým, ak je pozornosť riadená unimodálne (sluchová pozornosť riadená sluchom, zrková zrakom, a pod.) a krosmodálne (sluchová zrakom, zrková sluchom, ...).

Pozornosť sa ešte ďalej delí podľa toho, či to, na čo sa sústreďíme, sledujeme aj zrakom – tzv. „covert orienting“ alebo nie – „overt orienting“.

### 2.3 Multimodálne a krosmodálne vnímanie

Doterajší výskum súvisiaci s pozornosťou sa väčšinou zaoberal len jednou modalitou. Boli to napríklad experimenty zamerané na vizuálne vnímanie alebo selektívne počúvanie, málokteré však skúmali interakcie medzi modalitami. Pozornosť bola chápaná ako niečo špecifické pre danú modalitu a nezávislé od ostatných modalít [9].

Keďže však podnety z okolia vnímame multimodálne (t.j. viacerými zmyslami), je možné, že dochádza k interakciám, ktoré by sa mali brať do úvahy. Typickým príkladom je vnímanie reči. Ak sa hovoriaci nachádza v hlučnom prostredí, porozumenie reči sa stáva ťažším. Samotné vnímanie sluchom nestačí, tak si pomáhame aj zrakom – v praxi to znamená, že sledujeme pohyb jeho pier, jeho gestá, mimiku. Vizuálna informácia dokáže podstatne ovplyvniť to, čo počujeme, a dôkazom toho je McGurkov efekt [10]. Tento efekt sa väčšinou demonštruje na videu s detailným záberom na tvár človeka, ktorý vyslovuje slabiku “ga”, z reproduktorov však namiesto toho znie slabika “ba”. Na otázku, čo počujú, však ľudia napodiv neodpovedajú tú slabiku, ktorá znie z reproduktorov, avšak ani tú, ktorú vyčítajú z pier. Počutá a videná informácia totiž spolu nejakým spôsobom interagujú a výsledkom je, že ľudia počujú “da”.

Krosmodálne interakcie môžu mať na vnímanie pozitívny aj negatívny vplyv. Negatívny vplyv, ktorého prejavom bol aj McGurkov efekt, vzniká väčšinou vtedy, ak si sensorická informácia vnímaná rôznymi modalitami protirečí. To vedie ku 1) kombinácii podnetov, ktorá je určitým druhom váženého priemeru (McGurk), 2)



ponechaniu jedného z podnetov a potlačeníu druhého, alebo 3) vzájomnému potlačeníu podnetov a zmäteniu [1].

Pozitívny vplyv sa na neurálnej úrovni prejavuje tak, že neuróny vstupy z jednotlivých modalít sčítavajú a určitým spôsobom transformujú, zosilnenie signálu sa však prejavuje len vtedy, ak sú stimuly z jednotlivých modalít slabé. Ak bol jeden z unimodálnych stimulov silný, môže naopak zoslabiť druhý stimul, intenzita unimodálnych stimulov preto zohráva podstatnú úlohu.

Štúdium krosmodálnych interakcií pomáha skúmať pozornosť v tom zmysle, či je špecifická pre jednotlivé modalitty alebo zdieľaná. Keďže pozornosť je obmedzená, uvažuje sa o nej z hľadiska jej zdrojov. Ak sa pozornosťné zdroje nezdieľajú a dotyková informácia (napr. textúra) sa prezentuje vizuálne, zdroje, ktoré boli určené pre čisto vizuálne podnety, sa týmto znížia. Naopak, prezentácia vizuálnej informácie dotykom by mohla uvoľniť vizuálne zdroje [1]. Ak sa však pozornosťné zdroje zdieľajú, prezentovanie podnetu jednej modalitty inou modalitou spoločné zdroje len zníži, nikdy neuvoľní. O existencii spoločných zdrojov by mohlo svedčiť napríklad to, že ak očakávame podnet jedným zmyslom, spôsobí to očakávanie podnetu z daného miesta aj inými zmyslami. Ak je však podnet v tej modalitte, aká je očakávaná, spôsobí to menšiu záťaž na pozornosťné zdroje ako keď je v inej.

Ak sú človeku prezentované stimuly z rôznych modalít, môže ich integrovať do jednej udalosti alebo vnímať ako dve oddelené udalosti. Závisí to od viacerých faktorov, akými sú napríklad časová alebo priestorová blízkosť stimulov alebo napríklad aj od toho, či sú stimuly sémantický kongruentné. Ukázalo sa napríklad, že ak sa spolu so zvukom prezentuje aj vizuálny podnet, lokalizácia tohto zvuku sa posúva smerom k vizuálnemu podnetu. Ak sa však zvýši vzdialenosť alebo časové opozdenie medzi týmito stimulmi, tento efekt klesá [11].

## 2.4 „Cocktail party effect“

Kokteil párty efekt je schopnosť zamerať sluchovú pozornosť na jedného hovoriaceho spomedzi množstva konverzácií a šumu z pozadia [3].

Z pohľadu počúvajúceho je táto úloha jednoduchá a intuitívna, z hľadiska psychológie a fyziológie je to však zložitý problém, ktorý zahŕňa interakcie medzi signálom, sluchovým systémom a centrálnym nervovým systémom. Akusticky je tento

problém podobný problému vyčleniť reč jedného hovoriaceho zo spektrogramu obsahujúceho signály od viacerých hovoriacich a šum z pozadia, a to stroje zatiaľ nedokážu [3].

Výskum súvisiaci s cocktail party efektom prebieha od 50-tych rokov minulého storočia a doteraz nie je tento problém úplne vysvetlený [17]. Cocktail party efekt úzko súvisí so selektívnou pozornosťou – lebo človek vyčlení (selektuje) jednu z množstva konverzácií a na ňu sa zameria. Otázkou však je, ako to robí. Ako ľudia segregujú zvuky a ktoré podnety sú pri tom podstatné?

Prvý, kto sa zaoberal týmto problémom, bol Colin Cherry v práci [1]. Ako prvý zaviedol termín „cocktail party effect“ a popísal ho. Vo svojich experimentoch na rozpoznávanie správ prichádzajúcich do jedného alebo do oboch uší prišiel na to, že pri segregácii hlasov pomáhajú tieto faktory:

- hlasy prichádzajú z rôznych smerov
- čítanie z pier, gestá a pod.
- rôzne hlasy, stredné výšky tónov hlasov, stredné rýchlosti, muž vs. žena, atď.
- rôzne prízvuky
- prechodové pravdepodobnosti (založené na dynamike zvuku, syntaxi, ...)

Sedem fyzických atribútov zvuku ,ktoré môžu byť použité ako základ pre lokalizáciu zdroja zvuku.

1. Spektrálne oddeľovanie
2. Profil priestoru
3. Hormonizácia
4. Priestorové oddeľovanie
5. Temporálna separácia
6. Temporálny onset a offset
7. Temporálna modulácia

Všetky faktory okrem posledného odstránil Cherry tak, že nechal toho istého človeka nahovoriť dve správy „cez seba“ na magnetofónovú pásku. Napriek tomu sa správy dali rozlíšiť. Cherry na základe toho usúdil, že ľudia sú schopní predikovať sekvencie slov pričom majú v pamäti uložené veľké množstvo prechodových pravdepodobností, ktoré im pri tom pomáhajú.

Experimenty v tejto oblasti sa často robia formou **dichotického počúvania**. Podstata dichotického počúvania je v tom, že subjekt dostane slúchadlá a do každého ucha mu znie iný text a jeho úlohou je zamerať sa na jeden z nich. Popri tom ho má nahlas opakovať, čo sa nazýva „**shadowing**“ – **tieňovanie** [3]. V Cherryho experimentoch sa ukázalo, že táto úloha je ľahko zvládnuteľná, subjekt bez problémov zameria pozornosť na správu znejúcu do jedného ucha a druhú ignoruje a túto pozornosť môže ľahko „prepínať“ z jedného ucha na druhé. O obsahu textu znejúceho do tienneho ucha má subjekt malú predstavu, z textu znejúceho do ignorovaného ucha si však nepamätá prakticky nič.

V experimentoch sa skúmalo aj aké zmeny v texte znejúceho do ignorovaného ucha boli subjekty schopné zaregistrovať:

**zaregistrovali** tieto zmeny:

- zmena hlasu z mužského na ženský a naopak (zvyčajne)
- zmena reči na čistý tón (vždy)

**nezaregistrovali** tieto zmeny:

- zmena z angličtiny na nemčinu
- reverzne prehraná reč

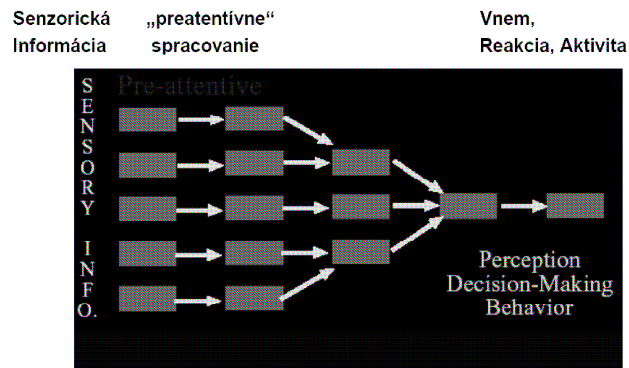
Tieto výsledky ukazujú, že zo správy, na ktorú nemáme zameranú pozornosť, sme schopní zaregistrovať všeobecné vlastnosti signálov (fyzikálne), avšak nie detaily ako jazyk, individuálne slová, sémantiku.

## 2.5 Modely pozornosti

### 2.5.1 Broadbentova teória – pozornosť ako filter (1958)

Na základe dovtedajších poznatkov o pozornosti získaných aj z Cherryho experimentov navrhol Broadbent teóriu, podľa ktorej pozornosť predstavuje filter, ktorý

vyberie, ktoré z množstva prijímaných signálov budú ďalej spracovávané (Obr. 9, Obr. 10)

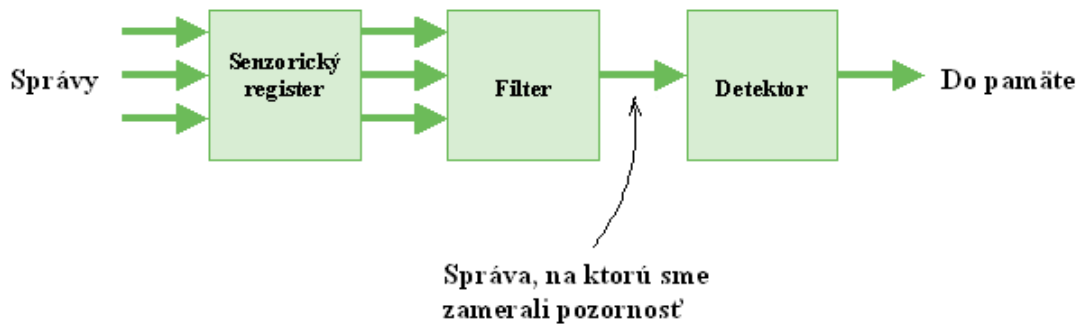


**Obr. 9 Pozornosť ako filter. Senzorické vstupy rozdelené do jednotlivých kanálov sú najskôr spracovávané paralelne (preatentívne = „predpozornostné“ spracovanie), niektoré sú prepustené na ďalšie podrobnejšie spracovávanie a tento výber realizuje pozornosť filter [1].**

Teóriu by sme mohli zhrnúť do týchto základných bodov:

- mozog separuje prichádzajúce zvuky na základe ich fyzikálnych charakteristík (napríklad toho, odkiaľ znejú) do kanálov
- iba niektoré subsignály prepustí na sémantickú analýzu (dešifrovanie významu), množstvo je limitované
- audio filter, ktorý je ovládaný vedomím, určí na ktoré kanály sa zameriame (selekcia)
- nevybrané kanály sa nespracujú vôbec

Broadbentov model je jedným z modelov skorej selekcie (lebo selekcia prebieha v skoršom štádiu procesu spracovávaní informácie – ešte predtým, ako sa určí jej význam) [1]. Tento model pozornosti bol však neskôr kritizovaný. Ak sa totiž nevybrané kanály ďalej vôbec nespracovávajú, ako je možné, že ak niekto, komu nevenujeme pozornosť (čo predstavuje nevybraný kanál), zrazu vysloví naše meno, automaticky presunieme pozornosť na to, čo hovorí?



**Obr. 10 Broadbentov model.** Senzorický register drží informáciu iba krátky čas a posúva ju filtru. Filter identifikuje správu, na ktorú sa zameria pozornosť pomocou fyzikálnych charakteristík (hlas, výška tónu, rýchlosť, akcent). Všetky ostatné informácie okrem tej „zameranej“ sú odstránené (odfiltrované). Správa, ktorá prešla filtrom, pokračuje do detektora, kde je ďalej spracovávaná, aby sa určil jej význam. Detektor spracuje všetky informácie, ktoré sa doň dostanú. Výstupom z detektora je krátkodobá pamäť. Informácie z nej sa po krátkom čase buď stratia, alebo prenesú do dlhodobej pamäte [12].

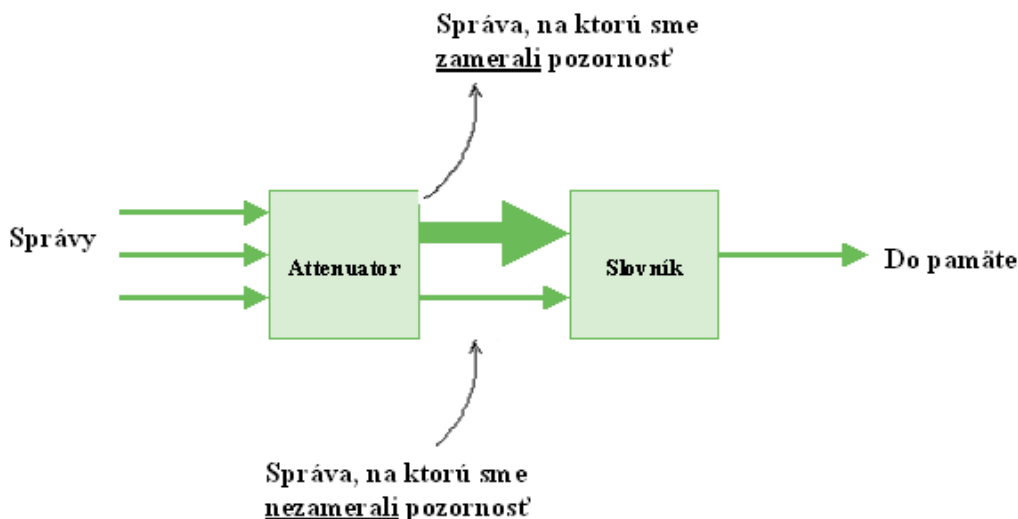
### 2.5.2 Model Anne Triesman – teória zoslabovania – model skorej (alebo strednej) selekcie (1964)

V úvode kapitoly 2.4 bolo napísané, že cocktail party effect je schopnosť zamerať sluchovú pozornosť na jedného hovoriaceho spomedzi množstva konverzácií a šumu z pozadia. Pojem „cocktail party effect“ však nesie aj druhý význam. Prvý súvisel so strategickou pozornosťou, vedomým zameraním sa na niečo. Druhý význam súvisí s automatickou pozornosťou. Je to situácia, keď pre nás zaujímavý podnet (napr. keď niekto vysloví naše meno) náhle automaticky pritiahne našu pozornosť. Týmto fenoménom sa zaoberal Moray a jeho výsledky poukazujú na nesprávnosť Broadbentovho modelu [1]. Hromadiace sa dôkazy proti Broadbentovmu modelu prispeli k vzniku nových modelov.

Anne Triesman robila experimenty formou dichotického počúvania. Medzi prezentovanými správami boli vety, ktoré začínali v jednom uchu a končili v druhom. Ukázalo sa, že ľudia participujúci na experimente v 30% prípadoch zaznamenávali tieto vety a nie to, čo počuli v tieňovanom uchu [2]. Jej zistenie aj ostatné dôkazy proti Broadbentovmu modelu ju viedli k vytvoreniu **teórie zoslabenia (attenuation theory)**, ktorej model je na Obr. 11.

Podľa tejto teórie pozornosťný systém nevymaže to, na čo sa nekoncentrujeme, úplne, len to potlačí. Potlačený vstup zväčša ignorujeme, niekedy však dôjde k prielomu – ak sa jedná o slová, ktoré sú pre nás podstatné (napr. naše meno) [1].

Problémom tohto modelu je, že nie je jasné, čo znamená „zoslabené spracovanie významu“.



Obr. 11 Model teórie zoslabovania. Filter z Broadbentovho modelu je nahradený „attenuatorom“, ktorý zoslabuje informácie, na ktoré sa nesústredíme. Attenuator analyzuje prichádzajúce správy podľa fyzikálnych charakteristík, jazyka, významu. Slovník je ako pamäť. Obsahuje uložené slová, ktoré majú rôzne prahy aktivácie. Nižší prah aktivácie pre dané slovo znamená aj vyššiu citlivosť na dané slovo [12].

### 2.5.3 Teória Deutscha & Deutschej – model neskorej selekcie (1963)

V modeloch neskorej selekcie, ako už vyplýva z názvu, dochádza k selekcii v neskoršom štádiu spracovania vstupu. Podľa tohto modelu k selekcii dochádza až po spracovaní významu počutého slova. Teda všetky podnety prechádzajú analýzou významu, ale iba niektoré sú vybrané pre odpoveď. Analýzou významu teda prechádzajú aj správy, ktorým nebola venovaná pozornosť [12]. Neprejdú síce do vedomia, ale dokážeme ich spoznať, vnímame ich podprahovo.

Tomuto modelu bolo vytknuté to, že aj keď je analýza významu v kanáli, ktorému nevenujeme pozornosť dobrá, v kanáli, ktorému pozornosť venujeme, je lepšia.

## 3 Experimentálna časť – vnímanie reči

### 3.1 Experiment „Cocktail Party Effect“

#### 3.1.1 Motivácia

Pozornosť zohráva veľkú úlohu aj pri vnímaní reči. Pomáha nám zorientovať sa v hlučných prostrediach a v spleti súbežne znejúcich konverzácií počúvať jedného konkrétneho človeka. Tento problém vyčleniť jeden „prúd reči“ z množstva ďalších a potlačiť vnímanie všetkých nepodstatných zvukov, sa nazýva kokteil party problém [3].

Tento experiment nadvzuje na pilotný experiment Ing. Beáty Tomoriovej [20] Štúdie a analýzy budú prispôbené podľa pilotného projektu a budú rozšírené o nové, na základe výsledkov z pilotnej štúdie.

Priestorová pozornosť v kokteil party probléme je témou viacerých štúdií. Niektoré ukázali, že ak poznáme polohu a hlas hovoriaceho, presnosť detekcie cieľovej správy medzi maskovacími správami je vyššia [15]. Aj v štúdiu [16] informácia o polohe hovoriaceho zlepšovala výsledky. Subjekty mali za úlohu identifikovať cieľovú správu medzi maskovacími správami. Správy boli priestorovo oddelené, zneli z piatich reproduktorov. Ak subjekty dostali informáciu, z ktorého reproduktora zaznie cieľová správa, zvýšilo to percento ich správnych odpovedí o 25%.

Tento experiment má podobné ciele ako spomínané, skúma vplyv priestorovej pozornosti na vnímanie reči, tak ako v [16] aj vplyv špecifickej polohy, odkiaľ zaznie správa, a navyše aj vplyv dĺžky páuz medzi jednotlivými slovami na ich identifikáciu.

Pri tomto experimente sme vybrali sluchový stimul (ďalej len cue) ,ako stimul ,ktorým sa budú subjekty riadiť pri lokalizácii cieľových zvukov.

#### 3.1.2 Hypotézy

- Ak bude subjekt poznať polohu hovoriaceho, bude vedieť lepšie identifikovať, čo hovorí, ako keď ju nebude poznať.
- Porozumenie sa zlepší s rastúcimi pauzami medzi slovami, keďže subjekty budú mať viac času „zanalyzovať“ čo počuli.

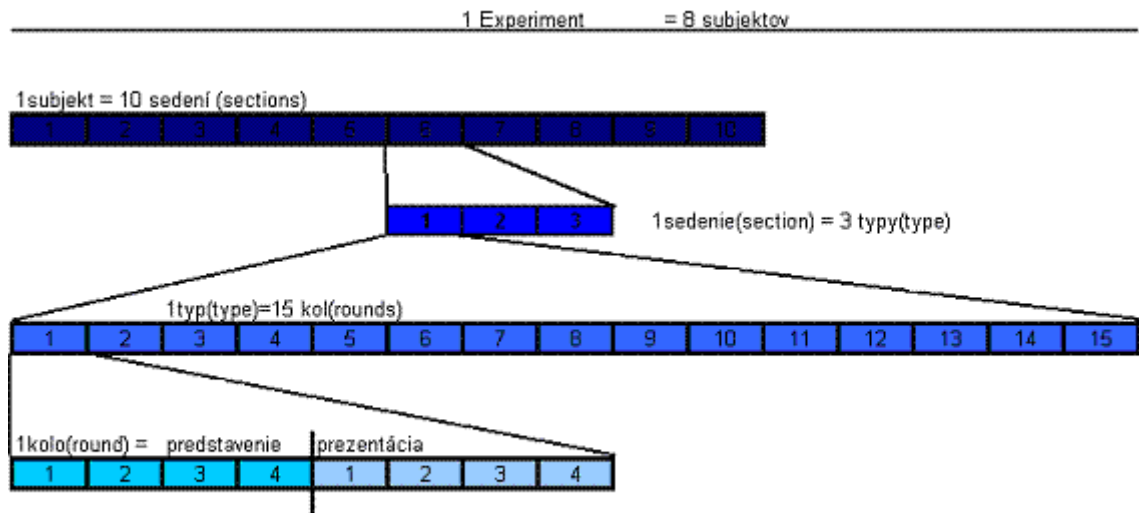
- Porozumenie bude lepšie pre krajné polohy hovoriacich ako pre stredové (vychádzajúc z [16]).
- Ak sa poloha hovoriacich nebude meniť, tak percento úspešnosti bude vyššie.
- Ak za stimul, ktorý bude určovať polohu cieľového hovoriaceho, vložíme pauzu, tak percento úspešnosti bude vyššie.

### 3.1.3 Metódy

#### 3.1.3.1 Popis experimentu

Experimentu sa zúčastnilo 8 subjektov (4 ženy a 4 muži vo veku 21 – 26 rokov) so sluchom v norme. Časti experimentu sú zobrazené na Obr.12. Experiment na jednom subjekte pozostával z 10 sedení. Jedno sedenie obsahovalo 3 bloky statický, synchronný a asynchronný, ktoré sa náhodne generovali. Bloky synchronný a asynchronný boli vytvorené z dynamického bloku v [20], ktorých rozdiel popíšem nižšie. Každý blok mal 15 kôl. V rámci bloku sa menila dĺžka páuz medzi vyslovovaním jednotlivých slov (čísel). Boli použité tri rôzne ISI (inter stimulus interval): 5ms, 500ms a 2000ms. Pri statickom bloku bola poloha hovoriaceho konštantná, pri synchronnom a asynchronnom nie. V rámci jedného bloku (statického, synchronného alebo asynchronného) sa náhodne vystriedali všetky možné kombinácie polohy a ISI, čo tvorí  $5 \times 3 = 15$  kôl. V jednom bloku bol výber cieľového hovoriaceho z jednotlivých hlasov rovnomerný (každý hlas bol cieľovým hlasom v jednom bloku práve 3krát). V rámci jedného kola, so zahrnutím aj predstavenia cieľového hlasu, bolo ISI konštantné. Počas jedného sedenia sa náhodne použilo každé ISI raz pre každý reproduktor. Jedno kolo sa skladalo z predstavenia a samotnej prezentácie. Predstavenie aj prezentácia sa skladali zo sekvencie 4 číslíc. Jedno kolo pozostávalo z prehratia cieľového stimulu cez reproduktory a zo zaznamenania odpovede subjektu. Úlohou subjektu bolo identifikovať čísla, ktoré zazneli počas prezentácie z reproduktorov určených vodítkom (cue). Cue bol zvukový stimul dlhý 5ms o frekvencii 20,05kHz. Počas prezentácie sa súčasne s cieľovým hlasom ozývali aj ďalšie čísla vyslovované inými hlasmi, ktoré maskovali túto cieľovú sekvenciu (sekvenciu 4 čísel budeme ďalej nazývať prezentácia). Pred každou prezentáciou sa im tento hlas predstavil – povedal náhodné 4 čísla (ďalej predstavenie alebo predstavenie v prezentácii).





Obr. 12 Rozdelenie Experimentu Koktail party efekt na jednotlivé kroky.

Každé sedenie pozostávalo z troch základných blokov o ktorých boli subjekty informované:

- Statický blok – cieľový hlas znel v rámci prezentácie stále z toho istého reproduktora – z toho, z ktorého sa predstavil. Cue sa predstavil pred každým zo štyroch slov pri predstavení ,a pri každom zo štyroch slov pri prezentácii pri cieľovom hlase. Tichá pauza s dĺžkou ISI predchádzala cue a cue predchádzalo cieľový hlas. Cue zaznelo päť milisekúnd pred cieľovým hlasom. Toto meranie poskytlo subjektom informáciu o polohe cieľového hlasu. ISI sa v týchto blokoch vyberali náhodne ,ale každé bolo použité raz pre každý reproduktor.
- Synchronný blok – cieľový hlas znel v každom segmente prezentácie aj predstavovania z náhodného reproduktora. Počas prezentácie sa cieľový hlas nevyskytol dvakrát na tej istej pozícii. Tichá pauza s dĺžkou ISI predchádzala cue a cue predchádzalo cieľový hlas. Cue zaznelo päť milisekúnd pred cieľovým hlasom. Toto meranie poskytlo subjektom informáciu o polohe cieľového hlasu stále päť milisekúnd pred cieľovým hlasom.

- Asynchrónny blok – cieľový hlas znel v každom segmente prezentácie aj predstavovania z náhodného reproduktora. Cue predchádzalo o päť milisekúnd pred tichou pauzou s dĺžkou ISI a Tichá pauza s dĺžkou ISI predchádzala cieľový hlas. Toto meranie poskytlo subjektom informáciu o polohe cieľového hlasu stále päť milisekúnd pred ISI. Pri tomto bloku má subjekt viac času zamerať svoju pozornosť. Cue predchádzalo ISI a ISI predchádzalo cieľový hlas.

Maskovacie správy boli v každom zo 4 čísel segmentu prezentácie predstavujúcich vyslovenie jedného čísla vyberané náhodne čo sa týka hlasov, čísel aj polohy. Z jedného reproduktora zaznel hlas v jednom čísle segmentu práve raz. Päť hovoriacich bolo náhodne vybraných z pätnástich na začiatku každého sedenia. Z týchto piatich sa nahodene vybral jeden cieľový a štyri maskovacie po každej prezentácii. Pre cieľový hlas boli štyri čísla z deviatich (okrem nuly) ,nahodne generované, a vybrané pre každú zo štyroch sekvencií ,tak ako pre predstavenie ,tak aj pre prezentáciu. Bolo však zabránené tomu, aby súčasne z viacerých reproduktorov naraz znelo to isté číslo alebo ten istý hlas. Tiež bolo zabránené ,aby sa cieľové číslo v rámci prezentácie opakovalo. Čísla pre maskovacie hlasy sa vygenerovali ,tak aby boli odlišné od cieľového a tiež tak aby v rámci prezentácie neznelo to isté číslo z viacerých reproduktorov naraz.

Keďže nie všetky nahrávky čísel mali tú istú dĺžku, za nahrávky s najdlhšou dĺžkou sa nepridávala žiadna pauza a za všetky ostatné slová sa pridal pauza ,aby dohrali rovnako ako najdlhšia nahrávka. Prezentácia s cieľovým hlasom bola dlhšia o pridané cue vzhľadom k maskovaným hlasom. Na odstránenie tejto časovej nerovnosti sme aj do maskovaných prezentácií vložili tiché intervaly dlhé 5ms na začiatok každého maskovacieho slova. To sa opakovalo pre každý segment prezentácia aj predstavenia.

V experimente boli použité číslice znejúce v anglickom jazyku, konkrétne čísla od 1 do 9 nahovorené 15 rôznymi hlasmi mužskými aj ženskými hlasmi. Tieto nahrávky sú súčasťou [22] databázy, odkiaľ boli prevzaté.

Experiment prebiehal v odzvučenej miestnosti Tepro –noise silencer o rozmeroch 1,8m x 2,9m x 3m. Bola použitá 7 kanálová zvuková karta Sound blaster live 24 , Creative technology a multikanálový zosilňovač Sony s výkonom 220W. Hlasy boli

prezentované z piatich reproduktorov rozostavených v azimutoch  $-40^\circ$ ,  $-20^\circ$ ,  $0^\circ$ ,  $20^\circ$ ,  $40^\circ$ , vzdialených približne 120cm od subjektu a umiestnených približne vo výške uší. Reprodukory boli umiestnené symetricky do stredu miestnosti, pre prípad nežiadúcich odrazov. Subjekty boli upozornené ,aby počas celého experimentu nehýbali hlavou a mali ju orientovanú čelom na prostredný reproduktor v azimute  $0^\circ$ . Pohyb očí nebol fixovaný. Počas experimentu sa subjekt riadil pokynmi zobrazovanými cez projektor za reproduktory na stenu. Obraz projektora bol umiestnený do azimutu  $0^\circ$  nad reprakmi., jeho spodný okraj bol vo výške horného okraja reproduktorov.

### 3.1.4 Tréning

Každý zo subjektov absolvoval 1 skúšobné kolo experimentu, aby sa oboznámil s experimentálnou procedúrou ,aby zvykol si na prezentované hlasy a aby sa mu jeho sluch prispôbil. Toto kolo nebolo zaratané do experimentu.

### 3.1.5 Experimentálna procedúra

Subjekt bol informovaný stále pred novým blokom experimentu o tom, aký blok nasleduje. Pred každou prezentáciou, ešte pred predstavením cieľového hlasu, bol subjekt informovaný o dĺžke ISI. Po tomto výpise sa mu predstavil cieľový hlas v predstavení spolu aj s cue. V statickom bloku znel stále z jedného reproduktora, v synchrónnom a asynchrónnom striedavo z náhodných. Hlas sa predstavil vyslovením náhodných 4 čísel, pauza medzi nimi odpovedala aktuálnemu ISI. Na predstavenie cieľového hlasu bez interakcie subjektu automaticky po uplynutí ISI nadviazala samotná prezentácia, počas ktorej už však zneli zo všetkých ostatných reproduktorov aj maskovacie správy. Subjekt po doznení prezentácie napísal sekvenciu 4 čísel, ktoré počul od cieľového hlasu ,čiže z pozície kde zaznelo vidítka-cue.

### 3.1.6 Spôsobu vyhodnocovania

Výsledky sa vyhodnocovali určením percenta správnych odpovedí. Použili sa na to dva spôsoby:

1. s celými sekvenciami – odpoveď sa považovala za správnu len vtedy, ak subjekt uhádol všetky čísla v prezentácii a v správnom poradí. Výsledky boli

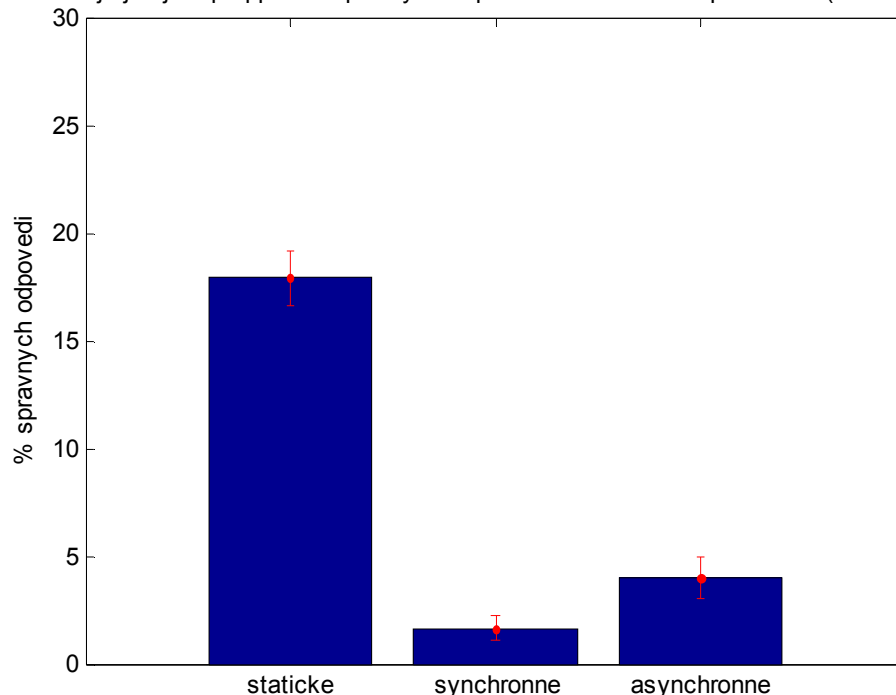
zokupované podľa prostredia a ISI. Pre každú kombináciu sa určilo percento správnych odpovedí.

2. s „rozbitými“ sekvenciami – hodnotila sa odpoveď v konkrétnom segmente prezentácie, to znamená, že bola správna, ak subjekt uhádol číslo v danom segmente a nesprávna, ak neuhádol. Výsledky boli zokupované podľa prostredia, polohy, ISI a sekvencie. Pre každú z týchto kombinácií sa určilo percento správnych odpovedí.

Väčšina vygenerovaných grafov bola na základe 2. spôsobu, pretože pri 1. spôsobe dosahovali subjekty len veľmi malé percento správnych odpovedí a taktiež pomocou 2. spôsobu bolo možné sledovať vplyv polohy hovoriaceho na výsledky.

### 3.1.7 Výsledky

bt dh2 es jk jk2 jkm ph pp : % správnych odpovedí v závislosti od prostredia (cele sekvencie)

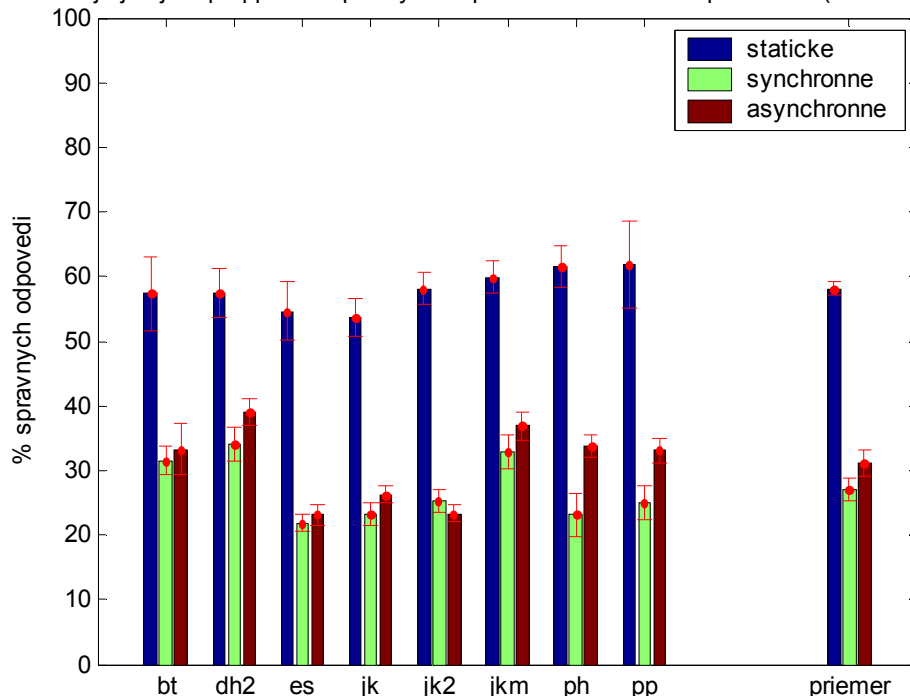


**Obr. 13** Percento správnych odpovedí v závislosti od prostredia (vyhodnocovanie podľa spôsobu 1), priemer cez ISI a subjekty, error bary znázorňujú štandardnú chybu.

Obr. 13 ukazuje percento správnych odpovedí v závislosti od prostredia. Je vidieť, že ak boli dáta vyhodnocované prvým zo spomínaných spôsobov, percento správnych

odpovedí je veľmi nízke, približne 18% v statickom prostredí ,3% v synchrónnom a v asynchrónnom 8%. Zase sa potvrdil výrazný rozdiel medzi statickým a synchrónnym ,či asynchrónnym prostredím. Tak isto ako v diplomovej práci Beaty Tomoriovej.

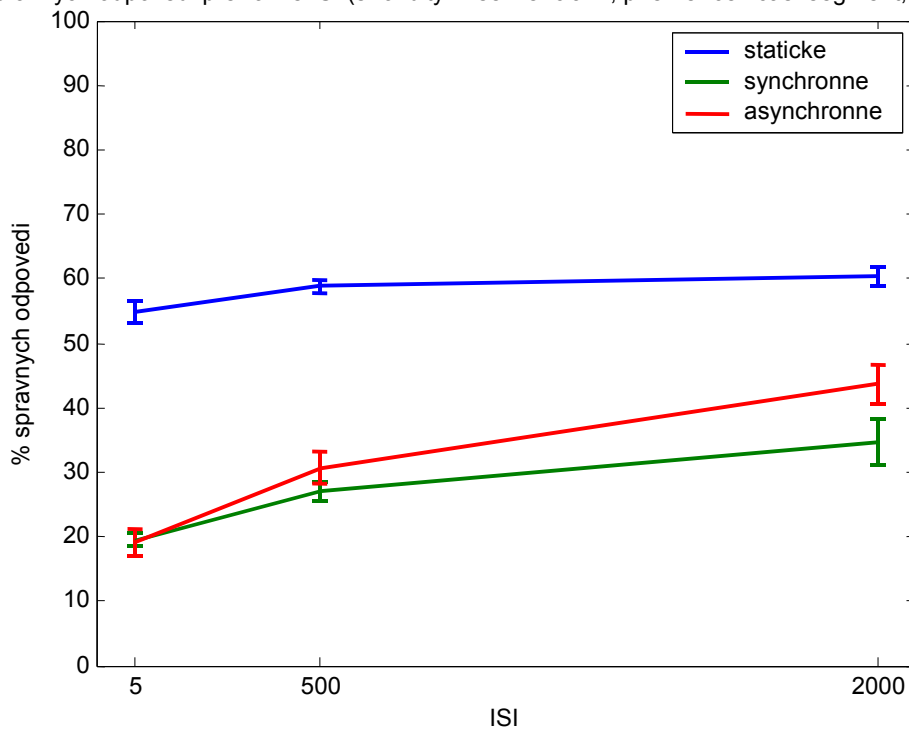
bt dh2 es jk jk2 jkm ph pp : % správnych odpovedí v závislosti od prostredia (rozbite sekvencie)



**Obr. 14** Percento správnych odpovedí v závislosti od prostredia pre jednotlivé subjekty a celkový priemer (v pravej časti obrázku), errorbary znázorňujú štandardnú chybu.

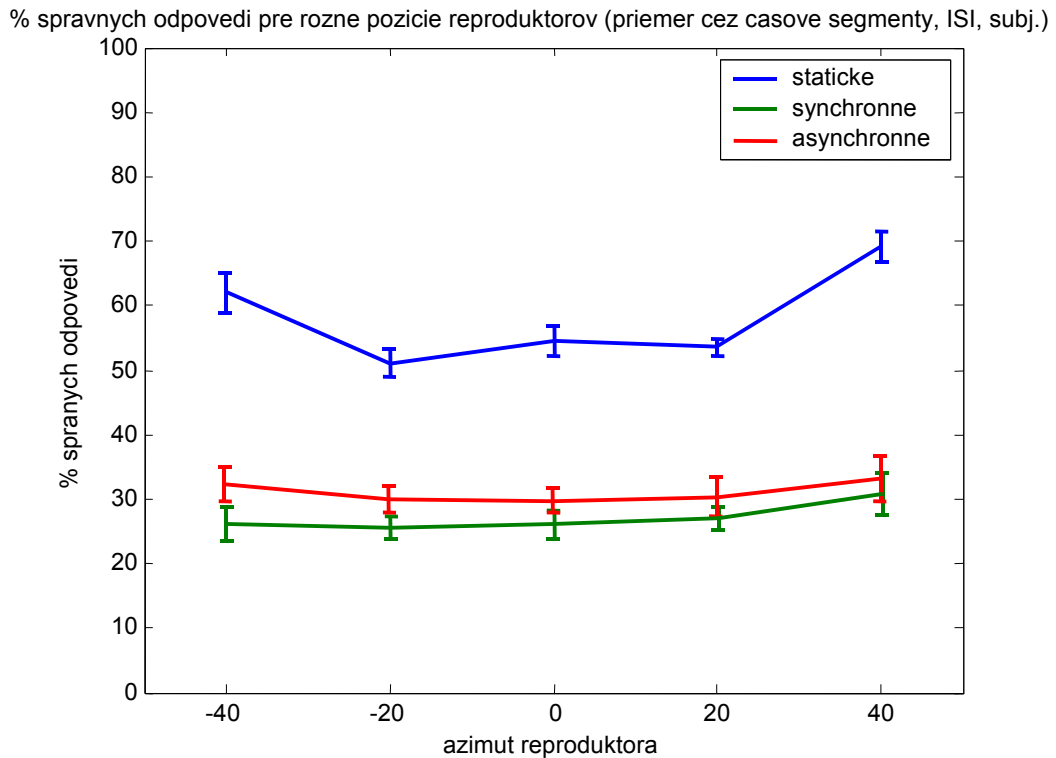
Pri druhom spôsobe, ktorým sú generované už všetky nasledujúce grafy, sú už percentá správnych odpovedí omnoho vyššie, statické kolá dosahujú presnosť približne 58%, synchrónne približne 27%, a asynchrónne približne 31% čo je vidieť na Obr. 14, ktorého ľavá časť predstavuje hodnotenia jednotlivých subjektov a pravá ich priemer. U každého subjektu boli výsledky v statických kolách približne o 25-35% lepšie ako v synchrónnych alebo asynchrónnych. Keďže subjekty „bt“, „dh2“ a „jkm“ strávili pri tréningu viac času ako ostatné subjekty a mali najlepšie výsledky, je potvrdenné čo bolo avízované v predchádzajúcej štúdii, že predĺžením tréningu sa presnosť u všetkých zvýšila. Tréning naučí lepšie rozpoznávať jednotlivé hlasy, čo pomôže pri ich separácii.

% spravnych odpovedi pre rozne ISI (s rozbitymi sekvenciami, priemer cez cas. segment, poziciu, su



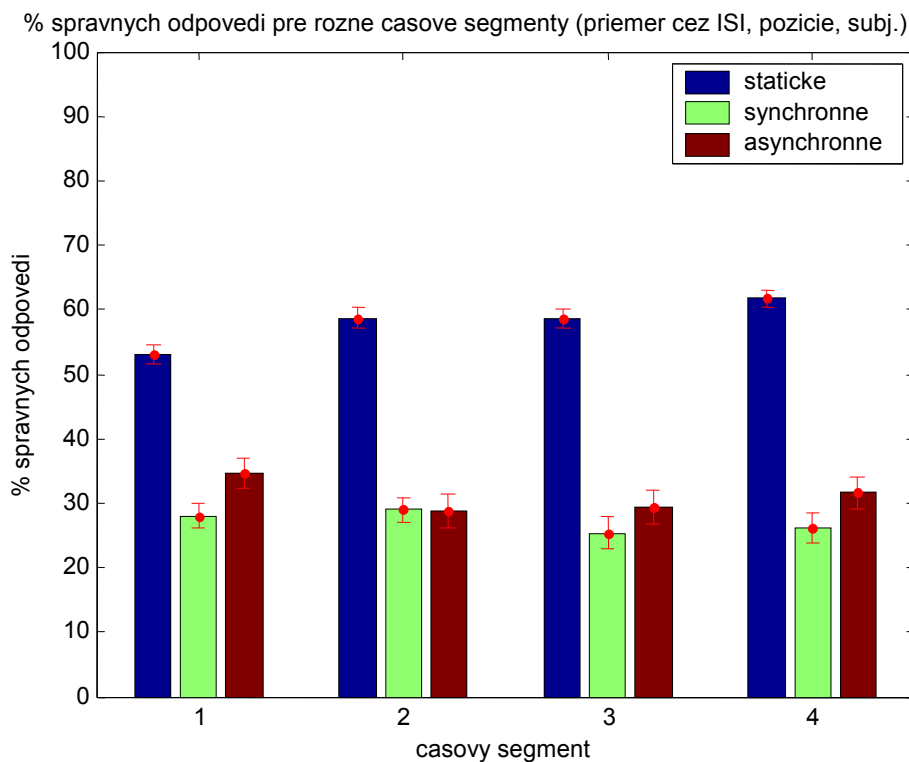
**Obr. 15** Percento správnych odpovedí v závislosti od ISI (priemer cez časové segmenty, pozíciu, subjekty), errorbary znázorňujú štandardnú chybu.

Výsledky tiež ukázali, že percento správnych odpovedí bolo viac-menej nezávislé od dĺžky pauzy pri statických kolách, ale pri synchronných kolách sa percento správnych odpovedí zvýšilo o 10% a pri asynchrónnych kolách až o 20% medzi jednotlivými slovami (Obr. 15).



**Obr. 16** Percento správnych odpovedí v závislosti od pozície (priemer cez časové segmenty, ISI a subjekty), errorbary znázorňujú štandardnú chybu.

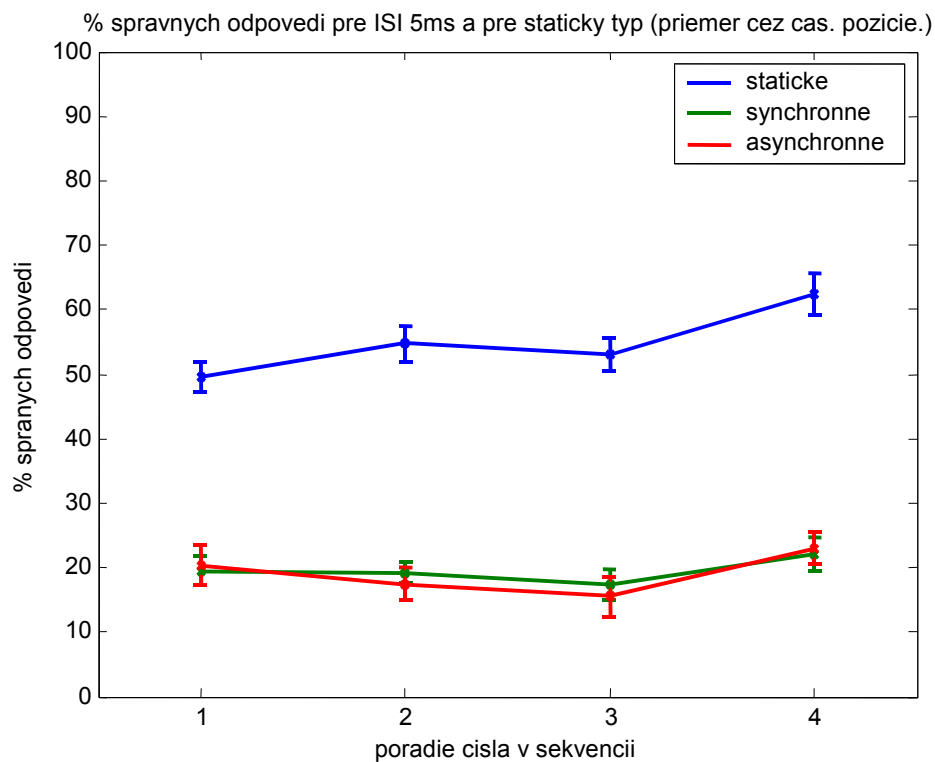
Obr. 16 ukazuje vplyv polohy cieľového hlasu na odpovede. V statickom prostredí boli odpovede výrazne presnejšie, ak cieľový hlas znel z niektorého z krajných reproduktorov. Presnosť bola približne o 20% vyššia v porovnaní s ostatnými azimutmi. V synchronnom a asynchronnom prostredí už rozdiel medzi krajnými azimutmi a ostatnými nebol taký výrazný.



**Obr. 17** Percento správnych odpovedí v závislosti od segmentu prezentácie (priemer cez ISI, pozície a subjekty), errorbary znázorňujú štandardnú chybu.

Z výsledkov na Obr.17 je tiež zrejmé, že pre statické dáta sa s časovým segmentom presnosť zvyšuje. Pri synchronnom a asynchronnom nie. Pre dynamické dáta asynchronná cue má vyššiu percentuálnu úspešnosť než synchronná, ale pre druhý segment. Najväčší rozdiel je v prvom segmente.

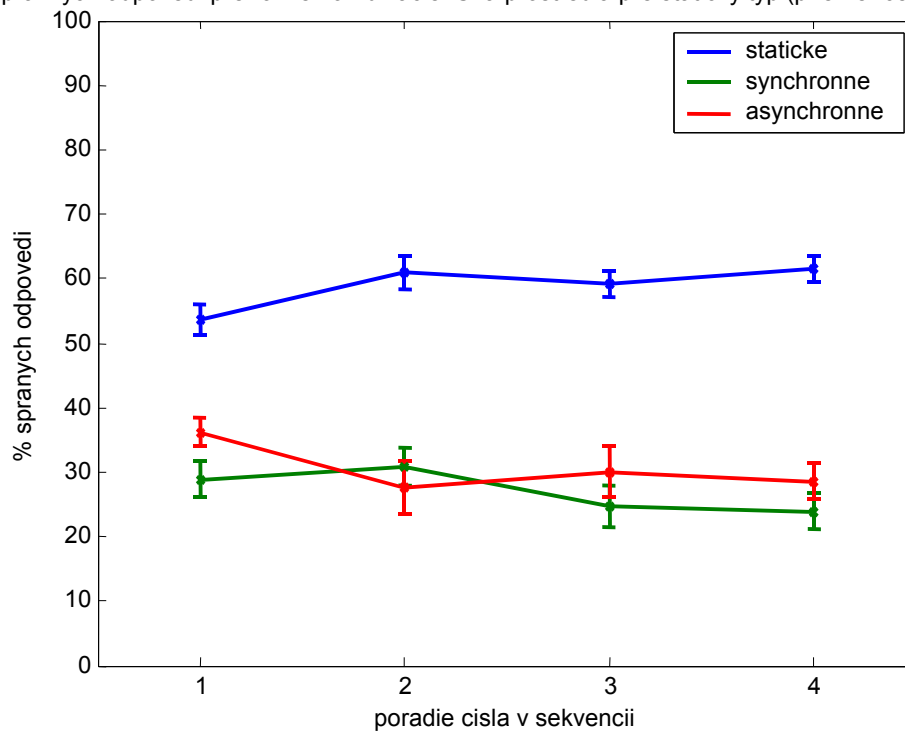




**Obr. 18** Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 5ms a subjekty), errorbary znázorňujú štandardnú chybu.

Z Obr. 18 vidíme, že pri 5ms dĺžke ISI sú výsledky lepšie na poslednom segmente vo všetkých troch typoch.

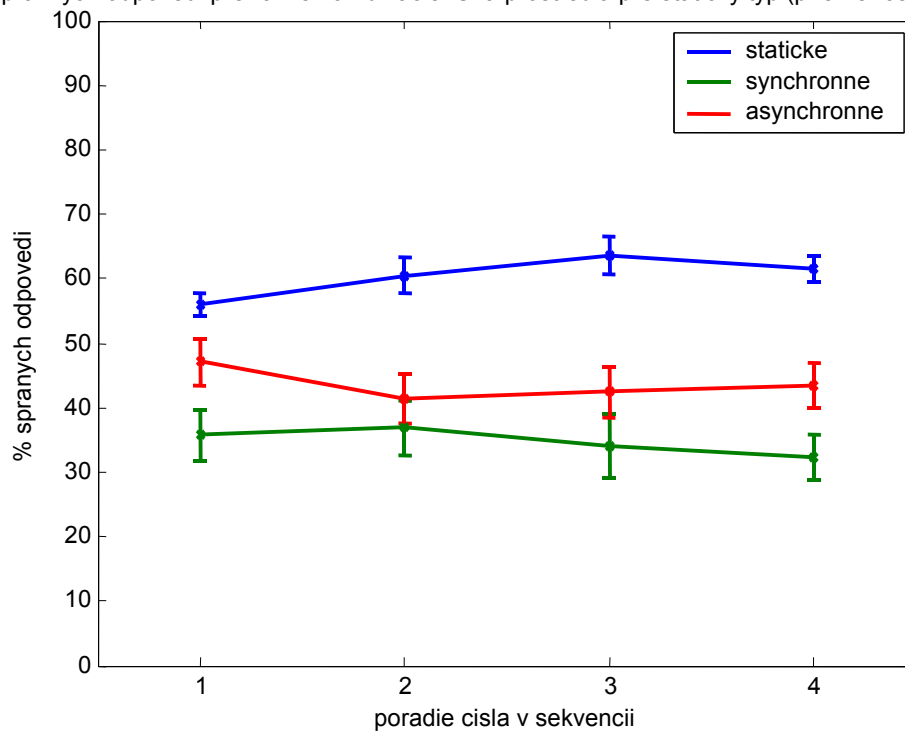
% správnych odpovedí pre rozne kombinácie ISI a prostredia pre staticky typ (priemer cez cas. poz



**Obr. 19** Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 500ms a subjekty), errorbary znázorňujú štandardnú chybu.

Z Obr. 19 vidíme ,že pri 500ms dĺžke ISI sú výsledky lepšie na poslednom segmente pre statický typ. Pre synchronný typ presnosť klesá v po druhom čísle a pre asynchronný typ klesá po prvom čísle v prezentácii.

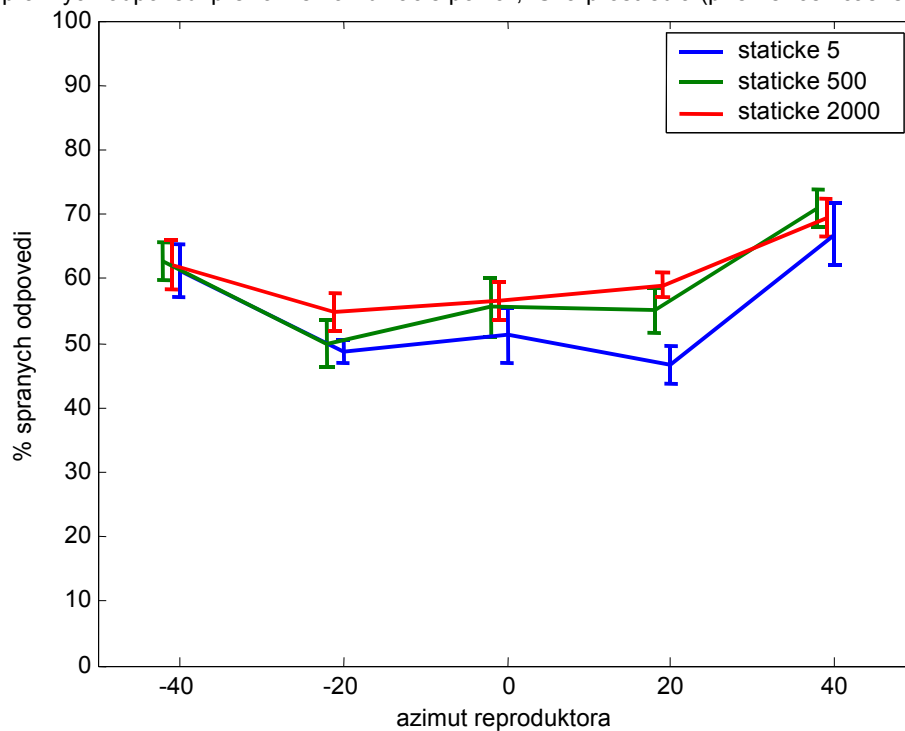
% správnych odpovedí pre rozne kombinácie ISI a prostredia pre staticky typ (priemer cez cas. poz



**Obr. 20** Percento správnych odpovedí v závislosti od segmentu prezentácie (pre ISI 2400ms a subjekty), errorbary znázorňujú štandardnú chybu.

Z Obr. 20 vidíme ,že výsledky sú oveľa lepšie ako v br. 18 a Obr. 19. Pre statický typ presnosť rastie s každým číslom v prezentácii. Pri dynamickými typoch je už rozdiel v presnosti odpovedí ,ale charakter krivky pre oba typy sa zachoval ako v Obr. 19 .

% správnych odpovedí pre rozne kombinacie pozícií, ISI a prostredia (priemer cez čas. segm., s ub.

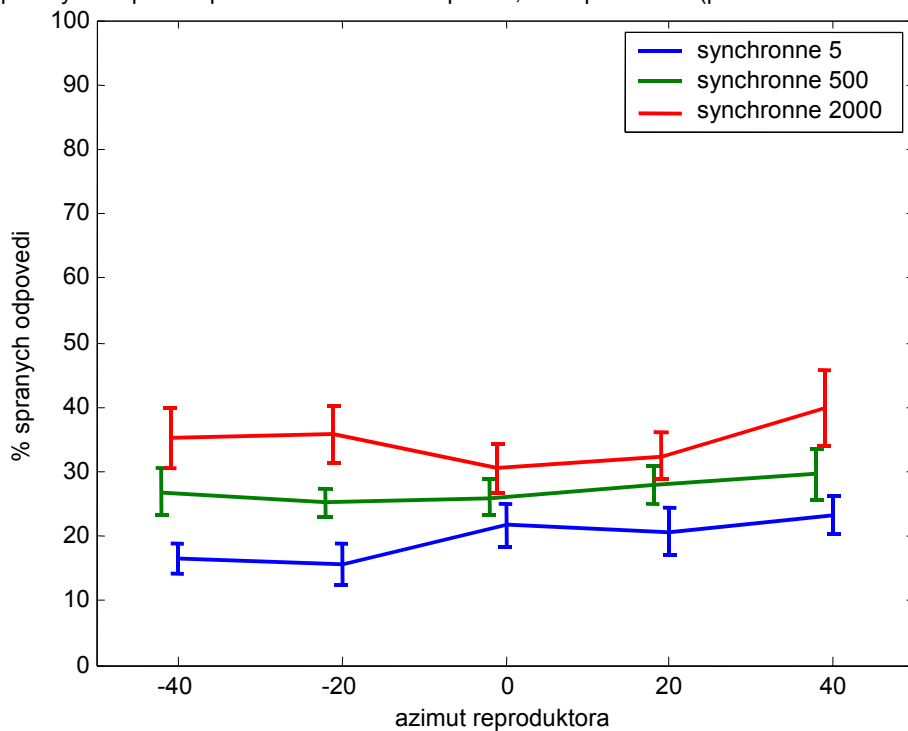


**Obr. 21** Percento správnych odpovedí v závislosti od pozície pre statický typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.

Obr. 18 ukazuje, že pri statickom type nemá pauza medzi solvami vplyv na úspešnosť odpovedí. Tiež môžeme vidieť, že úspešnosť odpovedí je vyššia pre krajné azimuty. ( $-40^\circ$  a  $40^\circ$ )

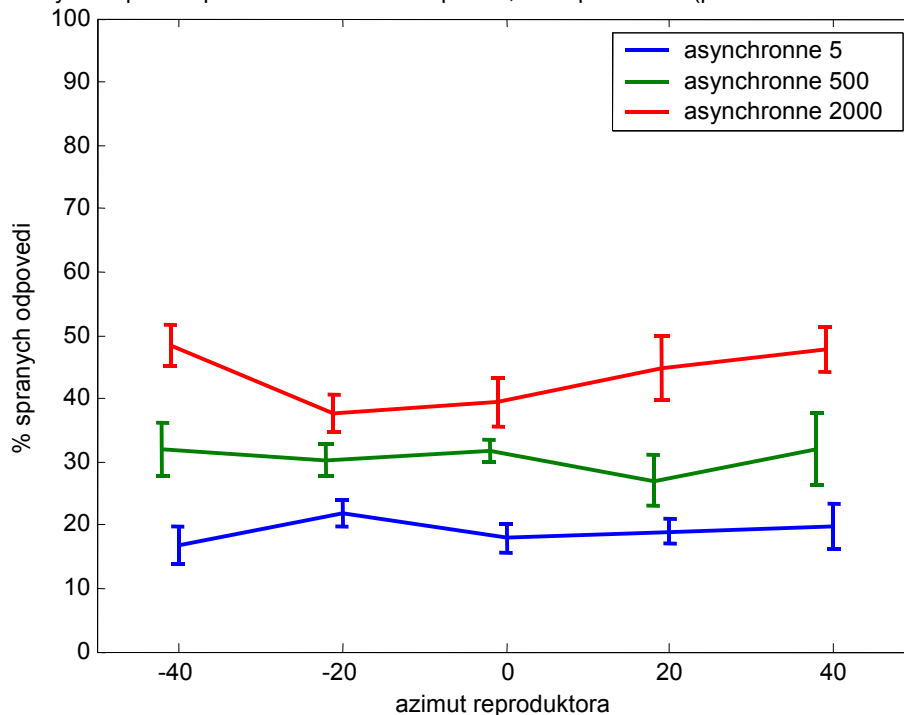
Na Obr. 19 je vidieť, že pauza medzi slovami výrazne ovplyvňuje presnosť a to miestami až o 20% .

% správnych odpovedí pre rôzne kombinácie pozícií, ISI a prostredia (priemer cez čas. segm., subj



**Obr. 22** Percento správnych odpovedí v závislosti od pozície pre synchronný typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.

% správnych odpovedí pre rôzne kombinácie pozícií, ISI a prostredia (priemer cez čas. segm., subj.)



**Obr. 23** Percento správnych odpovedí v závislosti od pozície pre asynchrónny typ (priemer cez časové segmenty a subjekty), errorbary znázorňujú štandardnú chybu.

Aj pri tomto type merania je úspešnosť vyššia s rastúcimi pauzami medzi slovami a to 25% - 30%. Pri pauze 2000ms je úspešnosť nepatrne vyššia na krajných azimutoch, tak ako sme to mohli sledovať pri statickom type.

### 3.1.8 Záver

#### Zhodnotenie hypotéz

- *Ak bude subjekt poznať polohu hovoriaceho, bude vedieť lepšie identifikovať, čo hovorí, ako keď ju nebude poznať.*

Hypotéza sa potvrdila hlavne pri statických kolách. Pri asynchrónnych kolách bola úspešnosť lepšia ako pri synchronných, ale nie výrazne.

- *Porozumenie sa zlepšuje s rastúcimi pauzami medzi slovami.*

Presnosť odpovedí bola viac-menej nezávislá od pauzy medzi slovami pri statických kolách, ale pri synchronných a asynchrónnych kolách bola úspešnosť lepšia

s rastúcimi pauzami medzi slovami, hyporéza sa potvrdila. Odporúčam v ďalších experimentoch používať viacero ISI ,ale nie pri statyckých typoch.

- *Porozumenie bude lepšie pre krajné polohy hovoriacich ako pre stredové (vychádzajúc z [16]).*

Hypotéza sa potvrdila hlavne pri statických prostrediach, pri synchronných a asynchronných už tento efekt nebol taký výrazný.

- *Ak za stimul ,ktorý bude určovať polohu cieľového hovoriaceho, vložíme pauzu ,tak percento úspešnosti bude vyššie.*

Táto hypotéza je pravdivá iba pri niektorých subjektoch a pri spriemerených hodnotách je percento úspešnosti vyššie približne len o 5%. Z toho vypláva ,že stimul ktorým sa odlišoval synchronný alebo asynchronný typ výrazne nepomohol pri lokalizácii čísla.

Výsledky tohto experimentu ukázali, že ak subjekt pozná polohu hovoriaceho lepšie identifikuje hovorené čísla od cieľového hovoriaceho ako keď sa poloha mení. Možno to vidieť na obrázkoch kde v statických prostrediach bola presnosť odpovedí oveľa vyššia ako v synchronných alebo asynchronných.

Presnosť odpovedí bola závislá aj od konkrétnej polohy, odkiaľ zaznel hlas, čo sa dokázalo hlavne v statických prostrediach, v ktorých mala krivka tvar „W“. Najvyššia presnosť bola v prípade, ak cieľový hlas zaznel z krajného reproduktora ( $-40^\circ$  alebo  $40^\circ$ ), pre azimuty  $-20^\circ$  a  $20^\circ$  presnosť klesla asi o 10% a v azimute  $0^\circ$  opäť stúpla o nepatrné percento. K nárastu presnosti na krajných azimutoch mohla prispieť , väčšia diferenciacia doznievania čísla do uši, takže subjekty mohli ľahšie lokalizovať a následne rozpoznať reprodukované číslo. Tiež k tejto zvýšenek presnosti mohol prispieť aj fakt ,že hlasy z týchto azimutov boli maskované len z jednej strany a tým pádom boli ľahšie rozpoznateľné. Keďže sa to neprejavilo v takej miere v synchronnom alebo asynchronnom prostredí, môžeme povedať, že presnosť bola ovplyvnená aj pozornosťou pri týchto krajných azimutoch. Vplyv pozície na presnosť odpovedí bol podobný ako v štúdiu [16], presnosť pri azimutoch  $-20^\circ$  a  $20^\circ$  tam však bola o niečo vyššia ako presnosť pri azimute  $0^\circ$ , pričom v tomto experimente je približne na rovnakej úrovni, až nižšia.

---

Z grafov je vidieť, že poradie slov má zplyv na presnosť odpovedí pri jednotlivých typoch merania. Ukázalo sa tiež, že na presnosť odpovedí nemá skoro žiaden vplyv dĺžka pauzy, pri statickom type, medzi jednotlivými slovami, čo je ovplyvnené tým, že naša schopnosť porozumieť reči je relatívne robusná voči rýchlosti hovorenia, ak nie je porozumenie miešané s ďalšími kognitívnymi úlohami. Pri synchronom a asynchrónnom type bola presnosť odpovedí vyššia s rastúcou pauzou medzi slovami.



## 4 Záver - zhrnutie experimentu

Experiment „cocktail party effect“ ukázal, že pozornosť je dôležitá pri vnímaní reči. Výsledné grafy ukázali, že ak poznáme polohu hovoriaceho v prostredí s viacerými hovoriacimi zlepšuje sa schopnosť porozumieť tomu, čo hovorí. Pri statickom type presnosť odpovedí nezávisí od dĺžky pauzy medzi slovami, (čiže aj od rýchlosti akou sú slová povedané). Tiež presnosť odpovedí nezávisí od toho, či je slovo na začiatku alebo na konci sekvencie slov, závisí však od polohy, odkiaľ slovo znie. Pri synchronnom a asynchronnom type dĺžka pauz medzi slovami, čiže rýchlosť a akou hovoríme, ovplyvňuje našu presnosť identifikovať cieľové slovo. Presnosť odpovedí závisí od toho, či je slovo na začiatku alebo na konci sekvencie slov, nezávisí však od polohy, odkiaľ slovo znie.

Na základe tohto experimentu vieme, že v ďalších experimentoch je potrebné správne zvoliť rôzne časové opozdenia medzi jednotlivými slovami. Tiež by do budúcnosti bolo zaujímavé detailnejšie sa zamerať na vplyv polohy reproduktorov pre jednotlivé typy.

Výsledky experimentov by sa mali do budúcnosti analyzovať pre každý hlas ktorý sa v experimente použil. Pre každý z vybraných hlasov zvlášť, tak aby sa dali hlasy s veľkou odlišnosťou (dominantnosťou) odseparovať, čo by mohlo viesť k ďalším hypotézam.

Vytvorenie ďalšieho typu prostredia, kde sa z dynamického v [20] vytvorili synchronny a asynchronny viedli k zaujímavým informáciám, hlavne pri roznej pauze medzi slovami. Obmenou v ďalšom experimente by mohol byť typ prostredia, ktorý by obsahoval niečo zo statického a niečo zo synchronného alebo asynchronného prostredia. Subjekty by v tomto prostredí dostali informáciu o tom, kadiaľ sa bude cieľový hovoriaci pohybovať (tým, že by sa z daných pozícií na začiatku predstavil). Mohlo by sa vyhodnocovať aj to, aké budú odpovede, ak hovoriaci bude preskakovať z jednej strany na druhú a aké ak pojde postupne zľava do prava alebo naopak. Tiež je zaujímavé ako by bola ovplyvnená naša vedomosť, ak by cieľový hlas nebol konštantný tj. Počas prezentácie by sa menil a subjekty by sa museli sústrediť iba na cue.

---

Čísla v anglickom jazyku boli pre každý subjekt zrozumiteľné, ale v nasledujúcich experimentoch by mohli subjekty použiť číslice v slovenskom jazyku, čo by mohlo ovplyvniť výsledky experimentov, keďže by to bolo pre subjekty prirodzenejšie.

## Zoznam použitej literatúry

- [1] KOPČO, N., *Výpočtová a kognitívna neuroveda*, prednášky k predmetu, 2007, dostupné na internete: <http://neuron-ai.tuke.sk/~kopco/kui440/>
- [2] SPIELER, *Attention*, prednáška dostupná na internete: <http://wwwhomes.doc.ic.ac.uk/~xh1/Referece/Others/a-visual-search-models.pdf>
- [3] ARONS, B. (1992), *A review of the Cocktail Party Effect*, Journal of the American Voice I/O Society, 12., 35-50.
- [4] CHERRY, E. C. (1953), *Some experiments on the recognition of speech, with one and with two ears*. Journal of Acoustic Society of America 25, 975--979.
- [5] BERNÁT M., *Sluchová priestorová pozornosť vyvolaná sluchovými a zrakovými podnetmi*, Diplomová práca, FEI TU KKUI, 2005
- [6] KOPČO, N. Úvod do neurovied, prednášky k predmetu, 2007, dostupné na internete: <http://neuron-ai.tuke.sk/~kopco/kui342/>
- [7] FRANCIS, G., *Sensory and perceptual processes*, prednášky k predmetu, dostupné na <http://www2.psych.purdue.edu/~gfrancis/Classes/PSY310/L31b.pdf>
- [8] BEAR, M.F., CONNORS, B.W., PARADISO M.A., *Neuroscience: exploring the brain (second edition)*, Lipincott Williams and Wilkins, Baltimore.
- [9] ANDOGA, R., KOPČO, N., *Mozgové štruktúry a mechanizmy riadenia strategickej pozornosti*, In: Kvasnička V., Pospíchal J., Kelemen J. Zborník z konferencie Kognícia a umelý život V, Smolenice, 2005
- [10] MCGURK, H., MACDONALD, J., (1976). *"Hearing lips and seeing voices"*, Nature 264, 746-748.
- [11] SPENCE, CH. *Audiovisual Multisensory Integration*, Acoustical Science and Technology Vol.: 28, No.: 2, 2007.
- [12] GOLDSTEIN, B., *Attention*, 2005, prezentácia dostupná na internete: <http://www2.una.edu/psychology/py385/goldstein%20chp4-1.ppt>
- [13] ANDOGA, R., BERNÁT, M., TOMORIOVÁ, B., KOPČO, N. (2006). *Spatial Modality-dependant attentional control in human sound localization*, Proceedings of the 10<sup>th</sup> International Conference on Cognitive and Neural Systems, Boston, MA, May 2006.
- [14] CUI Q., RAZAVI B., O'NEILL W., PAIGE G., *The influence of aging on Interactions Between Eye Position and Human Sound Localization*, poster prezentovaný na stretnutí "Association for Research in Otolaryngology", 2007
- [15] KITTERICK, P.T., SUMMERFIELD, A.Q., *The role of attention in the spatial perception of speech*, poster prezentovaný na stretnutí "Association for Research in Otolaryngology", 2007

- 
- [16] BEST, V., OZMERAL, E., SHINN-CUNNINGHAM, B., *Visually guided Attention Enhances Target Identification in a Complex Auditory Scene*, Journal of the Association for Research in Otolaryngology, 2007
- [17] Cherry C. (1953) „Some experiments on the recognition of speech, with one or two ears“ J. Acoust. Soc. Am. 25,975-979
- [18] Yost, 1992a, 1992b alebo segregácia zdroja zvuku (Bregman, 1990).
- [19] TOMORIOVA B. and KOPCO N ( 2008 ) Auditory Spatial Cuing for Speech Perception in Dynamic Multi-talker Environment. In: Proceedings of the 6th International Symposium on Applied Machine Intelligence and Informatics, Herlany, Slovakia, January 21 – 22, 2008
- [20] TOMORIOVA B. (2008) Unimodal and Crossmodal attentional control
- [21] [http://www.ausim3d.com/about/AuWeb\\_perception3.html](http://www.ausim3d.com/about/AuWeb_perception3.html)
- [22] <http://www.springerlink.com/content/fu751117221771xt/>

## Prílohy

- Príloha A: CD médium – diplomová práca v elektronickej podobe, prílohy v elektronickej podobe, zdrojové kódy a spustiteľná verzia programov, grafy a dáta k experimentu
- Príloha B: Používateľská príručka
- Príloha C: Systémová príručka
- Príloha D: Obrazová príloha