

**UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY**

**Význam priestorovej konceptualizácie pre znižovanie
kognitívnej záťaže v grafickom užívateľskom prostredí**

2012

Bc. Milan Uherčík

**UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY**



**Význam priestorovej konceptualizácie pre znižovanie
kognitívnej záťaže v grafickom užívateľskom prostredí**

Študijný program:

Kognitívna veda

Študijný odbor:

2503 Kognitívna veda

Školiace pracovisko:

Katedra aplikovanej informatiky

Vedúci diplomovej práce:

doc. PhDr. Ján Rybár, PhD.

Evidenčné číslo:



2012

Bc. Milan Uherčík



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Milan Uherčík
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.2.11. kognitívna veda
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Význam priestorovej konceptualizácie pre znižovanie kognitívnej záťaže v grafickom užívateľskom rozhraní

Cieľ:

1. Preskúmať problematiku produkcie nežiaducej kognitívnej záťaže v grafických užívateľských prostrediach, zhodnotiť vhodnosť metód jej merania a ich reálnu využiteľnosť.
2. Interpretovať problematiku kognitívnej záťaže v kontexte súčasných grafických užívateľských rozhraní (GUI) a interakcie človek-počítač (HCI) z hľadiska relevantných kognitívnovedných disciplín.
3. Vypracovať potenciálne riešenia a načrtnúť model kognitívne ergonomickej/efektívnej senzomotorickej interakcie pre GUI.

Anotácia: Práca analyzuje kľúčové problémy produkcie kognitívnej záťaže v HCI, interpretuje ich pomocou neuropsychologických modelov a teórií senzomotorickej interakcie a ponúka explicitnú formuláciu niekoľkých princípov kognitívnej ergonómie pre dizajn interakčných grafických užívateľských prostredí.

Vedúci: doc. PhDr. Ján Rybár, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky

Dátum zadania: 25.05.2010

Dátum schválenia: 16.10.2010

prof. RNDr. Pavol Zlatoš, PhD.
garant študijného programu

študent

vedúci práce

Čestné prehlásenie

Čestne prehlasujem, že som predloženú diplomovú prácu vypracoval samostatne s použitím uvedenej literatúry a pod vedením doc. PhDr. Jána Rybára, PhD. Týmto by som tiež rád poďakoval svojmu školiteľovi a blízkym za rady a podporu pri písaní tejto práce.

Milan Uherčík

Senec 25. máj 2012

Abstrakt

Práca analyzuje kľúčové problémy produkcie kognitívnej záťaže pri interakcii človek-počítač (HCI) v grafických užívateľských prostrediach (GUI) a interpretuje ich z neuropsychologického hľadiska. Vychádza z rozšírenia Baddeleyho modelu pracovnej pamäti a teórie paralelného spracovania vizuálneho signálu dorzálnou a ventrálnou neurálnou dráhou. Práca poukazuje na dôležitosť prispôsobovania GUI potrebám motorickej stránky interakcie v zmysle teórie stelesnenej kognície a popisuje niekoľko princípov priestorovej konceptualizácie redukujúcich kognitívnu záťaž v HCI.

Kľúčové slová: interakcia človek-počítač, grafické užívateľské prostredie, kognitívna záťaž, multimodálne teórie, Engelkamp, priestorová konceptualizácia, gesturálna interakcia, neuroergonómia.

Abstract

This paper analyses main problems which occur during cognitive load caused by human-computer interaction (HCI) in graphical user interface (GUI) and further interprets them from neuropsychological point of view. Our line of reasoning is based on an enhanced version of Baddeley's model of working memory and the two-streams hypothesis. We point out the importance of adjusting GUI to the motor skills used in the HCI, as described in the embodied cognition theory, and list principles of spatial conceptualization which reduce cognitive stress in HCI.

Keywords: Human-Computer Interaction, Graphical User-Interface, Cognitive Load, Multimodal theories, Engelkamp, Spatial Conceptualization, Spatialization, Gestural Interaction, Neuroergonomy.

Obsah

Úvod	4
1. Rozhranie človek-počítač.....	6
1.1 Užívateľské prostredia	6
1.2 Hodnotenie použiteľnosti	8
1.3 Užívateľská skúsenosť (UX)	10
1.4 Merateľné psychofyziologické prejavy a UX.....	11
1.5 Kognitívna záťaž	14
2. Pracovná pamäť a multimodálne teórie	15
2.1 Modely pracovnej pamäti	16
2.1.1 Baddeley-Hitchov model	16
2.1.2 Cowanov model	17
2.2 Pracovná pamäť vs. pozornosť	18
2.3. Multimodálne teórie a GUI	19
2.3.1 Paiviova teória duálneho kódovania.....	20
2.3.2 Engelkampova multimodálna teória.....	21
2.4 Benefity a úskalia multimodálneho kódovania	22
3. Hypotéza neuroergonomickej optimalizácie v GUI	24
3.1 Zhrnutie multimodálnych teórií a kognitívnej záťaže	24
3.2 Teória dorzálnej a ventrálnej dráhy.....	25
3.3 Formulácia hypotézy.....	26
4. Implicitné vizuálne prostriedky v GUI.....	30
4.1 Indikátory vnímania hĺbky.....	31
4.2 Vizuálne vodítka a Gestalt princípy.....	34
4.3 Užívateľské metaforý	40
4.4 Priestorová konceptualizácia v GUI	44
5. Motorická interakcia zo strany užívateľa.....	48
5.1. Ergonómia v HCI.....	49
5.2. Gestikulácia a gestúry	52
5.3. Gesturálna interakcia v HCI.....	55
6. Príklady aplikácii priestorovej konceptualizácie v kontexte historického vývoja ...	62
6.1 Post-WIMP a dotyková interakcia.....	66
6.2 NUI a konceptuálne riešenia pre GUI	67
6.3 Potenciál priestorovej konceptualizácie pre HCI	71
7. Záver.....	74
Bibliografia	77

Úvod

Pri interakcii človek - počítač (HCI), špeciálne v jej momentálne najpoužívanejšej forme, v grafickom užívateľskom rozhraní (GUI), človek aktívne využíva často len zlomok kognitívnych mechanizmov a procesov, ktoré bežne používa vo svojom prirodzenom životnom prostredí. Spomedzi perцепčných orgánov sa toto obmedzenie týka pri najbežnejšie používaných GUI hlavne redukcie získavania informácií formou vizuálnej percepcie, len s malým sprievodným podielom haptickej odozvy periférnych zariadení ako sú klávesnica a myš. Okrem zjavného obmedzenia perцепčných receptorov však pri GUI dochádza ešte aj k ďalším obmedzeniam. Zdanlivo obsiahle informačné spektrum vizuálnej percepcie je ďalej obmedzené na vnímanie sprostredkovaných informácií len z relatívne malej a staticky umiestnenej 2D zobrazovacej plochy. To v konečnom dôsledku znamená minimalizovanie ďalších prirodzených perцепčných kanálov ako vnímanie priestoru a hĺbky, bežne získavaných binokulárnym vnímaním a samostatným pohybom po priestore, s čím pochopiteľne úzko súvisia aj ďalšie obmedzenia vo forme fixovanej senzomotorickej koordinácie či obmedzenom rozsahu možností fyzickej interakcie s počítačom.

V súčasných bežne používaných GUI je vizuálna percepcia navyše jediným nosným kanálom pre tok sémantickej informácie. Kombináciou neprirodzenosti senzomotorickej interakcie HCI a šírkou informačného prietoku prenášaného vizuálnym kanálom tak môže ľahko dôjsť k nadmernej kognitívnej záťaži v podobe prehltenia pracovnej pamäti a pozornosti. Interferencie rôznych navigačných a sémantických prvkov, prílišná komplikovanosť či neprehľadnosť užívateľských prostredí sú častými problémami takejto interakcie. Jednou z hlavných problematík v oblasti HCI, je teda znižovanie celkovej záťaže užívateľa a zdokonaľovanie interakcie k jej čo najintuitívnejšej a najprirodzenejšej forme.

V mojej práci sa bližšie sústredím na pozadie produkcie kognitívnej záťaže v kontexte HCI a rozšírením Baddeleyho modelu pracovnej pamäti o teórie multimodálneho

kódovania, v práci popisujem model kognitívne ergonomickej senzomotorickej interakcie. Základom tohoto modelu je myšlienka rozkladu potenciálnej kognitívnej záťaže vhodným distribuovaním problematických procesov interakcie medzi paralelne pracujúce, neuralne nezávislé kortikálne regióny. Vychádzajúc z neuropsychologického pohľadu teórie dorzálnej a ventrálnej neurálnej dráhy poukazujem na potenciál integrovania priestorovej konceptualizácie do senzomotorickej interakcie v GUI.

1. Rozhranie človek-počítač

Prudký rozvoj technológií v oblasti úžitkovej elektroniky v posledných desiatkach rokov spôsobil, že pre veľkú časť produktívnej populácie sa stala výpočtová technika bežnou súčasťou ich života. Počítače sa stali nielen nástrojom nevyhnutným pre efektívne spravovanie nadrozmerných dát, ale aj virtuálnym životným prostredím, vrátane aspektu sociálnej a edukačnej interakcie. Je preto pochopiteľné, že v súčasnom, stále sa rozširujúcom rozsahu tejto interakcie, je samotná forma prostredia, ako aj jej interakčné prvky, predmetom intenzívneho zdokonaľovania nielen po stránke funkčnej, ale aj po stránke celkového užívateľského dojmu.

Interakcia medzi človekom a počítačom (ďalej len HCI, z angl. *Human-Computer Interaction*), je interdisciplinárny odbor zaoberajúci sa problematikou komunikácie medzi človekom (či už jednotlivcom alebo skupinami ľudí) a počítačom, pričom pod počítačom si môžeme predstaviť akýkoľvek interaktívny výpočtový systém.

1.1 Užívateľské prostredia

Jednou z ústredných tém, ktorou sa HCI zaoberá je užívateľské prostredie (tiež nazávané UI, z angl. User Interface), keďže tvorí komunikačnú bránu vstupov a výstupov medzi človekom a počítačom. Základnou snahou disciplíny HCI je snaha o zdokonaľovanie užívateľských prostredí, tak aby bola interakcia čo najintuitívnejšia a čo najviac prístupná aj pre menej skúsených alebo nových užívateľov.

Medzi najznámejšie užívateľské prostredia patrí textové a grafické užívateľské rozhranie. Napriek tomu, že **textové rozhranie** je historickým predchodcom grafického, svoje využitie nachádza v rôznych oblastiach expertných aplikácií aj v súčasnosti. Výhodou tohoto rozhrania je možnosť rýchleho dosiahnutia požadovanej akcie, pri minimálnej

záťaži koordináciou s polohovacími zariadeniami, či navigáciou v užívateľskom prostredí. Medzi nevýhody patrí relatívne veľký čas učenia apriórnych znalostí potrebných k plnohodnotnej interakcii, keďže ide o rozhranie zadávania vstupov v lexikálnej forme s presne vymedzenou syntaxou.

Grafické užívateľské rozhranie (ďalej len GUI, z angl. *Graphical User Interface*) je prirodzeným dôsledkom vývoja technológií a historickej snahy eliminovať neprístupnosť HCI textového rozhrania a priblížiť tak užívateľské prostredie aj menej skúseným užívateľom. Základnou črtou tohoto rozhrania je jeho možnosť ovládania pomocou interaktívnych grafických prvkov. Tieto prvky zväčša tvoria rozsiahlejší metaforizovaný systém ustálených vizuálnych formátov určujúcich ich užívateľský charakter (hyperlinky, tlačidlá, atď.). Pre fyzickú interakciu s GUI je zvyčajne využívané spolu s klávesnicou aj polohovacie zariadenie (myš, grafický tablet, trackball). Najčastejšie používané grafické užívateľské prostredia je možné rozdeliť do niekoľkých základných kategórií podľa ich charakteristických interakčných prvkov:

- **CLI** (*Command Line Interface*) – ide o bežné textové rozhranie využívajúce príkazový riadok ako užívateľský vstup a zvyšok zobrazovacej plochy pre príjem informácií komunikovaných zo strany počítača, pričom užívateľ ovláda prostredie primárne prostredníctvom klávesnice.
- **WIMP** (*Windows, Icons, Mouse, Pointer*) – je prostredie, ktorého základnými interakčnými prvkami sú okná, ikony, menu a kurzor. Toto prostredie využíva väčšina súčasných operačných systémov desktopových riešení osobných počítačov.
- **Webové rozhranie** – je graficky spracované prostredie hypertextových dokumentov navzájom prepojených klikateľnými odkazmi. V súčasnosti ide pravdepodobne o najčastejšie používané elektronické rozhranie pre získavanie informácií. Prostredie podlieha neustálym inováciám rozširovania funkčnej

stránky. Za jeho bežné charakteristiky možno považovať navigačnú centralizáciu domovskou stránkou, hierarchickú štruktúru podstránok, stabilné umiestnenie navigačného menu s klikateľnými odkazmi a textové polia.

- **Post-WIMP** – rozhranie je využívané prevažne v operačných systémoch prenosných zariadení (PC tablety a smartfóny). Ide o modifikované prostredie WIMP, upravené pre účely použitia v zariadeniach s dotykovými obrazovkami. Dotyková plocha tu nahrádza funkciu externého polohovacieho zariadenia a zjednocuje senzomotorickú interakciu do intuitívnej interakčnej plochy.

Intuitívnosť jednotlivých rozhraní je teda určená hlavne interakčnou charakteristikou kombinácie vstupov a výstupov a od čias textových užívateľských rozhraní prešla HCI výraznými zmenami smerom k užívateľskému pohodliu.

Na pokroku zdokonaľovania interakčnej stránky GUI má nepochybne významnú zásluhu testovanie ich použiteľnosti. Hlavným cieľom testovania použiteľnosti užívateľských prostredí je odstraňovanie neintuitívnych či konfliktných interakčných elementov, ktoré by mohli u užívateľa vyvolávať zbytočnú či nežiaducu kognitívnu záťaž.

1.2 Hodnotenie použiteľnosti

Prvými pokusmi o odborné testovanie a hodnotenie použiteľnosti užívateľského prostredia bolo uskutočnené spoločnosťou Xerox Palo Alto Research Center (PARC) už v roku 1981 pri vývoji počítača Xerox Star [78]. Išlo o jedného z prvých predchodcov súčasných osobných počítačov, ktorý podporoval napr. bitmapový grafický formát. Hodnotenie použiteľnosti v zmysle užívateľských prostredí nie je presne vymedzená disciplína a nemá dokonca ani štandardne akceptovanú definíciu. Ide skôr o súbor pragmaticky orientovaných metodológií ako overovať a zvyšovať mieru intuitívnosti a efektivity použitia produktu a schopnosti učiť sa s ním pracovať. Medzinárodný štandard ISO 9241-11 [49] rozlišuje pri hodnoteniach použiteľnosti tri hlavné aspekty, ktorými sú

účinnosť, výkonnosť a spokojnosť, no podľa priekopníka štúdií použiteľnosti Jakoba Nielsena [77] je nutné hodnotiť päť kvalitatívnych parametrov:

- **Schopnosť učenia sa** – resp. ako ľahko dokážu užívatelia vykonať základné úlohy pri prvom kontakte s užívateľským rozhraním,
- **Efektivita** – resp. ako rýchlo zvláda úlohy skúsený užívateľ,
- **Zapamätateľnosť** – ak sa užívateľ vráti k prostrediu po dlhšom čase, ako rýchlo nadobudne späť svoju obratnosť,
- **Chybovosť** – počet chýb, ktoré užívatelia pri testovaní urobia, ich periodicita a ako ľahko si dokáže užívateľ chybu svojpomocne uvedomiť a vyriešiť ju,
- **Spokojnosť** – ako prívetivo pôsobí prostredie na užívateľa,
- **Užitočnosť** – nakoľko splňuje dané prevedenie svoj požadovaný účel.

Napriek rozličným metodológiám, všetky štandardné štúdie použiteľnosti sa zhodujú v ich iteratívnom charaktere testovania, ktoré sa skladá z cyklu troch nasledovných fáz:

1. **Testovanie** – pre účely testovania konkrétneho rozhrania sa zväčša vyberá reprezentatívna vzorka reálnych cieľových užívateľov, ktorí dostávajú úlohy s rôznymi úrovňami zložitosti (v prípade testovania webového rozhrania to môžu byť rôzne druhy úloh od vyhľadania špecifickej informácie po uzatvorenie objednávky vopred určeného produktu). Počas testovania úloh sa zaznamenávajú údaje o postupe užívateľov.
2. **Analýza chýb a problémov** – v tejto fáze sa analyzujú dáta získané pri testovaní. Ide o identifikovanie a interpretáciu nečakaných užívateľských situácií (obchádzanie predpokladaných užívateľských postupov), sumarizáciu a lokalizáciu užívateľských chýb a hodnotenie reakčných časov.
3. **Návrh a implementácia riešení** – je individuálny proces v závislosti od konkrétneho prostredia a použitej metódy testovania.

Iterovaný charakter testovania teda znamená, že po identifikovaní a náprave problematických interakčných elementov nasleduje ďalšia fáza testovania, ktorá na jednej strane overí či mala implementácia riešení požadovaný efekt a na strane druhej, v prípade pozitívneho výsledku, môže ďalšie testovanie pomôcť odhaliť predtým neidentifikované problémy. Iterovanie sa v ideálnom prípade zvykne opakovať pokiaľ sa nedosiahne dopredu stanovená požadovaná miera parametrov.

Najrozšírenejšími a najčastejšie realizovanými štúdiami použiteľnosti sú v reálnej praxi štúdie komerčných web-stránok, nakoľko ide o veľmi obsiahle množstvo izolovaných a obsahovo rozdielnych prostredí. Štúdie popredných konzultantov použiteľnosti v tejto sfére ako Nielsen, Virzi [120], a Lewis [65] sa zhodli na závere, že pre jednu iteráciu testovania je efektívne postačujúci počet cca 5 participantov. Tvrdia, že piati užívatelia sú schopní odhaliť na jednu iteráciu štatisticky 80 % chýb použiteľnosti. Toto tvrdenie sa zakladá na štatistických výsledkoch množstva meraní a autori teórie interpretujú tento počet ako ekonomicky najefektívnejšiu možnosť pre testovanie. Tvrdenie obhajujú argumentom, že vzhľadom na iteratívny charakter testovania nemá zmysel používať vyšší počet užívateľov, keďže už pri druhom meraní s rovnakým počtom užívateľov je vysoká pravdepodobnosť pokryť aj zvyšných 20 % neodhalených chýb. Podľa Nielsenovho výskumu [77] je možné priblížiť sa k štatisticky 100 % počtu odhalených chýb až pri participácii 15 testovaných užívateľov pri prvej iterácii.

1.3 Užívateľská skúsenosť (UX)

Okrem štúdií použiteľnosti existuje aj širšie ponímanie tejto problematiky a tou je celková užívateľská skúsenosť, (ďalej len UX, z angl. *User Experience*). UX je vlastne štúdia použiteľnosti, ktorá v sebe zahŕňa nielen analýzu funkčného hľadiska užívateľského prostredia, ale aj celkové a čiastkové pocity užívateľa, teda emočnú rovinu prežívania práce s prostredím, ako aj niektoré sprievodné kognitívne mechanizmy.

Mnohé dodatkové informácie ako subjektívne pociťovanie nepohody, stresu, či nadmernej kognitívnej záťaže je možné merať rôznymi spôsobmi v závislosti od používanej metodológie. Prvou možnosťou je použitie kvalitatívnych metód dotazníkov alebo verbálne dopytovanie užívateľa počas úlohy, no s touto možnosťou sú spojené časté komplikácie v podobe nepresnosti takýchto výpovedí. Druhou možnosťou sú metódy kvantitatívne, akými sú behaviorálne psychofyzologické merania.

1.4 Merateľné psychofyzologické prejavy a UX

Psychofyziológia je disciplína študujúca vzájomné vzťahy medzi fyziologickými a psychologickými faktormi ľudského správania [105]. V porovnaní s kvalitatívnymi meraniami poskytujú psychofyzologické merania solídny štatistický a aj vedomostný základ, s ktorým je možné namerané výsledky efektívne porovnávať, vyhodnocovať a interpretovať. Pre účely štúdií HCI sa bežne používajú nasledovné typy meraní:

- **Sledovanie pohybu očí (*eye-tracking*) a pozornosť** – sledovaniu pohybu očí je v štúdiách použiteľnosti a UX kladený veľký význam, nakoľko sa z neho dá vyčítať množstvo dôležitých kognitívnych procesov ako správanie vizuálnej pozornosti vzhľadom na prezentovaný vizuálny stimul [51]. Pohyb očí býva najbežnejšie snímaný zariadeniami s infračervenou kamerou fixovanou priamo na hlave alebo pri niektorých prevedeniach kamerou umiestnenou na monitore. Pohyb očí je možné interpretovať na dvoch úrovniach: tzv. sakády (vedome neovplyvniteľné, veľmi rýchle a prerušované, ostro kmitajúce mikropohyby očí) a vedome cielená lokalizácia vizuálnej percepcie. Vyhodnocovanie meraní týchto pohybov je riešené obvykle buď formou tzv. „*heatmap*“ (pôvodný vizuálny stimul je topograficky prekrytý polotransparentnou vrstvou spojitej farebnej škály určujúcej dĺžku a miesto lokalizácie pohľadov), alebo formou tzv. „*gaze flow*“ (pôvodný vizuálny stimul je prekrytý chronologicky očíslovanými sakádami,

prepojenými líniami postupnosti ich pohybov). Pri niektorých sofistikovanejších eye-tracking metódach je možné zaznamenávať aj kontrakcie zrenice. Okrem reaktívneho zužovania pri vyššej a rozširovania pri vyššej svetelnej intenzite vizuálneho stimulu, sa priemer zrenice tiež zväčšuje pri náhlom zvýšení kognitívnej záťaže [43], pocite vzrušenia [104] alebo pri zvýšení záujmu užívateľa [52], pričom bolo preukázaný aj vplyv rozličných úrovní kognitívnej záťaže pri HCI na mieru rozširovania zrenice [47].



Obr. č. 1: Eye-tracking vizualizácie heatmap a gaze flow [44].

- Aktivita tvárového svalstva a emócie** – niektoré zmeny v emočnom prežívaní užívateľa je možné pozorovať ako mimické prejavy formou samovoľných kontrakcií svalovej hmoty na tvári [50]. Kontrakcie sú spúšťané elektrickými impulzami merateľnými pomocou elektromyografie (EMG), alebo zložitejšou a nepresnejšou metódou analýzy videozáznamu algoritmami počítačového videnia. Pre účely meraní UX sa jedná konkrétne o tri svalové skupiny: *corrugator supercilii* (lokalizované na začiatkoch obočia a ich kontrakciám je prisudzovaný výraz mračenia sa a negatívne emócie), *musculus zygomaticus* (lokalizované na končekoch pier a ich kontrakciám je prisudzovaný výraz úsmevu a pozitívnych emócií) a *orbicularis oculi* (lokalizované pod očnými viečkami a pri súčasných kontrakciách so zygomatickým svalstvom sa predpokladá, že užívateľ prežíva stav tzv. „duchenne“ úsmevu, ktorý je považovaný za prejav emócie šťastia) [105].

- **Produkcia potu a stres** – meraním odporu kože je možné sledovať mieru produkciu potu, o ktorom sa predpokladá, že reflektuje subjektívnu náročnosť úlohy z pohľadu užívateľa [87]. Nevýhodou tohoto merania je však pomalá reakcia tela a ďalšie súvisiace komplikácie. Rýchlosť fyziologickej odozvy na pokožke je cca až 6 sekúnd a presná dĺžka tohoto času navyše individuálne kolíše podľa jednotlivca, takže je obtiažne s istotou identifikovať presný moment vybudenia. Pri použití tejto metódy sa teda zvykne porovnávať celková produkcia potu počas testovania konkrétnej úlohy (ideálne trvajúcej nad 10 sekúnd) voči priemernej hodnote [105].
- **Variabilita počtu úderov srdca za minútu a pozornosť** - výskum v tejto oblasti poukazuje na to, že pri sústredenom vnímaní externého stimulu sa u pozorovateľa znižuje frekvencia úderov srdca [122]. Zvýšenie frekvencie je pozorovateľné naopak pri vnútornom sústredení pozornosti, akou je zvýšená mentálna aktivita spojená napr. s matematickými úlohami [69].

Existuje aj mnoho ďalších behaviorálnych príznakov ako meranie intenzity a frekvencie respirácie alebo sily ktorou užívateľ drží polohovacie zariadenie, no vzhľadom na náročnosť a relatívne slabú výpovednú hodnotu sa tieto metódy v praxi nepoužívajú.

Štúdie použiteľnosti, UX, ako aj niektoré experimentálne behaviorálne štúdie HCI sa formálne radia pod rámec disciplíny nazývanej kognitívna ergonómia. **Kognitívna ergonómia** je disciplína zaoberajúca sa štúdiou kognitívnych mechanizmov človeka v pracovnom prostredí vo vzťahu k technologickému designu. Na rozdiel od klasickej ergonomie, kognitívna ergonómia skúma interakčné procesy primárne z psychologického hľadiska. Kľúčovou snahou tejto disciplíny v HCI je zdokonaľovanie užívateľského rozhrania k čo nejefektívnejšej a naintuitívnejšej forme interakcie a k minimalizovaniu nežiadúceho javu kognitívnej záťaže.

1.5 Kognitívna záťaž

Štúdie kognitívnej záťaže siahajú až k historickým koreňom kognitívnej vedy, konkrétne k práci popisujúcej obmedzenie pracovnej pamäti G. A. Millera a fenoménu “sedem plus mínus dva” [74]. Tento fenomén popisuje obmedzenie pracovnej pamäti človeka popísaním maximálneho počtu nezávislých diskretných prvkov, ktoré je človek schopný súčasne udržať v krátkodobej pamäti. Niektoré výskumy však tvrdia, že toto číslo je zrejme nižšie a pohybuje sa medzi tromi až štyrmi prvkami [28]. Údaje sa však v pamäti môžu zlučovať do tzv. „*chunks*“, v ktorých je možné uchopiť zložitejšie mentálne schémy a týmto spôsobom limit zdanlivo navýšiť [16]. Teória kognitívnej záťaže J. Swellera [110] umožňuje poukázať na vplyv vyhotovenia informačného dizajnu na efektivitu pracovnej pamäti [15], kedy napríklad v prípade nevhodne navrhnutých interferujúcich informačných elementov môže nastať kognitívne preťaženie. Kognitívna záťaž má tiež výrazný vplyv na celkovú pozornosť, ktorá sa pri experimentoch v reálnych prostrediach prejavuje zvýšenými reakčnými časmi, ako aj chybovosťou. V prípadoch nadmernej záťaže, pracovná pamäť nie je schopná prijímať a spracovávať ďalšie informácie. Kognitívnu záťaž je možné bližšie rozdeliť na tri zložky [15]:

Vnútoraná (*intristic*) záťaž - popisuje obtiažnosť spracovania informácie alebo úlohy, teda jej kognitívnu komplexitu (teda napr. rozdiel náročnosti riešenia jednoduchšej aritmetickej úlohy a diferenciálnej rovnice). Rozkladaním do „*chunks*“ je však možné túto úroveň záťaže znižovať.

Vonkajšia (*extraneous*) záťaž - je záťaž, ktorá je dôsledkom rozhrania akým je informácia prezentovaná alebo štrukturovaná. Túto záťaž produkujú napríklad nevhodne navrhnuté, neprehľadné informačné prostredia s perцепčnými interferenciami. Spolu s vnútornou záťažou tvorí aditívny systém, ktorý určuje celkovú náročnosť spracovania úlohy v konkrétnom prostredí.

Konceptuálna (*germane*) záťaž - je žiaduci typ záťaže, potrebný pre akvizíciu informácie obsiahnutej v prostredí. Ide o súčasť procesu mentálneho spracovania, ktoré vedie k pochopeniu – teda k pragmaticky orientovanému spojeniu novej informácie s už existujúcim mentálnym modelom či konceptom.

Okrem vyššie popísaného perцепčno-kognitívneho spracovávanía informácie môže kognitívna záťaž súvisieť aj so sústredenou vizuomotorickou koncentráciou alebo koordináciou pohybov. Zvýšená záťaž spojená s komplexnejším spracovaním úloh produkuje navyše negatívne emočné vzrušenie (úzkosť) [55], čo môže mať za následok averziu či nechúť užívateľa voči prostrediu. Vyššie popísané zaťaženie jednotlivých zložiek je možné pozorovať či hodnotiť aj vďaka už spomínaným psychofyziológickým meraniam (pri náhlom vzraste vnútornej záťaže, napríklad meraním relatívneho rozšírenia zrenice alebo zmenou frekvencie úderov srdca, a pri vonkajšej napr. subjektívnym vyhodnocovaním charakteristických pohybov a fixácii očí vzhľadom na vizuálny stimul).

Hoci mnohé experimenty potvrdili predpoklad obmedzenú kapacitu pracovnej pamäte pre súčasné spracovavanie informácií, ukázalo sa, že efektivita sa u jednotlivcov silne líši vzhľadom na typ úlohy [97].

2. Pracovná pamäť a multimodálne teórie

V niektorých teoretických štúdiách sa často zamieňa termín krátkodobej pamäti s termínom pracovnej pamäti. Je dôležité zdôrazniť rozdiel medzi týmito dvoma podobnými označeniami, nakoľko krátkodobá pamäť je skôr „pasívny“ teoretický koncept spoločný pre neurovedy a kognitívnu psychológiu, pričom termín pracovnej pamäte je skôr konceptom kognitívnej psychológie a popisuje funkčné prepojenie s aktívnou manipuláciou s informáciami. Pracovnú pamäť by teda pre jej lepšie porozumenie vystihoval termín pracovnej pozornosti.

2.1 Modely pracovnej pamäti

2.1.1 Baddeley-Hitchov model

Baddeley a Hitch [9] popísali svoj tripartitný model v roku 1974 ako alternatívu k modelu pamäte Atkinsona a Shiffrina z roku 1968 [6]. Pôvodný model Atkinsona a Shiffrina, ktorý obnášal sériové prepojenie senzorickej, krátkodobej a dlhodobej pamäte, nahradili paralelizovaným modelom obsahujúcim dve podriadené zložky - fonologickú slučku a vizuálno-priestorový náčrtník, o ktorých využití rozhodovala hierarchicky nadradená centrálna exekutíva. Neskôr k dvom podriadeným zložkám pridal Baddeley aj epizodickú vyrovnávaciu pamäť [11].



Obr. č. 2: Baddeleyho model pracovnej pamäti.

Úloha **centrálnej exekutívy** spočíva v koordinovaní informácií do špecifických podriadených zložiek a v dohliadaní na všetky aktivity pracovnej pamäti. Pacienti s deficitom centrálnej exekutívy majú napríklad ťažkosti pri úlohách, ktoré vyžadujú zameranie sa na dôležitú informáciu (napr. úlohy vytvárania konceptov Wisconsinského testu triedenia kariet) [42].

Fonologická slučka plní funkciu krátkodobého úložiska zvukovej časti verbálnej informácie. Zahŕňa v sebe krátkodobú fonologickú a artikulačnú pamäť. Fonologická pamäť udrží zvukový materiál na krátky čas, po ktorom by sa materiál z pamäti stratil. Opakovaním je možné tento čas predĺžiť. Pacienti s poruchami fonologickej pamäti majú ťažkosti vybaviť si hovorené slová podobne, ako spomenúť si na ne aj v prípade

vnútorného opakovania. Existujú navyše štúdie poukazujúce zároveň aj na ich častú oslabenú vizuálnu pamäť [35].

Vizuálno-priestorový náčrtník dočasne zadržiava obsah určený na spracovávanie vizuálnych a priestorových typov informácií, pričom pod vizuálnou zložkou sa rozumejú parametre ako farba alebo tvar, a pod priestorovou zložkou rozmiestnenie, poloha a priestorové vzťahy elementov informácie. Toto rozdelenie podporujú nie len pozorovania behaviorálnych psychologických experimentov, ale aj poznatky neuropsychológie, ktoré poukazujú na funkčné zameranie dorzálnaj a ventrálnej dráhy vizuálnej informácie ako na "kde" a "čo" zložku [40].

Zložku **epizodickej pamäti** doplnil Baddeley do svojho modelu až v roku 2000, vzhľadom na chýbajúci článok popisujúci špeciálny druh neverbálnej pamäte pre sekvenčné a chronologické súvislosti (napríklad dej filmu, postup práce, atď.). Jednou z hlavných príčin začlenenia epizodickej zložky do modelu bolo tiež pozorovanie pacientov s amnéziou, ktorí neboli schopní kódovania nových poznatkov do dlhodobej pamäte, no nemali žiadne problémy s pamäťou krátkodobou, pričom išlo často o oveľa väčší rozsah sekvenčnej informácie, ako by bolo možné spracovať fonologickou slučkou [10].

2.1.2 Cowanov model

Alternatívny pohľad na pracovnú pamäť ponúka Cowanov model. **Cowan** nevníma pracovnú pamäť ako samostatný pamäťový systém, ale ako súčasť dlhodobej pamäte [20]. Predstavuje pozornosťný systém s kapacitou cca štyroch pamäťových reprezentácií („*chunks*“) ako jeho hlavnú črtu, pričom reprezentácie držané v pracovnej pamäti majú predstavovať istú podmnožinu reprezentácií z pamäti dlhodobej. Rozlišuje medzi dvoma úrovňami. Prvá úroveň sa skladá z aktivizovaných reprezentácií, ktorých počet nijako neobmedzuje. Druhá úroveň, ktorú Cowan popisuje, je zameranie pozornosti, pričom na rozdiel od Millera hovorí o limite cca štyroch aktivovaných reprezentácií.

Cowanov model rozšíril pridaním tretej úrovne **Oberauer** [82], ktorý ňou prezentuje ešte užšie, vedomé sústredenie pozornosti pracovnej pamäti. Opodstatnenosť tejto úrovne odôvodňuje tým, že človek pri konkrétnych mentálnych operáciách dokáže operovať súčasne iba s jedným zo spomínaných štyroch reprezentácií (demonštruje to príkladom s prirátavaním čísla 2 ku každej z cifier ľubovoľného štvorciferného čísla v pamäti, kde je väčšina ľudí nútená pri pripočítavaní mentálne prechádzať po jednej cifre) [56].

2.2 Pracovná pamäť vs. pozornosť

Napriek zdanlivej opozícii Baddeleyho a Cowanovho modelu, sa dá Cowanov model čiastočne aplikovať do paralelného systému exekutívy a podriadených zložiek pamäti. V prípade, že akceptujeme spomínaný funkčný pohľad na pracovnú pamäť ako na pracovnú pozornosť, Cowanov model v Oberauerovom dodatku bližšie špecifikuje Baddeleyho (svojim symbolistickým vyjadrením relatívne izolovaný model), v spôsobe akým jednotlivé podriadené zložky pamäte kooperujú, a to vzhľadom na *top-down* a *bottom-up* funkcie pozornosti. Koncept pracovnej pamäti je teda praktické používať pokiaľ sledujeme jej kvantitatívne dôsledky (napr. pri behaviorálnych experimentoch sledujúcich jej kapacitu), no pokiaľ by mal tento model slúžiť na funkcionálny popis interakcie z globálnejšieho pohľadu ďalších kognitívnych mechanizmov, je možno pragmatickejšie sa na koncept pozeráť ako na “pracovnú pozornosť”.

Kognitívna záťaž sa dá v ponímaní “pracovnej pozornosti” ďalej interpretovať ako prehltenie kapacít nielen sústredenej (*single-chunk*, vedomej) pozornosti, ktorej obsah sa neustále sekvenčným charakterom mení, ale hlavne ako prehltenie úrovne širšej pozornosti (Cowanových 4 „*chunks*“, resp. Millerových 7+-2), ktorá je schopná reagovať aj na ďalšie, vonkajšie – perцепčné (*bottom-up*) podnety alebo vnútorné – mentálne (*top-down*) podnety. Kognitívna záťaž má teda pochopiteľný dopad nielen na schopnosť paralelného spracovávanía, ale i vnímania periférnych informácií, keďže pozornostné

mechanizmy (v Baddeleyho modeli reprezentované vágnejším pojmom centrálnej exekutívy) priamo determinujú obsah nachádzajúci sa aktuálne v pracovnej pamäti.

V prípade užívateľských rozhraní sa užívateľ bežne dostáva do situácií, kedy je nútený prispôbiť sa rýchlym zmenám, čeliť informačným interferenciám alebo veľkému kvantu informácií, ktoré prichádzajúcich z prostredia. Pri jednoduchších implicitnejších úlohách sa ľudský mozog snaží informácie spracovávať paralelne, no úlohy, ktoré pre svoje vykonanie vyžadujú väčšie množstvo súvisiacich faktorov, ľahko vyčerpajú paralelné kapacity pracovnej pamäte, resp. pozornosti, a sú preto rozkladané na menšie podúlohy, zvládnuteľné sekvenčným spracovávaním, čo sa prejavuje zvýšenými reakčnými časmi [40].

2.3. Multimodálne teórie a GUI

Nečakaná potreba riešenia náročnejších úloh alebo sémantických interferencií v GUI môže byť často doprevádzaná frustráciou alebo stresom, čo navyše zhoršuje šance aktivácie konceptuálnej (*germane*) zložky kognitívnej záťaže, a teda aj šance transformácie inej ako emocionálnej užívateľskej skúsenosti do dlhodobej pamäti. Problematiku efektívnej akvizície znalostí a informácií, popisujú princípy tzv. inštruktážneho dizajnu (ďalej len ISD, z angl. *Instructional Systems Design*) [63]. História ISD siaha až ku obdobiu druhej svetovej vojny, kedy americká armáda potrebovala rýchlo naučiť veľké množstvo civilistov, zložitým technicky orientovaným úloham. Základným teoretickým motívom boli vtedy štúdie B. F. Skinnera o operantnom podmieňovaní [101] kombinované s cieľným rozkladaním zložitých úkonov na podproblémy. ISD sa postupne rozširovali o nové prostriedky inštruktážnych foriem ako multimediálne prezentácie a rôzne interaktívne prostredia GUI (napr. webové portály), ako aj o teóriu kognitívnej záťaže, kde fáza mentálneho porozumenia, tiež popisovaná konceptuálnou (*germane*) kognitívnou záťažou, tvorí v zmysle konštruktivistických teórií,

na ktorých inštruktážny dizajn stavia, nevyhnutný predstupeň k fáze akvizície a teda k finálnemu osvojeniu si prijímanej informácie. Pre problematiku GUI hrá navyše multimediálna forma inštruktážneho dizajnu fundamentálnu úlohu, keďže zrozumiteľný spôsob distribuovania informácií je jednou z najdôležitejších požadovaných vlastností pre perцепčnú zložku HCI interakcie. Efektívne formy multimediálnej prezentácie informácií z pohľadu perцепcie najlepšie teoreticky popisujú tzv. multimodálne teórie.

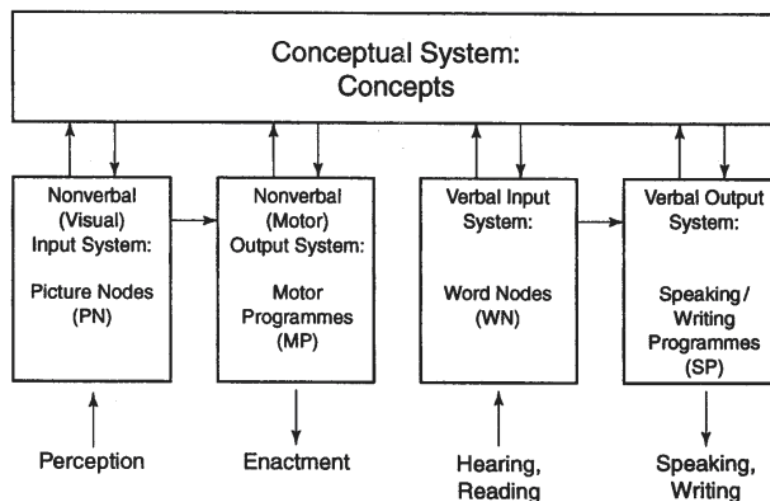
2.3.1 Paiviova teória duálneho kódovania

Paiviova duálna teória kódovania [69] predpokladá, že vizuálne a verbálne vnemy sú spracovávané dvoma samostatnými informačnými kanálmi kódovanými do pamäte rozdielnym spôsobom a každý informačný kanál takto vytvára oddelenú reprezentáciu prichádzajúcej informácie a iný samostatný pamäťový záznam. Paivio popisuje jeden informačný kanál pamäťového kódu založený na vizualnej predstavivosti (analogický) tiež nazývaný '*imagen*' a jeden verbálny (symbolický) kód nazývaný '*logogen*' [106]. Poukazuje na to, že vo väčšine prípadov sa na kognitívnom spracovávaní podieľajú oba informačno-pamäťové kanály súčasne, pričom špecifikosťou verbálnych vnemov je ich sekvenčné kódovanie. Ďalšie experimenty [64] poukázali tiež na to, že obrazotvornejšie slová sú ľahšie zapamätateľné a mentálne uchopiteľné ako abstraktné pojmy, nakoľko podľa Paiviovej teórie existuje predpoklad, že obrazotvorný verbálny vnem prechádza nielen verbálnym, ale aj vizuálnym informačným kanálom. Podľa Paivia, ľudia bežne využívajú analogickú zložku predstavivosti (*imageny*) pre úlohy kódovania slov do pamäte, čo im umožňuje ľahšie pamätanie, ako izolované používanie niektorých ich abstraktných seémantických reprezentácií. Príkladom podporujúcim toto tvrdenie je aj experiment Lee Brooksa [14]. Brooks potvrdil, že v kombinovaných úlohách, v ktorých boli participanti vystavení interferencii vizuálnej predstavivosti v porovnaní s rôznymi variantami neinterferujúcich úloh (úlohy, ktoré používali paralelné spracovanie odlišných informačných kanálov) boli namerané výrazne horšie výsledky, čo potvrdzuje paralelnosť spracovania duálnej teórie kódovania.

2.3.2 Engelkampova multimodálna teória

Paiviova teória duálneho kódovania síce hovorila o samostatných pamäťových kódach informačných kanálov, no jej zameranie bolo sústredené len na verbálnu a obrazovú informáciu. Multimodálna teória prezentovaná Engelkampom [25] popisuje existenciu ďalšieho pamäťového kódu - motorickej akcie. Mentálna reprezentácia motorických akcií je v kognitívnych vedách čoraz väčšmi študovanou oblasťou. **Teória stelesnenej kognície (*embodied cognition*)** [124] hovorí o významnom postavení mentálnych konceptov spojených s telesným prežívaním, ako aj o prepojení obrazových schém s dynamickými, mentálnymi a senzomotorickými akciami (napríklad gestikulácia, mimické prejavy atď.). Engelkamp sformuloval svoju multimodálnu teóriu ako výsledok empirického výskumu s rozsiahlymi variáciami experimentov. Vo svojich štúdiách poukázal na zistenia, že textom prezentované slovné frázy popisujúce motorickú akciu, ako napríklad "*bend the wire*" (ohnúť drôt) alebo "*nod your head*" (kývnuť hlavou), majú podobne ako v prípade Paiviovyh experimentov pri obrazotvorných verbálnych stimuloch, lepšie výsledky ich zapamätávania voči stimulom absentujúcim motorickú asociáciu. Rovnako sa výsledky potvrdili aj pri modifikácii experimentu prezentovaním verbálnych – zvukových fráz doprevádzaných ich fyzickou reprodukciou zo strany participantov a to či už virtuálnym, gesturálnym výstupom s virtuálnym znázorňovaním jednoduchých objektov alebo telesnou akciou znázorňujúcou akciu danej frázy.

Engelkampova multimodálna teória rozlišuje medzi dvoma modálne-špecifickými vstupnými a výstupnými systémami. Dva vstupné systémy obsahujú vizuálny a verbálny systém, ktoré sú základom aj pre Baddeleyho model pracovnej pamäte a Paiviovu duálnu teóriu kódovania. Pridanou hodnotou tohoto modelu je však popis vzťahu verbálneho a vizuálneho kanálu, kde Engelkamp demonštruje spôsob posilňovania stimulu cez konceptuálnu úroveň použitím špecifickej kombinácie informačných kanálov (viď. obr. č. 3).



Obr.č.3: Engelkampov model multimodálnej teórie [25].

Z hľadiska posilnenia senzomotorickej pamäte, je motorické znázorňovanie fráz na rozdiel od jednoduchého verbálneho výstupu výhodnejšie, keďže vyžaduje prevod na motorickú akciu isté sémantické spracovanie (prechod cez konceptuálnu úroveň) a teda aj určitú úroveň pochopenia a pamätania si prezentovanej frázy [103]. Spracovanie akcií, ktoré zahŕňajú prechod cez konceptuálnu úroveň, však pochopiteľne vyžadujú relatívne väčšiu kognitívnu záťaž a súvisiace reakčné časy v porovnaní s potenciálne reaktívnymi typmi ako je napr. plne verbálny prechod alebo vizuomotorická reakcia.

2.4 Benefity a úskalia multimodálneho kódovania

Potenciálny problém s koordináciou informačných kanálov cez konceptuálnu vrstvu môže vzniknúť pri presýtení pracovnej pamäte početnými mentálnymi operáciami. Autor teórie kognitívnej záťaže J. Sweller popisuje dve možnosti ako toto presýtenie znižovať:

1. **Automatickým spracovávaním** – podľa Engelkampovho modelu možnosťou spracovania informácie v rámci dvojíc orálno-sluchového a senzomotorického

kanálu. Automatické spracovávanie informácií vyžaduje menej pracovnej pamäte a jej kapacita tak môže byť využitá pre dôležitejšie úlohy či podúlohy.

2. **Akvizíciou mentálnych schém** – schémy sú organizované štruktúry vedomostí a zručností, ktoré zvyšujú potenciálne množstvo informácií v pracovnej pamäti zhlukovaním do spomínaných „*chunks*“. Pomocou rôznych metafor, či mentálnych konceptualizácií sa tak zvyšuje šanca uchopenia komplexnejších informácií.

Sweller vo svojej teórii rozlišuje primárne medzi vnútornou a vonkajšou kognitívnou záťažou. Vonkajšia záťaž je pre navrhovanie grafických užívateľských rozhraní obzvlášť dôležitým faktorom. Pri nevhodnom rozvrhnutí prvkov väčšieho vizuálneho celku a často nevyhnutnej popisnej textovej inštrukčnej, môže ľahko dôjsť k **efektu rozdvojenej pozornosti (*split-attention effect*)**. Tento nežiadúci efekt nastáva napríklad pri oddelenom použití diagramu a popisnej inštrukčnej legendy interpretujúcej ten istý informačný obsah inou vizuálnou formou. Problém teda súčasne nastáva nielen v konceptuálnej vrstve záťaže, kde v prípade nejasnosti reprezentácie grafickej alebo textovej zložky súčasne vznikajú izolovane dve rastúce akvizície schém bez jednotného korešpondujúceho konceptuálneho prepojenia, ale aj tým, že je na túto akvizíciu využívaný rovnaký (vizuálny) informačný kanál, čím sa zvyšuje pravdepodobnosť prehltenia pracovnej pamäti [111].

Druhým častým problémom je obdobná situácia, kde jedna z vizuálnych zložiek, napr. spomínaný grafický diagram, je samostatne plne zrozumiteľná mentálna schéma nevyžadujúca inštrukčnú, no je doplnená o nekorešpondujúci alebo neinformatívny popisný text. Tento prípad je nazývaný ako tzv. **efekt nadbytočnosti (*redundancy effect*)**, kedy doplňujúca informácia nežiaduco zaťažuje kognitívne kapacity a znemožňuje žiaducu aktivitu záťaže konceptuálnej (*germane CL*) [112].

Používanie viacerých informačných kanálov v užívateľských a inštruktážnych prostrediach má teda dve hlavné výhody. Prvou je výhoda posilnenej pamäťovej stopy viacerými kódmi. Pri GUI HCI to však nemusí nutne znamenať kombináciu vizuálnej a zvukovej percepcie, ale môže ísť tiež napríklad o kombináciu haptickej, motorickej a vizuálnej zložky gestúr zadávaných polohovacím zariadením. Druhou výhodou použitia viacerých pamäťových kódov je, že dokážu efektívne zabrániť interferenciám, akými sú efekt rozdelenej pozornosti a efekt nadbytočnej informácie.

3. Hypotéza neuroergonomickej optimalizácie v GUI

V nasledujúcich častiach zhrniem predchádzajúce modely a neuropsychologické štúdie do celistvej hypotézy modelu, ktorého zámerom je explicitná formulácia jadra základných princípov efektívnych GUI z hľadiska neuropsychologických poznatkov a kognitívnej ergonómie.

Implementácia poznatkov z neurofyziológických štúdií pri HCI je v súčasnosti novovznikajúcou a čerstvo otvorenou disciplínou, označovanou tiež ako **neuroergonómia**. Od kognitívnej ergonómie, ktorá sa primárne zaoberá psychologickými aspektami interakcie človeka s pracovným prostredím, prináša neuroergonómia v tejto problematike aj nový pohľad z hľadiska poznatkov neurálnej fyziológie či neuropsychológie.

Táto kapitola je rozdelená kvôli prehľadnosti do troch častí. V prvej časti zhrniem kľúčovú problematiku vyššie popísaných modelov a teórií. V druhej časti stručne popíšem jej neuropsychologické pozadie a v tretej časti sa pokúsím ponúknuť explicitnú formuláciu vyplývajúcich princípov kognitívne ergonomickej interakcie v GUI a poukázať tak na určitý model efektívnej senzomotorickej interakcie.

3.1 Zhrnutie multimodálnych teórií a kognitívnej záťaže

Teória duálneho kódovania popisuje relatívnu nezávislosť ukladania pamäťových – perцепčných kódov, ktorú Baddeley a Hitch rozširuje vo svojom modeli pracovnej pamäti pomocou fonologickej slučky, vizuálno-priestorového náčrtníka a epizodickej pamäti. Swellerova teória kognitívnej záťaže jeho model dopĺňa o parameter pamäťového zahltenia, ktorý determinuje rýchlosť spracovávania informácie, ako aj vyťaženie pozornosťných mechanizmov voči okolitým podnetom. Štúdie multimodálnych teórií kognitívnej záťaže tiež potvrdili, že jej intenzita je do veľkej miery závislá od toho, či je informácia alebo úloha spracovávaná paralelne viacerými perцепčnými kanálmi alebo len jedným kanálom, v ktorom môže vzniknúť preťaženie ľahšie, čo je práve prípad grafických užívateľských prostredí (príklad interferencií vizuálneho spracovania konceptov nekorešpondujúcich obrazových schém a inštruktážneho textu). Baddeley navyše hovorí o špecifickom sekvenčnom spracovaní sémanticky náročnejších, explicitných informácií prostredníctvom centrálnej exekutívy, ktorá sa dá interpretovať aj ako vedome sústredená pracovná pozornosť. Implicitné informácie, ktoré vyžadujú len reaktívnu formu spracovania mimo konceptuálnej vrstvy, sú však podľa Swellerovej teórie efektívne spracovateľné v pracovnej pamäti paralelne.

3.2 Teória dorzálnej a ventrálnej dráhy

Rozdelenie implicitnej a explicitnej zložky vizuálnej informácie má relevantnú analógiu nielen v rámci spomínaných modelov behaviorálnej psychológie, ale aj z pohľadu funkčnej neurofyziológie.

Prvou kortikálnou vrstvou, ktorá spracúva vizuálny signál z LGN (*lateral geniculate nucleus*) je primárna vizuálna kora (tiež označovaná ako V1) v záhlavnom laloku. V1 spracúva nízkoúrovňové informácie ako lokálne orientácie, farbu a kontrast vizuálneho stimulu v rámci malých receptívnych polí. Vyššie kortikálne úrovne vizuálnej kôry extrahujú komplexnejšie črty a pokrývajú širšie receptívne polia. Spracovanie signálu sa

ďalej delí na dve anatomicky rozdielne neurálne dráhy, na dorzálnu a ventrálnu, ktoré sú medzi sebou silne prepojené.

Ventrálna dráha smeruje k spánkovým lalokom, ktorým sú všeobecne prisudzované funkcie ako identifikácia objektov (infotemporálny lalok), rečové schopnosti (Brocova a Wernického oblasť), ako aj ďalšie sémanticky orientované kognitívne funkcie. **Dorzálna dráha** smeruje do parietálnej kôry, ktorej je prisudzovaná funkcia spracovávania priestorových vizuálnych informácií (*posterior parietal cortex*), ako aj spoluúčasť na senzomotorických koordináciách a správe motorických akcií všeobecne. Štúdie pacientov s poruchami dorzálnej dráhy preukázali početné diagnózy súvisiace s priestorovou koordináciou ako simultánna agnózia (porucha simultánneho vnímania len jedného objektu bez schopnosti uvedomovať si jeho kontext v rámci väčšieho celku), ataxia (porucha pohybovej senzomotorickej koordinácie), apraxia (porucha exekutívnej časti plánovanej motorickej reakcie), alebo akinetopsia (neschopnosť vnímať pohyb).

Z pohľadu neuropsychológie a klinických štúdií sa rozdeleniu dorzálnej a ventrálnej dráhy tiež prisudzuje pomenovanie “*where & what*” (kde a čo) [26]. Funkčné rozdelenie týchto dráh (priestorovo-implicitného (*where*) a deklaratívne-explicitného (*what*)) je teda v silnej korelácii s už spomínaným rozdelením zložiek vizuálno-priestorového náčrtníka, ako aj delenia modalitných pamäťových kódov podľa Paivia, Baddeleyho a Engelkampa.

3.3 Formulácia hypotézy

Ústrednou problematikou kognitívnej ergonómie a štúdií použiteľnosti užívateľských prostredí je nepochybne miera nežiaducej kognitívnej záťaže, ktorú produkuje neprirodená interakcia. Nasledovnou hypotézou sa pokúsím o explicitné sformulovanie hypotézy kľúčových princípov, ktoré by mohli slúžiť ako jadro efektívnej koncepcie GUI HCI, založené na neuropsychologických poznatkoch minimalizovania kognitívnej záťaže.

Vzhľadom na:

- 1) preukázanú efektivitu paralelného spracovania informácie viacerými perцепčno-informačnými kanálmi z hľadiska vyťaženia pracovnej pamäti [83][25],
- 2) neurofyziologické poznatky o disjunktnom, nezávislom spracovávaní priestorovej (implicitnej) a „sémantickej“ (explicitnej) zložky vizuálnej informácie cez samostatné informačné kanály dorzálnnej a ventrálnej neurálnej dráhy [26][38],
- 3) náchylnosť preťaženia pracovnej pamäti pri spracovaní náročnejšej informácie či úlohy len jedným informačným kanálom (v prípade GUI, sémantickou interferenciou explicitných, ventrálnou dráhou spracovaných vizuálnych črt),

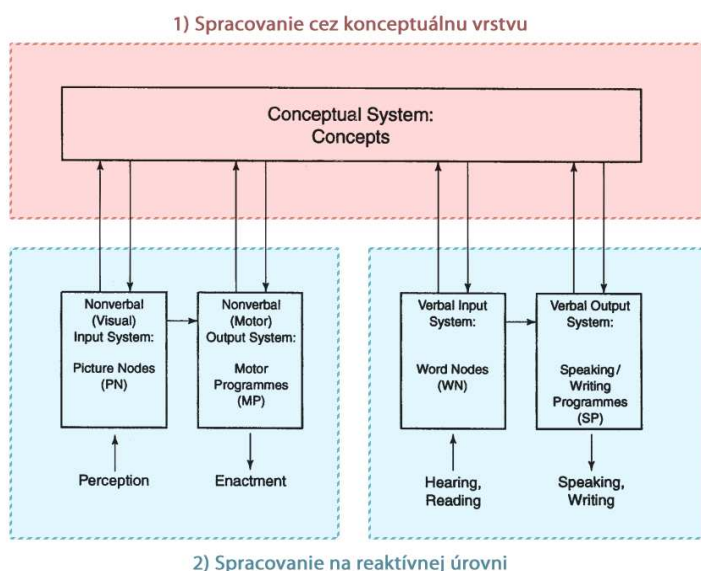
predpokladám nasledovné implikácie:

A) Vzťah vstupno-výstupných interakčných kanálov

Výsledky Engelkampových štúdií umožňujú identifikovať dve typologicky odlišné úrovne spracovania signálov a interpretovať ich špecifické využitie, a to pre fázu učenia, resp. zoznamovania sa s prostredím (snaha o zámerné presadzovanie perceptuálnej interakcie so zaťažením konceptuálnej zložky) a fázu bežnej operatívnej práce (snaha o elimináciu mentálneho prechodu konceptuálnej vrstvy a primárne využívanie reaktívnej úrovne interakcie):

- 1) **Spracovanie cez konceptuálnu vrstvu** - podľa Engelkampovho modelu nutne prebieha pri interakčných kombináciách (“slučkách”) vstupu vizuálnej percepcie a rečového výstupu, alebo pri sluchom prijímanej vstupnej informácii a výstupom formou motorickej akcie. Tieto dve kombinácie by mali podľa modelu posilňovať vedome uchopenie informácie a napomáhať budovaniu mentálnych schém. Je to teda ideálny spôsob perцепčnej distribúcie informácií pri zámere naučiť užívateľa

pracovať s novým prostredím. Pre tento druh spracovania je však typické vyššie kognitívne zaťaženie ako aj sekvenčnejšie spracovanie vnemov (viď. obr. č. 4).



Obr. č. 4: Delenie interakčných kanálov Engelkampovho modelu.

- 2) **Spracovanie na reaktívnej úrovni** - môže prebiehať podľa modelu len pri kombinácii interakcie vstupu vizuálnej percepcie a výstupu motorickej akcie alebo pri verbálnom vstupe sluchovou percepciou a verbálnym rečovým výstupom. Tieto dve možnosti modelovo autonómnych interakcií je navyše možné súčasne paralelizovať, keďže sa jedná o možnosť implicitného (automatizované) spracovávania, ktoré nevyžaduje prechod cez konceptuálnu vrstvu a teda ani nevyhnutne vedomé sústredenie pozornosti. Vzhľadom na potenciálnu reaktívnu úroveň týchto dvoch interakčných “slučiek” a ich možnosti paralelizácie, je tento spôsob interakcie ideálny pre bežnú (operatívnu) prácu v GUI, teda pre často opakované, štandardné, primitívne úlohy HCI (ako napr. navigačné akcie pre pohyb po GUI).

Keďže sa v mojej práci orientujem na vizuálno-priestorovú stránku GUI interakcie a súčasne technologické možnosti zatiaľ nie sú schopné plnohodnotnej a spoľahlivej, plne verbálnej časti HCI interakcie, budem sa ďalej v rámci tohoto konceptu sústreďovať len na vyhliadky vizuálno-motorickej reaktívnej slučky, pričom verbálnu interakčnú slučku reaktívnej úrovne spomeniem len marginálne.

B) Implicitný vs. explicitný charakter spracovania informácie

Nepočítajúc výukovú fázu (fázu zoznamovania sa s prostredím), dalo by sa zovšeobecniť, že cieľom vývoja GUI je jeho zjednodušovanie sústreďovaním bežných, často opakovaných užívateľských úkonov na čo najnižšiu operatívnu úroveň, eliminovaním explicitne spracovávaných úkonov (vyžadujúcich vedomú sústredenú pozornosť) a nahradzovaním riešení implicitnými (na reaktívnej, automatizovanej úrovni). Engelkampovým modelom som už naznačil teoretickú formu hypotézy, sústredením sa na nízkoúrovňovú interakciu vizuálnych vstupov a motorických výstupov. Je však nevyhnutné ďalej zdôrazniť dôležitosť duálnej neurálnej dráhy vstupného vizuálneho signálu. Vzhľadom na to, že prevažná časť explicitných (symbolických) vizuálnych informácií, ktoré sú silným potenciálnym zdrojom kognitívnej záťaže, je spracovávaná temporálnymi lalokmi (ventrálnou dráhou), považujem za perspektívne sústrediť sa na posilnenie implicitnej perцепčnej a motorickej stránky HCI a na taký druh interakcie, ktorého prislúchajúce úlohy by boli korelovateľné s dorzálnou dráhou. Pri štandardných GUI rozhraniach využívajúcich relatívne malú, staticky lokalizovanú 2D zobrazovaciu plochu, je dorzálna dráha z hľadiska potenciálu svojich kognitívnych kapacít vnímania priestoru a motorickej interakcie takmer nevyužitá. Nakoľko je jasne preukázaná efektivita znižovania kognitívnej záťaže pri paralelizácii viacerými nezávislými informačnými kanálmi, v prípade vizuálnej perцепcie sa zdá byť pre profit HCI vhodnou príležitosťou práve intenzívnejšie využitie funkcií prisudzovaných dorzálnou dráhou a to špeciálne posilnením využitia priestorovej zložky vizuálneho a motorického kanálu,

posilnením ich priestorovej konceptualizácie a ich vzájomným integrovaním (napr. formou gesturálnej interakcie a jej hlbším implementovaním aj do vizuálnej časti GUI).

V nasledujúcich kapitolách sa pokúsim bližšie poukázať na možnosti praktického využitia načrtnutých priestorovo orientovaných implicitných schém, ich vplyv na kognitívnu záťaž a vizuálnu pozornosť.

4. Implicitné vizuálne prostriedky v GUI

Špecifickou charakteristikou súčasných grafických užívateľských prostredí je silná dominancia vizuálneho kanálu, čo má za následok ľahšie preťaženie pracovnej pamäti. Ako som už naznačil, zvukový, resp. verbálny kanál by mohol byť ideálnym prostriedkom pre paralelizáciu a rozklad kognitívnej záťaže v zmysle Engelkampovho modelu, avšak pre bežné GUI prostredia osobných počítačov nie je perspektívny aspoň z dvoch dôvodov. Prvým dôvodom je, že súčasný stav technológií nie je schopný zabezpečiť obojstrannú úspokojivú úroveň, poskytujúcu plnohodnotnú verbálnu HCI interakciu. Druhým dôvodom je, že pri bežnom používaní osobných počítačov, bez izolovaných posluchového zariadenia ako slúchadlá, môžu zvukové emisie šíriace sa v prostredí pôsobiť veľmi rušivo – obzvlášť ak vezmeme do úvahy fakt, že v reálnych pracovných prostrediach sa často v tesnej blízkosti vyskytujú ďalší užívatelia s ďalšími osobnými počítačmi.

Okrem zjavného perceptuálneho obmedzenia GUI prostredí na vizuálny kanál je ďalej nutné poukázať aj na ďalšie skryté limitácie v rámci tohto kanálu. V bežných HCI interakciách osobných počítačov je vizuálna percepcia ochudobnená aj o binokulárnu percepciu hĺbky, prirodzeného mentálneho mapovania priestoru či globálnejších vizuálnych scén. Hĺbková a priestorová informácia má pre rýchle zorientovanie sa človeka kritický význam. Dizajnéri grafických užívateľských prostredí preto bežne

využívajú všetky dostupné možnosti piktorálnej kompenzácie, akými sú pri dvojrozmernom zobrazovaní monokulárne indikátory hĺbky alebo ďalšie implicitné vizuálne vodítka.

4.1 Indikátory vnímania hĺbky

Vnímanie hĺbky je prirodzenou súčasťou binokulárnej vizuálnej percepcie každého zdravého človeka. Hĺbku je možné vnímať prostredníctvom tzv. indikátorov (alebo tiež kľúčov), ktoré sa delia podľa spôsobu vnímania na binokulárne (stereoskopické) a monokulárne. Binokulárne indikátory sú charakteristické spracovávaním signálu z oboch očí a vyhodnocovaním ich rozdielnosti. Medzi binokulárne indikátory patria:

- **Binokulárna retinálna disparita (stereopsia)** – čím bližšie sa nachádza objekt, na ktorý človek smeruje upriamenú pozornosť, tým odlišnejší obraz dopadá na sieťnicu ľavého a pravého oka. Mozog je schopný pomocou týchto rozdielností vizuálnych obrazov vyhodnocovať fyzickú vzdialenosť.
- **Konvergencia svalstva** – mozog vyhodnocuje napätie okulomotorického svalstva potrebného ku konvergencii osí oboch zrakových orgánov. Podobne, ako pri binokulárnej disparite platí, že čím je pozorovaný bod alebo objekt v priestore bližšie k pozorovateľovi, tým menší uhol tieto osi zvierajú, a tým silnejší signál z vypätia svalstva mozog prijíma.
- **Horopter** – je množina korešpondujúcich bodov obidvoch sieťnic vytvárajúca kružnicu, ktorá popisuje rovinu jednej úrovne konvergence nedisparátnych obrazov. Charakteristickou vlastnosťou je, že pri binokulárnom vnímaní sa projekcie objektov, ktoré sú od pozorovateľa vzdialenejšie ako úroveň horoptera, neprekrížené rozdvajujú a projekcie objektov, ktoré sú od horoptera bližšie sa cez seba prekrývajú prekrížené.

Nakoľko sa pri bežne rozšírených GUI HCI používa len dvojrozmerné zobrazovacie médium, okrem potenciálne užitočných poznatkov z hľadiska neurofyziológie, binokulárne indikátory neposkytujú v súčasnosti žiadne priame praktické možnosti využitia pri tvorbe GUI.

Oveľa efektívnejším nástrojom sú však **monokulárne indikátory**, nakoľko sú vnímateľné aj pri pozorovaní len jedným okom a aj pri vnímaní dvojrozmernej reprodukcie zo zobrazovacej plochy GUI. Na rozdiel od binokulárnych indikátorov, je silný predpoklad, že mnohé monokulárne kľúče si človek adoptuje až postupne počas svojho mentálneho vývoja [12]. Existuje niekoľko druhov indikátorov s rozdielnym spôsobom ich mentálneho spracovania:

- **Lineárna perspektíva** – Ide o prirodzenú optickú vlastnosť ľudského oka, kde sa na sietnici vytvára dvojrozmerná projekcia obrazu okolitého sveta radiálnou konvergenciou paralelných línií, pričom najvzdialenejšie objekty sa limitne stretávajú v jednom konvergujúcom bode.
- **Relatívna veľkosť** – Ak sú v lineárnej perspektíve umiestnené dva rovnaké objekty v dvoch rozdielnych vzdialenostiach od pozorovateľa, je možné abstrahovať informáciu o hĺbke na základe ich rozdielného pomeru veľkosti priemetu na sietnici.
- **Interpozícia** – Objekty, ktoré sú umiestnené v priestore za sebou sa zvyknú bežne prekryvať. Pozorovateľ tak môže na základe takéhoto prekryvania vnímať nielen ich poradie v priestore, ale aj ich relatívnu vzdialenosť.
- **Osvetlenie a tieň** – Svetlo šíriace sa v priestore je nosným médiom pre vizuálnu percepciu. V prirodzených scénach sa bežne vyskytuje nejaký zdroj osvetlenia a objekty, ktoré môžu svetlo odrážať alebo pohlcovať. Okrem relatívnej vzdialenosti

objektov od zdroja svetla je možné vďaka vrhanému tieňu pozorovať napríklad aj ich naklonenie, veľkosť alebo vzdialenosť od podkladovej roviny.

- **Gradients textúr** – Každý, či už viac alebo menej transparentný objekt je za prirodzených podmienok osvetlenia viditeľný vďaka svojej textúre či štruktúre. V prípade, že pozorovaný objekt je v priestore zakrivený, členitý alebo má niektorú zo svojich plôch naklonenú voči pozorovateľovi neortogónalne, je pozorovateľná zvýšená hustota textúry a jej gradientu. Hustota sa úmerne zvyšuje so vzdialenosťou pozorovaného bodu na objekte v priestore.
- **Vzdušná perspektíva** – Mimo vákua, sa každé prostredie vyznačuje určitou hustotou. Čím ďalej je v prostredí umiestnený nejaký pozorovaný objekt, tým viac častíc rozptýlených v prostredí bráni jeho viditeľnosti, čo má tiež za následok úmerné znižovanie parametrov ako kontrast, intenzita alebo charakter farby.
- **Pohybová paralaxa** – Pri pohybe pozorovateľa v priestore, bližšie staticky umiestnené objekty menia svoju relatívnu vizuálnu polohu rýchlejšie ako objekty vzdialenejšie. Pri takomto pohybe priestorom je človek schopný mapovať súčasne celé priestorové usporiadanie objektov v scéne, keďže rýchlosť ich zdanlivého pohybu určuje ich aktuálnu vzdialenosť.
- **Akomodácia oka** – Pre zaostrený priemet obrazu na sietnici musí ľudské oko prispôbovať polomer zakrivenia šošovky. Deje sa tak pomocou ciliárneho svalu, ktorého intenzita stiahnutia dáva mozgu parciálnu informáciu o blízkosti pozorovaného objektu.
- **Hĺbka ostrosti** – Keďže si ľudské oko dynamicky prispôbuje zaostrenie podľa aktuálnej vzdialenosti objektu, nedokáže vnímať zaostrene celý priestor. Objekty

mimo rozsah zaostrenej hĺbky, sú preto vnímané ako rozmazané, čo napomáha schopnosti odlišovať objekt od pozadia.

Až na výnimku akomodácie a hĺbky ostrosti, je väčšina monokulárnych indikátorov dôsledkom princípov lineárnej perspektívy. Pri vizuálnej percepcii sa pochopiteľne súčasne využíva kombinácia viacerých indikátorov. Existujú preto rôzne dohady a predpoklady o dominancii a citlivosti na jednotlivé indikátory, ich prioritizáciu, či vplyv na vizuálnu pozornosť pri špecifických scénach.

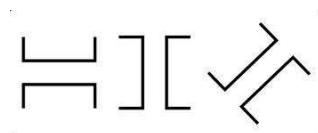
4.2 Vizuálne vodítka a Gestalt princípy

Okrem indikátorov hĺbky existujú aj elementárnejšie vizuálne prostriedky, ktoré majú vplyv na vizuálnu pozornosť, spôsob prehliadania scény, či samotný spôsob spracovávania zobrazovanej informácie, ktorými sú vizuálne vodítka. V súčasnosti je možné pozorovať využitie týchto vodítok asi najlepšie v prostredí komerčných webov. Keďže návštevník webovej stránky často len povrchno prehliada obsah webu, dizajnéri sú nútení používať agresívnejšie metódy vnucovania konkrétnych informácií, argumentov, či kľúčových slov. Medzi najčastejšie používané prvky patria okrem sémanticky orientovaných piktogramov a grafikonov, rôzne oddeľovacie línie, rámce, ohraničenia a kontrastné plochy. Väčšinu týchto prostriedkov však spájajú princípy popísané Gestalt teóriou, ktorej základy siahajú až do roku 1924 [46].

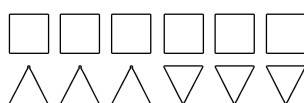
Gestalt princípy sú súborom deskriptívnych pozorovaní psychologických aspektov vizuálnej percepcie, ktoré umožňujú efektívne interpretovať systematickosť organizovania vizuálne samostatných prvkov do zmysluplných polí a štruktúr vytvárajúcich nový celok [107]. Je tiež všeobecne známe, že zlepšujú perceptuálnu interakciu ako aj schopnosť učenia a efektivitu grafických inštruktážnych prostredí [90]. Charakteristickými znakmi Gestalt princípov sú: **emergencia globálnej informácie**

(súvislosť hierarchicky podriadených elementov vytvárajúca novú, hierarchicky nadradenú informáciu), **zhmotnenie** (automatické dopĺňanie chýbajúcich elementov či vlastností), **multistabilnosť** (možnosť vnímania dvoch a viac emergentných obrazov) a **invariancia** (odolnosť identifikácie útvarov voči miernym deformáciim, rotáciám alebo voči použitiu iných, hierarchicky podriadených elementov). Pre účely grafického dizajnu sa najčastejšie používajú princípy zoskupovania:

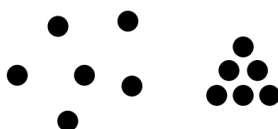
Zákon symetrie a rovnováhy – tendencia vnímania celku, v prípade, že má celá skupina vizuálnych prvkov vlastnosť stredovej alebo osovej súmernosti:



Zákon podobnosti – tendencia vytvárania skupín objektov odlišením podobných a nepodobných objektov:



Zákon blízkosti – tendencia vnímania zhlukov objektov ako skupín alebo sérií:



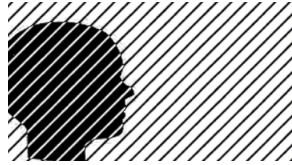
Zákon periodicity – tendencia vidieť vizuálny rytmus opakovania určitého vzoru:



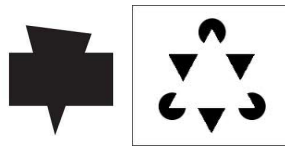
Zákon plynulosti pokračovania a smeru – tendencia hľadania súvislých tvarov s neprerušným pokračovaním:



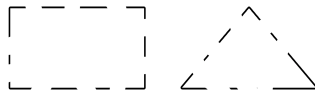
Zákon oddeľovania objektu od pozadia – schopnosť zameriavať pozornosť na zmysluplný objekt a ignorovať jeho okolie:



Zákon výstižnosti (Prägnanz) – tendencia vidieť najjednoduchší tvar:



Zákon uzavretosti – tendencia vnímať celistvý vizuálny útvar napriek jeho neúplnosti:

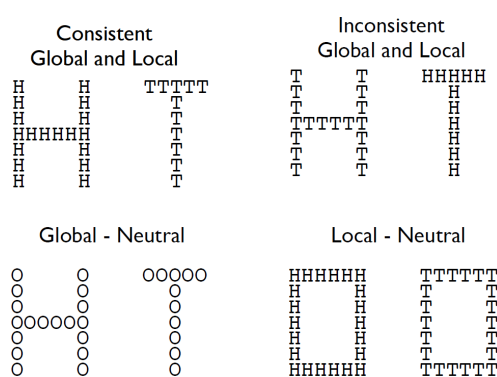


Gestalt princípy zoskupovania sú špecifické svojim nízkoúrovňovým prepojením vnímania priestorovej zložky a explicitného symbolického vnímania, teda z pohľadu neuropsychológie predstavujú príklad fundamentálneho funkčného prepojenia dorzálnnej a ventrálnej dráhy spracovania vizuálneho signálu [31].

Súvislosti potvrdzujú tiež prípadové štúdie pacientov s poruchou simultánnej agnózie [68] a bilaterálnymi léziami parieto-temporálnej oblasti mozgu [119], pri ktorých bolo preukázané, že takýto pacienti mali vážny deficit rozpoznávania globálnej, Gestalt úrovne vizuálnej scény, pričom rozpoznávanie podkladových lokálnych objektov bolo neovplyvnené. Marginálne stojí tiež za pozornosť Wiliamsov [115] Dawnov syndróm [13]. Wiliamsov syndróm je zriedkavá porucha neurálneho vývoja, pričom pacienti s touto poruchou nie sú schopní vnímať globálnu úroveň hierarchického vizuálneho stimulu, kdežto pacienti s Dawnovým syndrómom naopak nie sú schopní vnímať úroveň lokálnu.

Behaviorálne experimenty skúmajúce priestorový **vzťah percepcie globálneho a lokálneho aspektu** vizuálnej pozornosti boli prvýkrát realizované už v roku 1977

Navonom [76]. Navon použil vo svojom experimente tri variácie konzistentného, nekonzistentného a neutrálneho hierarchického podnetu (veľké písmeno skladané z malých) a testoval reakčné časy identifikovania rozdielností lokálnych a globálnych písmen za podmienok sústredenia pozornosti na globálny a lokálny aspekt. Výsledky jeho experimentu poukázali na fakt, že participanti dokázali identifikovať lepšie nekonzistentnosť na globálnej úrovni než na lokálnej, z čoho bol vyvodený predpoklad prioritizácie vnímania globálnej zložky.



Obr. č. 5: Prevedenie podkladov Navonovho experimentu [76].

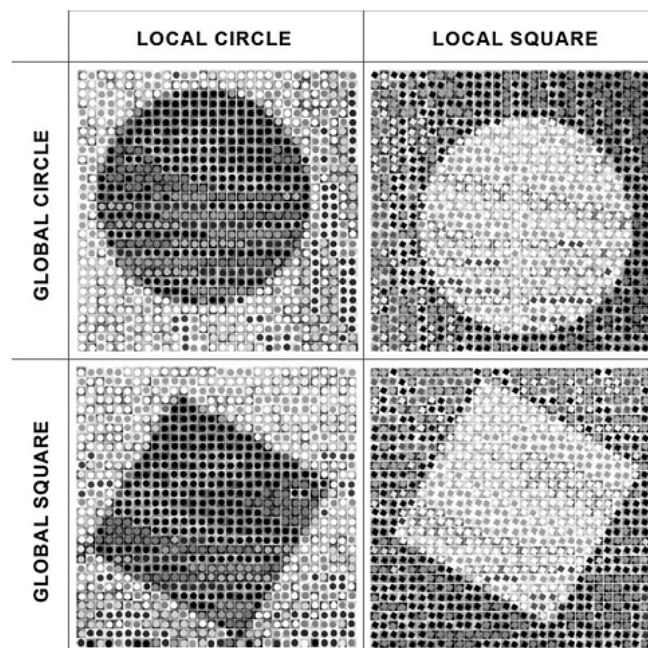
Prioritizácia vnímania globálnej zložky bola tiež skúmaná štúdiami, ktoré poukázali na silný vplyv faktorov ako je veľkosť lokálnych elementov [54] a ich priestorová frekvencia (hustota) [8][100], pričom vplyv kontrastu alebo farby nebol potvrdený.

Ako vysvetlenie tohto javu sa ponúka interpretácia neurofyzologických poznatkov o paralelnom spracovávaní vizuálneho signálu cez **“rýchlu” magnocelulárnu neurálnu dráhu** a **“pomalú” parvocelulárnu dráhu**. Toto neurálne rozdelenie začína na úrovni gangliových buniek sietnice a je im prisudzovaná tiež vlastnosť charakteristického spracovania priestorových frekvencií [70]. Pre magnocelulárnu (rýchlu) dráhu ide o spracovávanie nízkych priestorových frekvencií (globálny pohľad na scénu), [99] pričom pre parvocelulárnu dráhu ide o spracovanie detailných vizuálnych informácií, reprezentovaných vysokými priestorovými frekvenciami [18]. Rýchle magnocelulárne

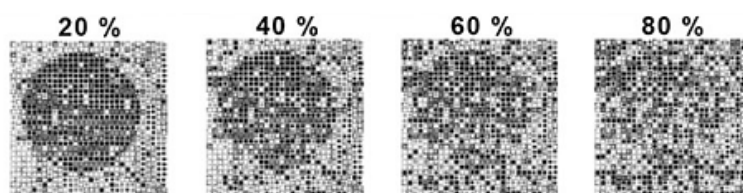
spracovanie nízkej priestorovej informácie je teda pokladané za iniciátora *top-down* procesov pozornosťných mechanizmov, predikcií a iných osvojených užívateľských zvykov v GUI [59].

Existujú ďalej štúdie popisujúce súvislosť vnímania Gestalt zákonov s neurálnou aktivitou špecifických kortikálnych oblastí ako *temporo-parietal junction* (TPJ, časť rozdeľujúca temenný lalok od spánkového), *anterior cingulate cortex* (ACC, predná časť prepojovacieho mostíku *corpus callosum*, spájajúceho mozgové hemisféry) a *precuneus* (prostredná časť *superior parietal cortex*) [45]. Oblasť TPJ navyše korešponduje s mozgovými oblasťami typicky poškodenými u pacientov so simultánnou agnóziou. Títo pacienti sa prejavujú tiež deficitom vnímania globálnej Gestaltovej zložky z vnímanej scény. Nedávna štúdia **Huberle & Karnath** [45] poukázala na prepojenie medzi behaviorálnymi prejavmi a fMRI aktiváciou v oblasti TPJ. Experiment pozostával z merania aktivácii pri percepcii globálnych stimulov s rozdielnou zreteľnosťou vnímateľnosti (kruhy a štvorce, pozostávajúce z dvojrozmerného poľa menších kruhov a štvorcov - vid'. obr. č. 6) a bol testovaný na zdravých jedincoch.

Výsledky poukázali na signifikantný vplyv vnímania Gestaltovej zložky s aktiváciou v oblasti TPJ, pričom boli pozorované rozdielne aktivácie úmerné zmenám zreteľnosti globálnej zložky. Je dôležité pripomenúť, že TPJ je oblasť mozgovej kôry, ktorá je neurálnym rozcestníkom dorzálnej a ventrálnej vizuálnej dráhy, a teda experiment opäť podporuje hypotézu disjunktného delenia ventrálnej a dorzálnej dráhy, resp. priestorovej a sémantickej časti vizuálnej informácie.



a)



b)

Obr. č. 6: Podklady experimentu Huberle & Karnath [45]:

- a) kombinácie globálnej a lokálnej vizuálnej zložky použité v experimente,
 b) prevedenie rozdielnej vnímateľnosti globálnej zložky.

Zo všeobecnejšieho pohľadu psychológie by sa dali prostriedky vnímania vizuálnej informácie rozdeliť na dve základné perspektívy. Prvou perspektívou je pohľad štrukturalizmu, ktorý vysvetľuje percepciu pomocou spomenutých Gestalt princípov a vodítok, teda pomocou viac – menej pevne daných pravidiel v súvislosti s kontextom situácie. Druhou perspektívou je Gibsonov ekologický prístup pomocou tzv. affordancií, na ktorých ďalej stavia teória obrazových schém a metafor.

4.3 Užívateľské metafory

Affordancie

Gibson použil termín affordancií prvýkrát vo svojej publikácii v roku 1977 [33], kde ho definoval ako množinu možných prirodzených akcií (pohnútok) pozorovateľa latentne vyplývajúcich z prostredia. Treba zdôrazniť, že v zmysle tohto konceptu bol neoddeliteľnou súčasťou aj samotný pozorovateľ a jeho fyzická schránka, a teda išlo o vzájomný vzťah medzi prostredím a jeho pozorovateľom. Norman [80] túto formuláciu zúžil z jej objektívneho charakteru vyplývajúceho z prostredia len na affordancie, ktoré sú vnímané pozorovateľom, nakoľko nemá zmysel reálne uvažovať o možnosti affordancie, pokiaľ je mimo dosah percepcie. Príkladmi affordancií sú napr. uchopenie a otvorenie pre kľučku na dverách, gúľanie a hádzanie pre loptu, strihanie pre nožnice, pričom svoju dôležitosť pre konkrétnu akciu tvorí hlavne design vyhotoveného objektu (napr. rôzne tvary kľučiek – guľové, pákové, atď.). Špecifickým príkladom affordancie pre GUI je napr. virtuálne grafické tlačítko, ktoré využitím indikátorov hĺbky, ako sú tieňovanie, odlesk či gradienty, svojim vizuálnym vystúpením z priestoru napovedá o možnosti stlačiteľnosti, často podporenej tzv. hover parametrom (presunutie kurzora ponad objekt). Pri HCI majú affordancie ako súbor všetkých možností vyplývajúcich z danej situácie dôležité a intenzívne využitie [81]. Umožňujú totiž objektívnejšie interpretovať interakčný charakter designu, efektívnu analýzu prostredia, možnosti užívateľa v konkrétny moment, ale tiež syntézu v podobe predpokladov jeho potenciálneho správania.

Obrazové schémy

S konceptom affordancií silne súvisí tiež koncept obrazových schém. Obrazové schémy majú svoje významové ukotvenie v oblastiach telesnej interakcie a lingvistiky. Ide o dynamický, motorický, mentálny vzor, ktorého hlavnými popisnými zložkami sú dianie v čase a reprezentácia v priestore. Mnohé obrazové schémy sú tiež príznačné tým, že sú

popísateľné jednoduchou trajektóriou. Príkladom takejto obrazovej schémy je schéma reprezentujúca výraz "von", tzn. jeho reprezentácia vyjadruje počiatočný stav nachádzajúci sa v nejakom lokálnom bode priestoru a jeho koncový stav trajektórie sa nachádza mimo pôvodne ohraničené územie pôvodnej polohy (ďalšími príkladmi môžu byť výrazy ako cesta, cyklus, naprieč, ponad, odklon, atď.). Hoci najprirodzenejšou reprezentačnou formou sa javí byť vizuálny kanál, obrazové schémy môžu byť reprezentované multimodálne, a teda aj samotným stelesneným prežívaním imaginácie motorickej akcie, bez vizuálnej zložky.

Takéto "stelesnené prežívanie" má hlboké opodstatnenie pri vytváraní obrazových schém ako aj affordancií a najlepšie ho popisuje koncept stelesnenej kognície (*embodied cognition*). Stelesnená kognícia vychádza z predpokladu, že akékoľvek kognitívne operácie sú ovplyvňované a determinované spoluúčasťou telesnej schránky, v ktorej je ukotvených množstvo mentálnych konceptov a kauzálit kategorizácii objektov.

Metafory všeobecne

Obrazové schémy a ich stelesnenie hrajú z hľadiska lingvistiky dôležitú rolu pri evolúcii komplikovanejšieho kognitívneho konceptu, ktorým je metafora. Obrazová schéma v tomto prípade môže slúžiť ako reprezentačný nástroj k uchopeniu abstrakcie nejakého stavu, deja alebo situácie. Ako príklad môže poslúžiť metafora „emočnej vertikály“ [61], kde šťastie či úspech je reprezentované ako nachádzanie sa na životnej pozícii ‚hore‘ a pocit nešťastia, či neúspechu ako pozícia ‚dole‘. Tá zrejme vyplýva zo stelesneného prežívania situácie ‚dole‘, kedy je v prípade nešťastia, či neúspechu fyzicky pozorovateľne skleslý, nevzpriamený, unavený. Lakoff a Turner vo svojej štúdii [60], v ktorej sumarizovali známe lingvistické metafory (anglického jazyka), prišli k pozoruhodnému záveru, že 2/3 všetkých zaužívaných slovných metafor sú priestorového charakteru (viď. tiež príklad vertikálnej reprezentácie emočného stavu hore vs. dole, alebo metafory ako život je cesta, atď.)

Metafora vo všeobecnosti umožňuje ľuďom porozumieť jednej veci na základe inej, bez nutnosti vnímania identity [109]. Po obrazových schémach, ktoré sú voči metaforám na relatívne nízkej úrovni abstraktizácie, a ktoré majú prevažne len deskriptívny charakter, sa metafora líši práve vo svojej možnosti využitia ako nástroja pre aktívnu abstraktizáciu uchopením situácie v rámci inej, už existujúcej mentálnej schémy.

Crossmodalita a synestézia

Množstvo metafor je založených na projekcii nielen medzi dvoma rôznymi prostrediami, ale aj medzi rôznymi modalitami. Najčastejším prípadom sú bežne používané lexikálne metafora, ktoré okrem svojej abstraktnej konceptuálnej roviny zahŕňajú aj paralelné, nízkoúrovňové spracovanie vo forme cross-modálnej percepcie. Pre príklad najnižšej abstraktizačnej úrovne môže slúžiť Köhlerov experiment "Takete & Baluba" [57] a Ramachandranov experiment 'Kiki & Buoba' [94]. V experimente mali participanti za úlohu priradiť k dvom vizuálnym objektom, jednému s ostrými hranami a druhému s oblými, podľa vlastných subjektívnych predpokladov fonologickej príslušnosti, pomenovania 'Kiki' alebo 'Buoba', pričom dané slová v jazyku experimentu nemali žiadny sémantický význam. Spomedzi všetkých participantov, 95 % až 98 % označilo ako 'Kiki' objekt s ostrými hranami a 'Buoba' s oblými. Ramachandran a Hubbard [93] predpokladajú, že tento jav môže napomáhať k objasneniu evolúcie jazyka, nakoľko pomenovanie objektov evidentne nebolo arbitrárneho charakteru. Podobné javy boli objavené aj pre mnohé ďalšie crossmodalities. Podolsky [88] vo svojom experimente s premaľovaním kancelárie, poukázal na súvislosť farby a vnímania tepla. Pri premaľovaní kancelárie na modro, mali testovaní zamestnanci pocit chladu a termostat klimatizácie nastavovali na vyššie hodnoty voči normálu, pričom pri premaľovaní na žlté, boli hodnoty termostatu klimatizácie nastavované na nižšie, keďže zamestnanci sa v takejto miestnosti cítili subjektívne teplejšie.

Okrem prejavov crossmodálneho efektu v bežne známej miere však existujú jedinci s abnormálne vyvinutými asociačnými prepojeniami, napríklad so schopnosťou vnímať

imaginárne farby pri percepcii čísiel, písmen či slov, pričom sú tieto imaginácie opakovateľne vyvolateľné, automatické, neovplyvniteľné vlastnou vôľou a ich charakter pretrváva v čase [102]. Takéto crossmodálne ovplyvňovanie percepcie sa nazýva synestézia. Okrem najčastejšie vyskytujúcej sa syntestézie projekcii grafém do farieb, boli dokumentované aj projekcie pohybu do zvuku, čísiel do vzdialenosti, slov do chuti, zvukov do farieb a ďalších vyše 60 druhov [118]. U mnemonika Solomona Shereshevského bola údajne dokumentovaná súčasná synestezia až 5 zmyslov resp. modalít [67].

Podľa Ramachandrana a Smileka [102], ktorí sa venovali experimentom na jedincoch s projekciou synestezie grafém do farieb, môže byť za jav synestezie zodpovedné tesné susedstvo kortikálnych regiónov, ktorým je prisudzované spracovávanie asociovaných druhov modalít. Argumentujú napríklad štúdiami, ktoré popisujú príbuznosť kortikálneho spracovania farby [41] a grafém [92] v oblasti *'fusiform gyrus'*, ktorá ďalej susedí s vizuálnou kôrou V4. Je teda minimálne diskutabilné či niektoré metaforické či crossmodálne spojenia nemajú pôvod v obdobných neurálnych prepojeniach.

Z globálnejšieho pohľadu ide o potenciálne podobný prípad kortikálneho regiónu angular gyrus, ktorý je neurálnou križovatkou medzi oblasťami záhlavného, spánkového a temenného laloku, teda križovatkou medzi vizuálnou kôrou a ventrálnou a dorsálnou dráhou. Tejto oblasti je prisudzovaná zodpovednosť za procesy spojené s jazykom, matematikou a niektorým ďalším, logickým, kognitívnym funkciám. Ramachandran vo svojom výskume s pacientami s poškodenou kôrou v tejto oblasti poukázal na to, že takýto jedinci neboli schopní porozumieť metaforickému významu slovných fráz, pričom rozumeli len doslovnému obsahu (čo je analogickou situáciou porúch vnímania Gestaltovej vizuálnej zložky). V inom experimente poukázal tiež na ich neschopnosť vnímať crossmodálny efekt *'Kiki & Buoba'*, bežne fungujúci u zdravých jedincov. V súvislosti s výsledkami svojich experimentov a pozorovania úmerne zvyšujúcej sa vyvinutosti tejto mozgovej oblasti evolučne smerom od jednoduchých cicavcov, cez

opice, primátov až k človeku teda predpokladá, že táto oblasť je minimálne čiastočne zodpovedná za porozumenie metaforám [92].

Užívateľské metafor v HCI

Metafory si získali v oblasti HCI význačné postavenie a stali sa súčasťou terminológie teórie dizajnu ako aj praxe. Cieľom metafor využívaných v HCI je však skôr ako schopnosť naučiť sa s prostredím pracovať (learnability) práve v úvode spomenutá použiteľnosť, nakoľko ich prirodzene pochopiteľná pointa a predpokladané funkcionálne komponenty sú základným krokom k jednoduchosti a intuitívnosti, ktorá činí systém použitelným [125]. Dôvodom k používaniu metafor v HCI a v grafických užívateľských prostrediach je niekoľko. Najsilnejším je zrejme dôvod, že informačným technológiám a HCI chýbajú akékoľvek prirodzené affordancie [34], z čoho ďalej vyplýva, že užívateľské prostredie neobsahuje žiadne intuitívne interakčné vodítka alebo implicitnú informáciu o tom, ako ho použiť. Ďalším dôvodom k použitiu metafor je, že užívateľovi svojou podstatou poskytujú chýbajúce implicitné affordancie a umožňujú definovať konkrétny spôsob interakcie (pochopiteľne odhliadnuc od možných problémov s neintuitívnosťou periférneho, interakčného hardwareu).

4.4. Priestorová konceptualizácia v GUI

Nakoľko súčasné technologické pozadie GUI poskytuje pri HCI oveľa väčšie možnosti ako v dobe terminálov a textových rozhraní, svoje využitie si čoraz väčšmi nachádza aj estetické využitie počítačovej grafiky, ktorej zámerom je vytvárať pohodlnejšie a užívateľský prístupnejšie prostredie. V rámci evolúcie GUI je jednoznačne pozorovateľný progres hlavne v oblasti efektívneho využívania priestoru zobrazovacej plochy. Tento proces sa v oblasti teórie dizajnu nazýva priestorová konceptualizácia (alebo tiež priestorizácia, z angl. spatialization). Najväčším prínosom priestorovej konceptualizácie je v predchádzajúcich kapitolách bližšie popísaný efekt znižovania

kognitívnej záťaže. Kognitívna záťaž sa takto môže znižovať vytvorením rôznych priestorových konceptov ponad explicitnými, zväčša textovými dátami, čím sa vytvára implicitne prepojenie ich vzťahov. Pozitívny efekt podporuje aj teória multimodálneho kódovania v súvislosti s teóriou dorzálnnej a ventrálnej dráhy vizuálneho signálu, keďže vnímanie priestoru je po neurálnej stránke spracovávané ako samostatná modalita. Pre vytvorenie takýchto implicitných priestorových vzťahov sa dajú použiť spravidla dve metódy:

- Prvou metódou sú spomínané **Gestalt zákonitosti zoskupovania** (automatické mentálne zlučovanie priestorovo usporiadaných elementov do pomyselných kategórií vytvárajúcich organizovanú, typologicky samostatnú množinu).
- Druhou metódou sú **priestorové užívateľské metafory** (využívajúce abstraktizované koncepty osvojené užívateľmi mimo prostredia HCI) akým je napr. metafora pracovného stolu (desktopu) alebo metafora okien resp. tabúl prekrývajúcich sa ponad seba v priestore.

Pre zjednodušené operacionalizovanie a sparalelizovanie týchto dvoch metód, si je možné predstaviť Gestalt zákonitosti zoskupovania ako súbor istých affordancií vizuálnej percepcie.

Pre priestorovú percepciu využiteľnú v GUI existujú dva **referenčné rámce**. Prvým je **egocentrický**, založený na aktuálnej pozícii a priestorovej orientácii pozorovateľovej telesnej schránky (koncepty ako hore, dole, vpredu, vzadu, vľavo, vpravo). Egocentrický rámec je možné charakterizovať pohľadom „z prvej osoby“. Druhý referenčný rámec je **allocentrický**, kde je pozorovateľ, resp. objekt pozorovania lokalizovaný vzhľadom na globálny, staticky koordinačný systém (napríklad sever, juh, východ, západ) a takýto rámec je zväčša charakterizovaný pohľadom „z tretej osoby“ (typický napr. pre pohyb

kurzora po ohraničenej zobrazovacej ploche monitora). Pri grafických užívateľských prostrediach sa používajú oba rámce kombinovane vzhľadom na konkrétnu užívateľskú metaforu a vzhľadom na to, či je v zmysle kontextu manipulovaná poloha samotného užívateľa v prostredí (egocentrickým fiktívnym pohybom, teda zobrazením určitej špecifickej užívateľskej scény, prechod medzi stránkami alebo jej časťami) alebo je manipulovaný konkrétny grafický objekt, v aktuálnom, dočasne ustálenom prostredí (allocentrická manipulácia s dátovými súbormi, formulármi, atď.).

Anomáliou, ktorá je špecifická pre oba referenčné rámce je tzv. vertikálna asymetria [117]. V prvom rade ide o odlišnosť vnímania horizontálnych a vertikálnych vizuálnych zoskupení objektov, kde špecifickosť vnímania vertikálnych zoskupení má svoj základ ukotvený v percepcii [30]. Toto ukotvenie možno pozorovať napr. aj pri optickej ilúzii Shepardových stolov [98]. Špecifickosť vertikálnej osi tiež súčasne koreluje s psychologickým vnímaním viacerých fenoménov reálneho sveta ako gravitácia, vertikálne rozdelenie horizontu (obloha a zem), ale aj s ľudskou anatómiou, teda napr. obmedzenou možnosťou otáčania hlavy a očí v smere tejto osi (čo je zrejme tiež dôvodom pomeru strán zobrazovacích zariadení bežných osobných počítačov a ich koncepcie „na šírku“). Dôsledkom pre tvorbu dizajnu v GUI je napríklad prehľadnejšie delenie početných informácií do stĺpcov, či vertikálnych zoznamov. S vertikálnou asymetriou tiež korešpondujú jazykové metafory typu premostenia výrazov „menej“, „horšie“, či „slabšie“ so smerovaním nadol a ich pozitívne protiklady so smerovaním nahor (napr. „zvyšovanie“ úrovne, a pod.) [17].

V súčasnosti sa v prostredí GUI používa veľké množstvo metaforicky orientovaných, priestorových konceptualizácií. Uvediem teda niekoľko fundamentálnych príkladov, ktoré sa vyskytujú vo väčšine prostredí WIMP, post-WIMP, webových stránkach a ďalších GUI:

- **Absolútne hierarchické usporiadanie vrstiev objektov, interpozícia** – prejavuje sa globálnym pohľadom užívateľa „z vrchu“, pričom v prostrediach WIMP sa kurzor nachádza na najvrchnejšej vrstve ako manipulátor. Vo WIMP si navyše na rozdiel od reálneho sveta, udržujú objekty svoju stálu hierarchickú pozíciu aj v prípade, že dve okná položíme vedľa seba bez prekryvania.
- **Koncept obsiahnuteľnosti** – ide v podstate o priestorovú schému, kde špecifický typ objektov (napr. priečinkov) má vlastnosť, že môže obsahovať iné, hierarchicky nižšie objekty (napr. konkrétne dátové súbory).
- **Koncept samostatných priestorových scén, stránok** – komplexnejšie vizuálne prostredia v GUI sú zoskupované do samostatných scén, ktoré môžu využívať odlišné metafory užívateľského rozhrania. Scény orientované na textový obsah sú napr. často metaforizované k podobe fyzických (papierových) dokumentov [79].
- **Koncept štítkov / kariet (*tabs*)** – ide o efektívny grafický nástroj zvyrazňujúci aktuálnu pozíciu v rámci horizontálne alebo vertikálne usporiadaného zoznamu viacerých kariet (stránok), pričom aktuálny obsah hlavnej časti zobrazovacieho pola zodpovedá kategórii popísanej výrazným nadpisom danej karty. Základom tejto metafory v reálnom svete je spôsob prezerania kartografických dokumentov prehľadne označených vyvýšenými popisnými štítkami.
- **Koncept priamej manipulácie** – ide o koncept dynamickej, časovo spojitaj možnosti manipulácie s grafickými a dátovými objektmi. Zjavnými príkladmi tohto konceptu sú funkcie ako „*drag & drop*“, dynamická zmena polohy a veľkosti okien.
- **Koncept hyper-reálnosti (*super-realita*)** – je prípad metaforizácie, kde rozšírením konceptu pochádzajúceho z reálneho sveta, projekcia metafory nielenže funguje

rovnako intuitívne, spôsobom, akým užívateľ očakáva, no jej použitie sa javí ešte efektívnejšie ako v realite (napríklad multitouch gestúry v post-WIMP prostrediach pre približovanie alebo otáčanie fotografií).

- **Umelé asociované verbálne koncepty** – ide o umelé pomenovania užívateľských funkcionalít či prvkov, ktoré nemajú žiadnu prirodzenú spojitosť so svojim označením predurčujúcim ich spôsob metaforického používania. Príkladom je metafora záložiek (bookmarks), či pomenovanie znakového reťazca určujúceho fyzickú lokáciu súboru v dátovej štruktúre „cestou“ (path), či v prípade web portálov URL lokácia pomenovaním „adresa“.

Pri verbálne metaforizovaných, umelo asociovaných konceptoch je nutné upozorniť na možnosť konfliktnej kombinácie viacerých metafor, prichádzajúcich z odlišných oblastí informačných technológií. V prípade dátových súborov by napríklad použitie metafory „cesty“ prirodzene nekorešpondovalo s metaforou „záložky“ zaužívanou pre „adresu“ webstránky („záložka v adresári“ vs. „záložka cesty“). Zavádzanie nových (nielen verbálnych) metafor do praxe, je preto pri niektorých evolučne ustálených, zaužívaných prostrediach značne obmedzené a problematické. Konfliktné interferencie metafor vyžadujú zvýšenú abstraktizáciu pre prepojenie oboch metafor, čím uniká ich primárna pointa intuitívneho použitia a zvyšujú kognitívnu záťaž užívateľa.

5. Motorická interakcia zo strany užívateľa

Spomínané metaforu používané v GUI boli predovšetkým perцепčného charakteru, teda smeru komunikácie informácie od počítača k užívateľovi. V rámci HCI je však nemenej dôležitý pohľad aj zo strany užívateľa k počítaču, ktorý stavia na benefitoch perцепčného rozhrania vo forme affordancií, teda istej množiny ponúkajúcich sa možností k akciám, alebo metafor, ktoré poskytujú užívateľovi intuitívny konceptuálny priestor celkového prostredia.

Potenciálne akcie užívateľa v GUI by sa dali hierarchicky zoradiť podľa šírky záberu svojho konceptuálneho priestoru do nasledovných bodov:

- 1. Navigácia medzi užívateľskými scénami** – ide o prechádzanie medzi rôznymi druhmi komplexnejších prostredí ako sekcia nastavení, domovská stránka, pracovná plocha, alebo inými osobitými prostrediami. Pohyb medzi scénami je z užívateľského hľadiska najčastejšie realizovaný egocentrickým rámcom.
- 2. Voľba prvkov záujmu** – ide o akcie súvisiace s bližším sústredením sa na obsahovú stránku v lokálnejšom prostredí scény (napr. allocentricky orientovaný pohyb kurzorom v prostredí v kombinácii s konceptom „kliknutia“ ako finálnym aktom rozhodnutia).
- 3. Manipulácia s konkrétnymi objektmi** – selekcia a manipulácia s fyzickými dátami, operácie s ich obsahom, nastavovanie parametrov, zadávanie numerických a textových hodnôt premenných.

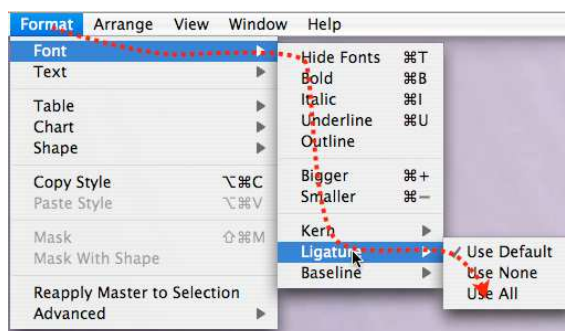
V súčasnosti najbežnejšie používanými prostriedkami motorickej interakcie v GUI zo strany užívateľa sú klávesnica pre zadávanie textových vstupov a polohovacie zariadenie (myš, trackball, touchpad, joystick, tablet alebo dotyková obrazovka) pre priestorovo orientované lokalizácie záujmov a manipulačné činnosti užívateľa. Pre efektívnu interakciu musia oba druhy týchto zariadení spĺňať isté ergonomické požiadavky.

5.1. Ergonómia v HCI

Ergonómia je dôležitý interdisciplinárny odbor zaoberajúci sa optimalizáciou efektivity pracovných podmienok pre konkrétne ľudské činnosti z hľadiska motoriky a kognitívnej psychológie. Ergonomické princípy motorickej interakcie v HCI sú často využívané aj pri

navrhovaní percpečnej stránky, napríklad pri modelovaní rozmiestnenia rôznych funkčných elementov grafického prostredia ako aj analýzu ich použiteľnosti.

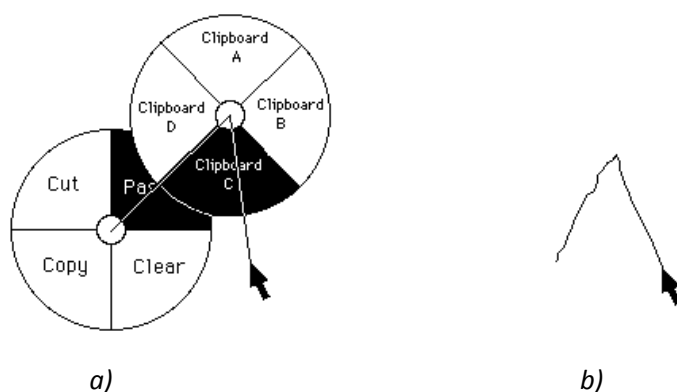
Fittsov zákon [29] je matematický model, ktorý umožňuje pre potreby HCI odhadnúť čas potrebný k rýchlemu premiestneniu lokácie kurzora od aktuálnej k cieľovej lokácii. Fitts tento pohyb popisuje závislosťou od vzdialenosti a veľkosti cieľovej plochy. Rozvinutou verziou tohto zákona je tzv. **riadiaci zákon (Steering law)** [95] popisujúci čas potrebný na presun medzi dvoma lokáciami pri definovaní cesty dvojrozmerným ohraničením (tzv. tunelom).



Obr. č. 7: Pri priechode medzi dvoma výpadkovými menu je ohraničenie popisované riadiacim zákonom užšie a vyžaduje vyššiu senzomotorickú koncentráciu.

Accot & Zhai pomocou týchto dvoch zákonov poukázali na fakt, že prejdienie kurzorom cez určitú lokáciu vyžaduje radikálne menší čas a menšiu potrebu senzomotorickej pozornosti ako presné cielenie kurzora na túto lokáciu [1]. V zmysle tejto myšlienky vytvorili koncept interakcie „**crossing-based interface**“, kde užívateľ využíva namiesto presného klikania iba prechod kurzorom cez túto oblasť. Potenciál pre zvýšenie efektivity práce môže byť využitý hlavne v prostrediach hypertextových dokumentov alebo pri zariadeniach s menšou zobrazovacou plochou, kde presné kliknutie na odkaz spôsobuje užívateľovi zbytočnú kognitívnu záťaž v podobe senzomotorickej koncentrácie. Podobným príkladom aplikácie Fittsovhovho riadiaceho zákona je ergonomická alternatíva klasických vypadkových menu použitím tzv. **koláčových menu**

(Pie-menu) s radiálnym usporiadaním položiek. Tie svojim usporiadaním podobne znižujú celkový čas a koncentráciu užívateľa potrebnú k dosiahnutiu cieľovej položky. Ďalšou výhodou niektorých implementácií Pie-menu je možnosť urýchleného výberu hierarchicky nižšej položky použitím gestúry symbolizujúcej trajektóriu pohybu prechodu jej položkami (viď obr. č. 8).



Obr. č. 8: a) Ukážka "koláčového" menu, b) znázornenie trajektórie urýchlenej voľby gestúrou

[58]

Oba koncepty, crossing-based interface a pie-menu, sú zreteľnými ukážkami využitia potenciálu gestúr a priestorovej konceptualizácie nielen percepčnej, ale aj motorickej interakcie v GUI.

Z pohľadu ergonómie sa dá vzniesť určitá fundamentálna kritika aj ku motorickej stránke interakcie s klávesnicou, konkrétne k neefektívnej koncepcii priestorového rozloženia kláves. Súčasný štandard rozloženia „**QWERTY**“ bol totiž riešením vyvinutým k použitiu pre mechanické písacie stroje. Princíp rozloženia spočíval v rozdelení najpravdepodobnejších následností dvoch písmen na opačné strany klávesnice z dôvodu minimalizovania pravdepodobnosti zaseknutia dvoch po sebe nasledujúcich, raziacich hláv písmen [66]. Oveľa efektívnejším rozložením je „**Dvorak**“ [23], ktorého koncepcia spočíva v minimalizovaní množstva pohybu pri písaní slov (pozn. rozloženie Dvorak bolo vytvorené analýzou slov anglického jazyka).

Je nutné si uvedomiť, že nepočítajúc funkčné tlačidlá, vstup cez klávesnicu, bol vytvorený primárne pre interakciu lexikálnych vstupov pre HCI, ktoré nebolo možné komunikovať cez verbálny kanál. V prípade, že by takáto interakcia možná bola, podľa Engelkampovho modelu by užívateľ nemusel transformovať verbálne symboly cez konceptuálnu vrstvu do umelých motorických sekvencií, ktoré vyžadujú senzomotorické sústredenie a vplyvajú tiež na produkciu kognitívnej záťaže.

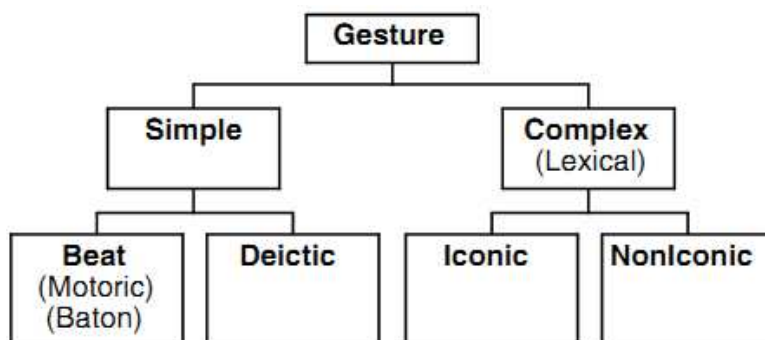
5.2. Gestikulácia a gestúry

Pri verbálnej interakcii je jej častým sprievodným javom gestikulácia. Zaujímavou skutočnosťou je, že pri verbálnych prejavoch obsahujúcich priestorové informácie bola zaznamenaná jej zvýšená hustota [2]. Samovoľné gesturálne prejavy vypovedajú o multimodálnom kódovaní mentálnych reprezentácií, konkrétne o úzkom prepojení motorickej a priestorovej reprezentácie. Tomuto úzkemu prepojeniu navyše zodpovedá aj neuroanatomická štruktúra prepojenia motorickej kôry a dorzálnej priestorovej dráhy, ako aj prepojenie modalít s možnosťou automatizovaných reakcií podľa Engelkampovho modelu.

Trafton & kol. sa vo svojej štúdii [116] snažil pozorovať zákonitosti výskytu gestikulácií, ich kontext, typológiu a koreláciu voči elementárnym lexikálnym prvkom. Analyzoval napríklad gestikulácie verbálnych prejavov zamestnancov fMRI laboratória a meteorológov, u ktorých bol predpokladaný vyšší stupeň používania priestorových reprezentácií pracovnej pamäte a priestorovo orientovaného uväzovania. V oboch prípadoch pozorovaní použil pre rozlišované skupiny expertných pracovníkov (predpoklad vyššej lexikálnej úrovne) a pracovníkov technickej podpory v danej oblasti. Pri analýze korelácii vychádzal z **Golledgeovho lexikálneho rozdelenia** [36] elementárnych priestorových prvkov (spatial primitives) a sústredil podľa nich svoje pozorovanie na lexikálne prvky typu:

- **Poloha / vzťah** – lexikálne prvky popisujúce geometrické vzťahy, egocentrickú (relatívnu) alebo allocentrickú (absolútnu) pozíciu (napr. vedľa, ponad, pozdĺžnosť, atď.)
- **Veľkosť / množstvo** – prvky popisujúce kvantitatívne priestorové informácie (napr. početnosť, hustota, relatívna veľkosť atď.)
- **Transformácia v čase** – prvky vyjadrujúce priestorovú zmenu objektu či prostredia (napr. príchod, odchod, vyprchanie, narastanie, atď.)

Trafton & kol. analyzoval videozáznamy s gestikuláciami podľa schémy gesturálnych charakteristík McNeilla [71]. Schéma rozlišuje základné rozdelenie na jednoduché gestúry (bez priraditeľného sémantického významu) a komplexné gestúry (reprezentujúce určité priestorové charakteristiky lexikálneho výrazu), pričom v každej rozlišuje dve podriadené kategórie (viď. obr. č. 9):



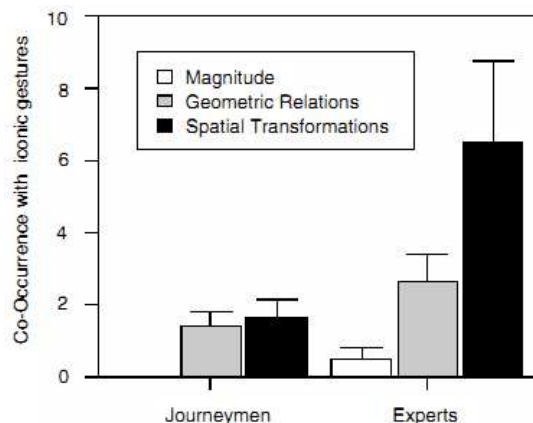
Obr. č. 9: McNeillovo delenie gestúr [71] použité v experimente Trafton & kol.

- Jednoduché gestúry:
 - o **Úderné** – nemajú žiadny lexikálny súvis s hovoreným, majú rytmický charakter súvisiaci s dôrazom intonácie.
 - o **Názorné** – gestúry pre poukazovanie a referovanie na konkrétne objekty alebo časti.

- Komplexné (lexikálne) gestúry:
 - o **Ikonické** – priamo reprezentujú črty lexikálne reprodukovateľného objektu či udalosti (napr. “točiť sa v kruhoch” a točenie dlaňou či prstom).
 - o **Neikonické, metaforické** – majú lexikálny súvis, no nemajú priamo interpretovateľnú vizuálnu koreláciu priestorovej reprezentácie, prípadne ide o symboly zaužívaných metaforických gest (napr. zdvihnutie ukazováku do vertikálnej polohy ako znak pre zdôraznenie pozornosti).

Experiment prebiehal v dvoch fázach. V prvej fáze bolo úlohou participantov verbálne popisovať viditeľný trojrozmerný model a v druhej fáze išlo o voľné reprodukovanie z pamäti, konkrétne problematiky prislúchajúcej ich pracovnej praxi za predpokladu použitia priestorovej imaginácie. Vzhľadom na cieľ štúdie – získať prehľad o vplyve priestorových lexikálnych zložiek indukujúcich gesturálne prejavy – bol týmto experimentom primárne sledovaný ikonický typ gestúr. Ikonické gestúry sú totiž jediná z uvedených typologických kategórií, ktorá predstavuje priamo interpretovateľný súvis medzi priestorovou gesturálnou reprezentáciou a lexikálnym významom.

Výsledky štúdie potvrdili zvýšený výskyt gesturálnych prejavov u expertných pracovníkov v porovnaní so zamestnancami technickej podpory, a poukázali tiež na zvýšený výskyt použitia gestúr pri experimente s voľnou reprodukciou z pamäti v porovnaní s experimentom s viditeľným 3D modelom. Najhodnotnejším prínosom tejto štúdie však bolo identifikovanie priestorového typu ‘transformácie v čase’ z Golledgeových kategórií elementárnych lexikálnych prvkov ako najčastejšie sa vyskytujúceho ikonického typu.



Obr. č. 10: Výsledky štúdie Trafton & kol. použitím Goldegovho delenia gestikulácie poukázali na výrazne zvýšený výskyt priestorového typu gestikulácie u zamestnancov laboratória (experts) a relatívne nízky výskyt u zamestnancov technickej podpory (journeymen) [116].

Toto zistenie poukazuje na existenciu prirodzených asociatívnych typov motorických akcií interpretovateľných v podobe akýchsi affordancií mentálnych reprezentácií. Pre GUI má teda hypotetický význam podrobnejšie mapovanie prirodzených asociácií týchto pohybov z hľadiska potenciálneho využitia pre prirodzenú motorickú interakciu v HCI.

Gestikulácia svojou podstatou spadá podobne ako mimické prejavy pod spomenuté psychofyziologické prejavy. Na rozdiel od mimických prejavov, ktoré reprezentujú vnútorný emočný stav jedinca, gestikulácia má schopnosť vypovedať aj o mentálnych reprezentáciách konkrétnych lexikálnych prvkov a poukázať na niektoré prirodzené, priestorovo asociované, motorické akcie, potenciálne využiteľné aj pre účely intuitívnej HCI interakcie.

5.3. Gesturálna interakcia v HCI

HCI je špecifická tým, že je zo strany užívateľa spoločne vnímateľná ako určitá komunikácia s virtuálnou entitou (napr. formou dialógových okien, alebo príkazového

riadku – t.j. výmenou určitej vstupnej informácie za výstupnú) a z druhej polovice ako manipulácia s nástrojom vo virtuálnom prostredí (napr. editovanie textu, multimédií).

V situácii oboznamovania sa užívateľa s novým prostredím, alebo jeho neznámymi funkcionalitami, je teda hypoteticky možné, že z psychologického hľadiska bude užívateľ vzhľadom na jeho nedostatok zručností či skúseností vnímať interakciu skôr ako personifikovanú formu sekvenčnej komunikácie vstupov a prijímaných výstupov, než ako plynulú dynamickú manipuláciu s nástrojom.

V zmysle McNeillovho rozdelenia prirodzených gesturálnych prejavov je teda teoreticky možné interpretovať aj fundamentálny typ interakcie užívateľa s počítačom, akým je pohyb polohovacím zariadením dvojakým spôsobom: ako deiktickú (názornú – ukazovaciu) gestúru komunikácie alebo ako stelesnenú manipuláciu s kurzorom.

Nakoľko bolo preukázané, že mimické prejavy emočných stavov sú nepomerne frekventovanejšie pri sociálnej interakcii (pri vedomí, že jedinec je niekým pozorovaný) [86], potenciálnou odpoveďou na otázku subjektívneho vnímania HCI by mohli byť pravé dáta namerané z UX štúdií, ktoré snímajú mimickú reakciu užívateľov pri riešení úloh v GUI. Problém s týmito dátami však spočíva v tom, že vzhľadom na etický kódex explicitného informovania participantov o vyhotovení videozáznamu, sú vopred týmto sociálnym faktorom ovplyvnení. V rámci navrhovania interakčných užívateľských rozhraní má teda pre eliminovanie možných interferencií s prirodzeným vnímaním interakcie význam zvážiť tiež jej orientovanie, buď k forme sekvenčnej “komunikácie” alebo dynamickej “manipulácie” (pochopteľné vzhľadom na konkrétny kontext použitia).

Bez ohľadu na subjektívne vnímanie užívateľa sú však v GUI rozlíšiteľné dva fundamentálne typy motorickej interakcie:

- **Symbolická** – charakteristická sekvenčnou výmenou vstupov a výstupov (napr. akékoľvek klávesnicou zadávané lexikálne vstupy a chordické klávesové skratky, gestúry polohovacieho zariadenia).
- **Operatívna / manipulatívna** – charakteristická dynamickou interakciou vo forme manipulácie s objektmi (interakcia podnietená affordanciami, dynamické gestúry ako drag&drop, listovanie na multitouch zariadeniach, atď.).

Napriek ich odlišným charakteristikám sú však oba druhy interakcie dekodované zo strany počítača na explicitne symbolické dáta. Možnosti využitia symbolických a operatívnych gestúr pre HCI sa preto výrazne líšia v závislosti od technologickej implementácie snímania motorickej interakcie. Tá je najbežnejšie determinovaná použitím konkrétneho hardwareového rozhrania snímajúcim buď kontaktný alebo bezkontaktný typ motorickej interakcie:

- **Kontaktná interakcia** (2D plocha) – pri prostrediach WIMP osobných počítačov ide zväčša o snímanie metaforických symbolov vytváraných trajektoriou polohovacieho zariadenia, v kombinácii s použitím napr. pravého tlačidla myši pre "zápis". V prípade post-WIMP zariadení využívajúcich dotykové obrazovky namiesto polohovacieho zariadenia, hrajú gestúry kľúčovú úlohu pre interaktívnu navigáciu v GUI, kde použitie a tvar výslednej trajektórie gestúry vychádza z umelo generovaných affordancií grafického prostredia, ktoré podnecujú užívateľa k ich vnímaniu ako nesymbolických intuitívnych gestúr.
- **Bezkontaktná interakcia** (3D priestor) – Pre snímanie priestorových gestúr sa používa najmä snímanie pomocou real-time počítačovej analýzy video-signálu z kamier (napr. Xbox Kinect) alebo snímanie pomocou špeciálnych polohovacích zariadení (napr. hracie konzoly Wii). Vzhľadom na súčasne obmedzený rozsah využitia priestorových gestúr je táto technológia momentálne rozšírená prevažne

v hracích konzolách alebo ako periférny doplnok pre aplikácie marginálneho použitia.

Pri oboch druhoch interakcie je z pohľadu kódovania na strane počítača nutné rozlišovať tri fázy: 1, fáza motorického pohybu pred vykonaním konkrétnej gestúry alebo manipulácie; 2, gestúra alebo akt manipulácie a 3, fáza motorického pohybu nasledujúca po vykonaní gestúry. Odlíšenie gestúry od sprievodných pohybových fáz je pri použití periférnych polohovacích zariadení a dotykových obrazovkách efektívne riešené podržaním tlačidla alebo priamo dĺžkou dotyku s obrazovkou. Pre rozpoznávanie gestúr pomocou algoritmov počítačového videnia tento proces predstavuje zložitejší výpočtový problém, obzvlášť pri zvážení potreby odlíšenia od ďalších samovoľných pohybov pri interakcii.

Motorická interakcia je efektívne kategorizovateľná a interpretovateľná McNeillovým rozdelením gestúr [x], pričom každá z týchto kategórií ponúka priestor pre rozdielnu oblasť využitia v HCI:

- **Úderné gestúry** – nakoľko ide zväčša o samovoľné prejavy dôrazu intonácie popri verbálnej komunikácii, ich potenciálne využitie pre HCI spočíva v možnej identifikácii emočného rozpoloženia užívateľa (dynamika pohybov) alebo ako potenciálna dátová výpomoc pre rozpoznávanie individuálnych charakteristík pohybov konkrétneho užívateľa. V súčasnosti sú však úderné gestúry skôr zdrojom nežiaducich chýb pri algoritmickej rozpoznávaní a kategorizácii.
- **Názorné gestúry** – sú interpretovateľné ako prostriedok lokalizovania záujmu či navigačnej voľby, tradične formou polohovacieho zariadenia (myš, tablet) alebo priamym dotykovým poukázaním na zariadeniach s dotykovým displejom.
- **Ikonické gestúry** – zahŕňajú oba typy symbolickej aj manipulačnej interakcie v závislosti od použitej technológie. Podľa toho či ide o sekvenčne zadávanú

symbolickú gestúru (napr. myšou „kreslené“ gestúry) alebo dynamickú, manipulačnú gestúru (napr. drag&drop, multitouch listovanie, zväčšovanie alebo rotácia).

- **Metaforické gestúry** – využitie predovšetkým pri vopred špecifikovaných symbolických vstupoch, podobne ako sekvenčné ikonické gestúry.

Špecifickou vlastnosťou motorickej interakcie klasických osobných počítačov (klávesnica, myš, monitor) je špecializácia ľavej a pravej ruky. Vo zvýšenej miere je pozorovateľná najmä u pokročilejších užívateľov, resp. expertných užívateľov komplexnejších, produktívne orientovaných softwareov pracujúcich s priestorom (napr. AutoCAD, Photoshop, 3D max, atď.), kde pravá ruka plní zväčša funkciu precíznej priestorovej lokalizácie záujmu, potvrdzuje výber a má objektovo-manipulatívny charakter. Expertní užívatelia pri tomto type softwareov využívajú ľavú ruku vo zvýšenej miere pre početné chordické gestúry (klávesové skratky), ktoré plnia úlohu urýchleného výberu akcie alebo nástroja. Tento model motorického správania je pozorovateľný už pri elementárnych behaviorálnych prejavoch lateralitý rúk (preferencie využívania jednej z rúk), napríklad pri manipulácii s bežnými predmetmi dennej potreby (uchopenie nádoby do ľavej ruky a precíznejšia a rýchlejšia manipulácia s jej obsahom pravou rukou). Pri HCI bežných osobných počítačov je však pochopiteľné, že takéto hardwareové rozhranie ponúka prirodzenejšiu a plnohodnotnejšiu ergonómiu interakcie primárne pravorukým užívateľom.

Pri zvážení počtu využiteľných (nesystémových) klávesových skratiek a kognitívnej záťaže súvisiacej s potrebou učenia a pamätania špecifických kombinácií kláves, sa však ako efektívnejšie riešenie ponúka alternatívne využitie touchpadu pre zadávanie gestúr (minimálne pre hardware prenosných počítačov). Duplicita funkcie dotykovej plochy touchpadu pri použití periférneho polohovacieho zariadenia tak môže byť plne nahradená využitím tejto plochy pre zadávanie priestorovo reprezentovaných

gesturálnych symbolov ľavou rukou, a to paralelne pri použití pravej ruky pre precízne lokalizovanie/polohovanie kurzora. Gestúry majú voči klasickým klávesovým skratkám výhodu v tom, že sú ľahšie zapamätateľné a intuitívnejšie vďaka svojej priestorovej reprezentácii a pôvodu v prirodzených obrazových a motorických schémach. Ďalšou výhodou je prakticky neobmedzený počet definovania vlastných gesturálnych symbolov. Okrem zadávania symbolických gestúr môže touchpad navyše suplovať aj funkciu jednoduchších alebo doplnkových manipulačných operácií, obzvlášť v prípade rozširujúcich sa touchpadov s podporou multitouch technológie (napr. použitie známych dvojprstových gestúr pre operácie zväčšovania, otáčania, atď.).

Na rozdiel od symbolických gestúr, manipulačné operácie sú špecifické okrem potreby dynamickej vizuálnej odozvy, aj potrebou voľby typu funkcie či nástroja ešte pred ich použitím. Symbolické operácie (gestúry, funkčné tlačidlá, príkazy) môžu prebiehať v dvoch rôznych scenároch:

- **Výber dátových objektov -> voľba typu akcie** – tento typ scenára je špecifický a výhodný hlavne pre situácie, kde je potrebné vykonať identickú akciu na väčšom počte dátových objektov. Odhliadnuc od procesu selekcie objektov, tento scenár svojou abstraktnou podstatou nevyužíva žiadne výhody priestorovej konceptualizácie. Užívateľskými nevýhodami tohto scenáru sú abstraktnosť a jeho sekvenčná charakteristika, v prípade zložitejších operácií potreba komplexného plánovania a s tým súvisiace riziko zvýšenej kognitívnej záťaže.
- **Voľba typu akcie (nástroja) -> vyber dátového objektu** – je typickým scenárom analogickým k manipulácii človeka s objektmi v prirodzenom prostredí. Pri HCI sa v tomto prípade najprv volí funkcia alebo nástroj, ktorý je následne využívaný pre požadované úpravy a akcie na zvolenom objekte. Hlavnou nevýhodou tohto

scenáru je však možnosť pracovania súčasne len s jedným nástrojom na jednom objekte.

Rozdelenie úloh ľavej a pravej ruky v HCI má navyše z neuropsychologického hľadiska ergonómie pozitívny efekt na rozklad kognitívnej záťaže nielen medzi dorzálnou a ventrálnou neurálnou dráhou, ale aj medzi kontralaterálnymi dráhami motorickej kôry, dvoch samostatne pracujúcich hemisfér. Senzomotorickú interakciu síce vo všeobecnosti nie je možné kvôli jej komplexnosti generalizovať na konkrétne kortikálne oblasti, no je možné popísať korelácie špecifických oblastí pri konkrétnych typoch interakcie. Pre symbolický typ gesturálnej interakcie boli preukázané korelácie s Brocovou a Wernického oblasťami [114], ktorým je prisudzovaná tiež zodpovednosť za produkciu a porozumenie reči. Pre dynamický, manipulatívno-operatívny typ interakcie sú špecifické napríklad korelácie oblastí zodpovedné za koordinácie medzi zrakom a pohybom ruky (*eye-hand movements*), konkrétne *temporo-parietal junction* (TPO) [37], ale aj subkortikálne oblasti ako bazálne gangliá [96] alebo mozoček (zodpovedný za činnosť rutinne zautomatizovaných pohybov) [72].

Pri oboch typoch interakcie sú pochopiteľne súčasne aktivované aj ďalšie mozgové oblasti ako somato-senzorická kôra (spracovávanie dotykových a iných telesných vnemov), primárna motorická kôra (vôľou riadené pohyby), prefrontálna kôra (plánovanie a vyššie kognitívne funkcie), ale v neposlednom rade komplexný vizuálny systém.

Pre eliminovanie nežiaducej kognitívnej záťaže pri HCI má teda z pohľadu neuropsychológie význam najmä paralelné distribuovanie a rozkladanie interakčných kognitívnych mechanizmov medzi samostatne pracujúce neurálne dráhy, aké popisuje už spomínaná teória dorzálnnej a ventrálnnej dráhy. Charakteristickým problémom súčasných GUI je využívanie relatívne obmedzeného okruhu interakčných modalít, a to nielen z perцепčnej, ale aj z motorickej stránky HCI. Ide hlavne o nadmerné zaťaženie užívateľa explicitnými, vizuálne prezentovanými, sémantickými informáciami, ako aj

potrebou neverbálnej, sekvenčnej interakcie, ktorých neurálne spracovanie je prisudzované ventrálnej dráhe.

V nasledujúcej kapitole sa pokúsím poukázať na niektoré konceptuálne riešenia GUI, ktoré v konečnom dôsledku využívajú vo zvýšenej miere potenciál mechanizmov dorzálnej dráhy, a to formou priestorovej konceptualizácie prostredia a dynamicky orientovanou, primárne manipulatívnu, senzomotorickou interakciou.

6. Príklady aplikácii priestorovej konceptualizácie v kontexte historického vývoja

Pri pohľade do histórie HCI je možné pozorovať určité smerovanie vývoja užívateľských rozhraní. Nepočítajúc éru sálových počítačov a diernych štítkov, prostredie príkazového riadku CLI možno považovať za historicky prvé prostredie využívajúce grafické rozhranie. Charakteristickými vlastnosťami CLI je sekvenčná, textová forma komunikácie s presne definovanou syntaxou a potreba apriórnej znalosti veľkého množstva výrazov pre plnohodnotnú interakciu. Bez predchádzajúcej direktívnej výuky práce s prostredím nemá nový, neskúsený užívateľ reálne možnosti interagovať s prostredím. Ďalšou historicky dôležitou metou užívateľských prostredí sú prostredia GUI. Popri nesporných historických faktoroch cenového a rozmerového sprístupnenia počítačov, bolo implementovanie grafického prostredia výrazným krokom vpred pri integrovaní HCI do každodenného života, a to nielen pre výpočtové účely, ale aj ako prostriedok získavania a výmeny informácií, multimédií a v neposlednej rade ako komunikačný sociálny nástroj. Vzhľadom na evolučne nemenné, štandardné hardwareové rozhranie strany užívateľskej interakcie (monitor, klávesnica a polohovacie zariadenie) prechádzali užívateľské prostredia zmenami hlavne po perцепčnej a konceptuálnej stránke HCI. V porovnaní s textovou formou CLI sa prostredia GUI vyznačujú rozsiahlym využívaním grafických prostriedkov a obsahujú tiež početnú sadu užívateľských metafor a intuitívnych prvkov, umožňujúcich neskúsenému užívateľovi samostatné oboznamovanie sa s prostredím a jeho funkciami.

Počiatky GUI možno sledovať pravé akýmsi pridaním vrstvy priestorovej konceptualizácie v textových prostrediach – napr. používaním odrážok či pracovaním s prázdnu plochou pre organizovanejšie výpisy informácii na obrazovke, alebo rôznymi inými princípmi zprehľadňovania textovej informácie známymi prevažne z oblasti typografie.

Ďalším historickým posunom od textového ku dnes známemu grafickému rozhraniu bola okrem zavedenia jednoduchých tabuľkových rámcov, metafora práce na virtuálnom dokumente ako na papierovom dokumente prostredníctvom tzv. WYSIWIG textových editorov (WYSIWIG = „*what you see is what you get*“, čo znamená, že aktuálny vizuálny stav dokumentu, ktorý užívateľ vidí na obrazovke vo fáze jeho úprav, je aj finálnym stavom dokumentu). Dôležitým novým prvkom tohto editora bola rozšírená funkcia kurzoru (ovládaného spočiatku len klávesnicou), ktorý sa už mohol pohybovať po dvojrozmernom priestore dokumentu a predstavoval isté allocentrické stelesnenie polohy užívateľa vo virtuálnom prostredí.

Implementáciou pokročilých nástrojov pre formátovanie textu postupne pribudli tiež funkčné tlačidlá, dodnes používaná funkcia vypadkových menu či jednoduché dialógové okná. S vývojom výpočtovej sily hardwareu tiež postupne pribúdala aj grafická podpora v podobe nahrádzania textových popisov klikateľných tlačidiel piktograficky reprezentovanými ikonami, či využitie monokulárnych indikátorov hĺbky pre posilnenie vizuálneho oddelenia jednotlivých prvkov alebo častí pracovného priestoru. Monokulárne indikátory hĺbky sú tiež často využívaným vizuálnym prostriedkom pri dotváraní ilúzie hĺbky ako tretieho rozmeru GUI prekrývajúcich sa dvojrozmerných plôch aplikačných či dialógových okien.

Ďalším dôležitým krokom priestorovej konceptualizácie v historickom vývoji GUI bola popri allocentrickom pohybe kurzora ponad hierarchickými vrstvami okien aplikácii, implementácia egocentrického typu pohybu užívateľa v prostredí. Ide o pohyb užívateľa z pohľadu prvej osoby, ktorý mu umožňuje rýchlejší a intuitívnejší prechod cez obsiahlejšie užívateľské scény, napríklad využívaním metafory prechodu cez viaceré pracovné plochy (*virtual desktops*). Užívateľsky prívetivý, dynamický a priestorovo

vizualizovaný prechod umožňuje napr. funkcia „Mission Control“ (pôvodne „Exposé“) operačného systému OS X [89]. Funkcia je hybridným spojením klasického konceptu WIMP a abstrakcie konceptu inšpirovaného z alternatívneho rozhrania **škálovateľných užívateľských prostredí** (ZUI, z angl. *Zoomable User Interface*).



Obr. č. 11: Prechod kategorickým rozložením okien Mission Control.

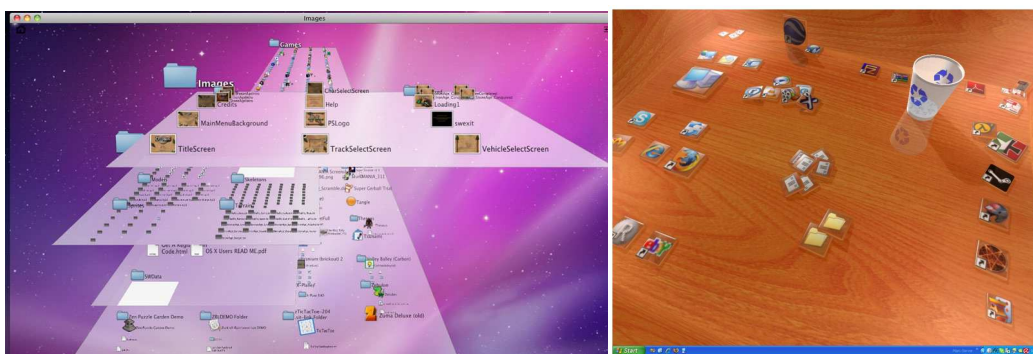
ZUI je v porovnaní s klasickým WIMP rozhraním alternatívna forma GUI, kde sa na rozdiel od mnohokrát neprehľadného vrstvenia aplikačných okien v iluzórnom treťom rozmere, nachádza celé užívateľské prostredie v jedinej úrovni dvojrozsmernej roviny a užívateľ sa pohybuje dynamicky ponad celú plochu zväčšovaním alebo zmenšovaním zobrazovaného pola jej informácií (zväčša tzv. „scrollovaním“). Hierarchická štruktúra dát je reprezentovaná možnosťou nekonečného vnárania diskretné ohraničených podoblastí nižšej úrovne v rámci hraníc jej nadriadenej oblasti. Príkladmi takýchto prostredí sú konceptuálne pôsobivá demonštrácia ZUI rozhrania formou webstránky <http://zoomism.com/> [123], aplikácia prieskumníka súborov Eagle mode [24], alebo z interakčného hľadiska analogické prostredia využívajúce princípy ZUI, ako sú geograficky špecializované aplikácie Google Earth, či Google Maps [39].



Obr. č. 12: Ukážka ZUI prostredia Eagle Mode – rámce priečinkov obsahujú ďalšie hierarchicky podriadené priečinky, pričom dynamický priechod je medzi nimi umožňovaný približovaním.

ZUI je jedným z príkladov riešenia poukazujúceho na to, že okrem implementácie priestorovej konceptualizácie do presunu medzi užívateľskými scénami, môže takéto rozšírenie v GUI sprehľadniť a zefektívniť prácu aj v iných jeho oblastiach, ako je prechádzanie súborovým systémom alebo administrácia aplikácií či okien.

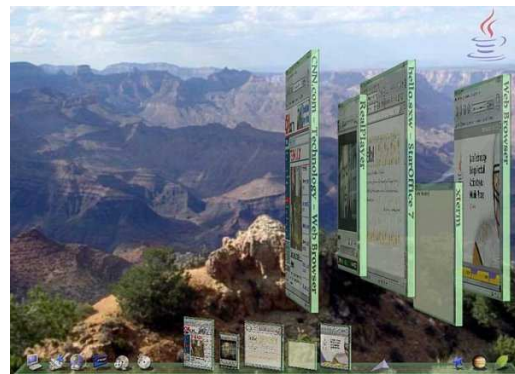
S rastúcou výpočtovou silou grafických kariet ako aj procesorov sa postupne rozvíjali aj možnosti využitia 3D grafiky od jednoduchých objektov až po komplexne virtuálne prostredia. Implementácie priestorovej zložky možno nájsť v mnohých globálnych či parciálnych aplikáciách GUI rozšírenia metaforu pracovnej plochy o trojrozmerné prostredie (napr. 3D fileSpace [22], BumbTop [21]), alebo tiež rozšírením o trojrozmernú administráciu okien aplikácií (napr. SphereXP [3], Tactile3D [4], Project Looking Glass [5]).



a)



b)



Obr. č. 13: a) Modifikácie pracovnej plochy 3D fileSpace a BumpTop,
b) trojrozmerné spravovanie okien Project looking glass.

Z interakčného hľadiska ide však pri plne trojrozmerných konceptoch skôr o akúsi demonštráciu súčasných technológií počítačovej grafiky, než o efektívne použiteľné užívateľské rozhranie, nakoľko pohyb a manipulácia v takomto 3D prostredí sú silne obmedzené neprirodzenou motorickou interakciou bežné používaných, polohovacích zariadení.

6.1 Post-WIMP a dotyková interakcia

Okrem neustáleho posilňovania výpočtovej kapacity hardwareu zariadení sa iteratívnym zhromažďovaním empirických skúsenosti, vyhodnocovaním štúdií UX a použiteľnosti popisovaných na začiatku tejto práce, evolúcia prostredí GUI postupne prepracovala k potrebe obohacovania nielen percepčnej časti komunikovanej informácie, ale aj k obohacovaniu interakcie zo strany užívateľa k počítaču. Nakoľko však vývoj komplexnejších GUI (napr. operačné systémy) spadá primárne pod komerčný sektor, existuje bohužiaľ len veľmi slabá, verejnosti prístupná, faktografická dokumentácia, poukazujúca na pozadie a príčiny aplikovania rôznych funkcionalít, sledovateľná viac-menej len historickým vývojom softwareu a hardwareu finálnych produktov na trhu.

Príchod súčasných post-WIMP mobilných zariadení, tabletov či tzv. smartfónov, často disponujúcich dotykovými obrazovkami, umožnil svojou užívateľsky prirodzenejšou a intuitívnejšou formou interakcie sprístupnenie HCI aj pre menej skúsených užívateľov. Potenciálnym vysvetlením môže byť, že takáto priamejšia forma interakcie bez nutnosti neprirodzeného periférneho polohovacieho zariadenia (myš, grafický tablet), poskytuje užívateľovi dojem prirodzenejšej manipulácie. Zobrazovaná informácia „pod prstom“ sa totiž javí byť priamo fyzicky ovplyvňovaná, a to bez nutnosti reprezentácie kurzora ako indikátora polohy užívateľa vo virtuálnom prostredí. Súčasné post-WIMP dotykové zariadenia navyše efektívne kombinujú vierohodne iluzórne affordancie modernej počítačovej grafiky (napr. imitácie mechanicky uchopiteľných či posúvateľných častí panelov grafickým znázornením drážok, spolu s prepracovanejšou dynamicky reagujúcou animovanou odozvou týchto vizuálnych elementov). Poskytujú tak prirodzenejšiu implementáciu užívateľských metafor a priamej, dotykovej gesturálnej interakcie (ako napr. listovanie fotografií v galérii, multitouch operácie ako približovanie alebo natáčanie, atď.). Post-WIMP zariadenia teda reálne poukazujú na efektívnosť využívania intuitívnych konceptov a zároveň na význam intenzívnejšieho prepojenia percepčnej a motorickej stránky HCI.

6.2 NUI a konceptuálne riešenia pre GUI

V súčasnosti je trend posilňovania intuitívneho prepojenia percepčnej a motorickej stránky v terminológii HCI definovaný konceptom tzv. užívateľsky prirodzeného prostredia **NUI (*natural user interface*)**. NUI predstavuje istú idealistickú víziu interakcie, kde užívateľ pri komunikácii s počítačom využíva jemu prirodzené interakčné prostriedky určené kontextom situácie, na rozdiel od CLI, kde potrebuje pre prácu s prostredím rozsiahle apriórne znalosti použitia symbolických výrazov a ich syntaxe, alebo na rozdiel od klasického GUI, ktoré síce užívateľovi ponúka čiastočnú možnosť samostatného

oboznamovania sa, no pre plnohodnotnú interakciu je potrebný istý čas strávený samovýukou práce s týmto prostredím.

V konceptuálnej vízii NUI by teda užívateľ nemal byť nútený k žiadnym neprirodzeným abstraktným reprezentáciám navigačných či operačných akcií prostredia.

Príkladmi takto implementovaných komplexnejších prostredí NUI sú Perceptive Pixel [27], Microsoft Surface [73], alebo projekt interaktívneho zvukového generátora reacTable [48].



Obr. č. 14: Koncepty NUI: Microsoft surface a reacTable.

Uvedené príklady využívajú inovácie ako rozšírenie multitouch technológie pre podporu súčasnej interakcie viacerých užívateľov, bezdrôtovú synchronizáciu a výmenu dát s elektronickými zariadeniami položenými na kontaktnú plochu, alebo plne dynamickú vizuálnu odozvu interakcie grafického prostredia. Problém takýchto prostredí však často býva v ich obmedzenej univerzálnosti použitia pri dodržaní princípov NUI, čím sú v rôznych krajných scenároch nútené využívať metaforizované princípy známe z klasických rozhraní GUI.

Potenciálnym riešením tohto obmedzenia môže byť v budúcnosti koncept tzv. **multimodálnych užívateľských prostredí** (MUI, z angl. *Multimodal User Interface*), kde by mal mať užívateľ možnosť komunikovať s počítačom popri bežne používaných interakčných senzomotorických vstupoch, napríklad aj prostredníctvom verbálnej interakcie. Bohužiaľ počítačové rozpoznávanie reči dosahuje v súčasnosti stále relatívne

vysokú mieru chybovosti, takže prostredia MUI zostávajú skôr akousi konceptuálnou víziou budúcnosti, než plnohodnotnou alternatívou. Prostredia MUI tak v prípade kombinácie verbálneho komunikačného kanálu a senzomotorickej interakcie umožňujú veľmi vhodné mapovanie spomínanej možnosti nízkoúrovňovej interakcie Engelkampovho modelu, čím je možné znížiť mieru kognitívnej záťaže spôsobenú transformáciami modalít cez konceptuálnu vrstvu a priblížiť sa k teoreticky prirodzenejšej interakcii HCI.

Zaujímavou alternatívou GUI kombinujúcou princípy prostredí MUI a NUI je rozhranie tzv. **rozšírenej reality (*augmented reality*)**. Príkladom takýchto konceptov sú interaktívne monokulárne monitory EyeTap [7] alebo Project Glass od Google, kde užívateľ v konečnom dôsledku vníma popri pôvodných vizuálnych vnemoch z prirodzeného prostredia aj doplnkové vizuálne informácie o prostredí ich premietaním na transparentnú plochu umiestnenú tesne pred jeho okom. Oba systémy využívajú monokulárny aparát s implementovanou kamerou, ktorou systém získava vstupné vizuálne informácie o okolí. Signál je ďalej spracovávaný v zariadení algoritmami počítačového videnia a podľa použitej softwareovej aplikácie sú výstupné dáta premietané užívateľovi graficky naspäť do zorného pola, čím sa mu rozširuje rozsah vizuálnych vnemov o „rozšírenú realitu“. Project Glass navyše interaktívne spracúva aj verbálne užívateľské vstupy, čím načrtáva potenciálnu víziu HCI budúcnosti. Podobným konceptuálnym riešením budúcnosti je koncept ‘Sixth Sense’ Pranav Mistryho [75] z dielne MIT, ktorý však premieta výstupný signál do reálneho trojrozmerného priestoru a pre interakciu používa gesturálne vstupy.



a)



b)

Obr. č. 15: Koncepty rozšírenej reality:

a) Project Glass spoločnosti Google, b) Sixth Sense Pranav Mistryho.

Koncept Sixth Sense je však príkladom, ktorý čelí viacerým potenciálnym komplikáciám súvisiacich s neintuitívnym a nespoľahlivým zadávaním symbolických vstupov. Tie rieši elegantným spôsobom Project Glass ich interakciou cez verbálny kanál. Budúcnosť užívateľských prostredí počítačov, ktoré by spĺňali princípy NUI má však okrem možnosti implementácie multimodálnej interakcie ešte jednu perspektívu, a to v ich užšej špecializácii na konkrétny účel. Čím menej univerzálne - konkrétnejšie použitie totiž od počítača očakávame a potrebujeme, tým jednoduchšia a konkrétnejšia sadá prostriedkov môže byť použitá k prirodzenej, intuitívnej interakcii v jeho užívateľskom prostredí. Príkladom NUI môžu byť aj zariadenia, pri ktorých si svojou presne vymedzenou oblasťou použitia užívateľ často ani neuvedomuje, že interaguje s počítačom (napríklad GUI niektorých moderných integrovaných počítačov v automobiloch či fotoaparátach). Takéto účelovo orientované rozhrania plnia úlohu

konkrétnych nástrojov a efektívne obchádzajú komplikácie pri neprirodzených virtuálnych prostrediach, ako sú zmeny voľby funkcie či nástroja, pri prechodoch medzi komplexnejšími užívateľskými scénami alebo komplikáciách v podobe psychologickkej interferencie reálneho a virtuálneho prostredia.

6.3 Potenciál priestorovej konceptualizácie pre HCI

Pozorovaním charakteristík vývoja GUI od CLI k NUI je nepochybne zjavné, že výhody priestorovo orientovaných riešení sa násobia sústredením nielen na jej perцепčnú, ale aj na motorickú stránku, pre ktorú je typická dynamická interakcia v priestore.

Využitie priestorovej konceptualizácie by sa dalo spolu s niektorými vyššie popísanými princípmi kognitívnej ergonomie sumarizovať z pohľadu dlhodobej a krátkodobej interakcie.

Z pohľadu perцепčnej stránky:

- **Pri krátkodobej (okamžitej) interakcii** – ponúka efektívne sprehľadňovanie vnímania viacrozmerných alebo rôznorodých dát pomocou implicitných priestorových vzťahov. (napr. intenzívne využívanie Gestalt princípov pre vizuálnu kategorizáciu objektov do skupín podobnosti v 2D, použitím jednoduchých indikátorov hĺbky pre iluzórne usporiadanie objektov rozšíriteľné aj na tretí rozmer).
- **Pri dlhodobejšej interakcii** – pomáha vytvárať určité mentálne užívateľské mapy prostredia a zabezpečuje rýchle a nekonfliktné používanie komplexnejších prostredí stabilným rozmiestnením vizuálnych prvkov (napr. využívanie priestorových metafor ako metafora pracovnej plochy, či textového dokumentu a inými ustálenými konvenciami priestorových zvyklostí, ako napr. umiestňovanie vyhľadávacieho pola na webstránkach v pravom hornom rohu, alebo centralizovaním častejšie sa meniaceho obsahu smerom na stred

obrazovky popri distribúcii navigačných, funkčných či iných stálejších užívateľských elementov na perifériu.

Z pohľadu motorickej stránky:

- **Pri krátkodobej (okamžitej) interakcii** – umožňuje akýkoľvek dynamický pohyb a manipulačné operácie užívateľa v priestore GUI. Užívateľ pri tomto type temporálnej interakcie hlavne motoricky reflektuje vizuálne komunikované affordancie vzhľadom na cieľ jeho akcie (napr. allocentrický typ pohybu kurzorom po obrazovke, či rozsiahlych dokumentoch, manipulácia s dátovými objektmi alebo interakčnými prvkami).
- **Pri dlhodobejšej interakcii** – opakovaným používaním štandardizovaných motorických akcií buduje u užívateľa motorickú pamäť. Nadobudnutím sady nových zautomatizovaných zvykov, alebo využitím metaforizovaných motorických zvykov osvojených v prirodzenom prostredí (mimo HCI) sa posilňuje rýchlosť a prirodzenosť reakcie užívateľa v prostredí (napr. gestúry a klávesové skratky odľahčujú kognitívnu záťaž od potreby hľadať tlačidlo konkrétnej akcie v paneli nástrojov či vo výpadkových menu). Gestúry sú navyše prirodzene metaforizovateľné využitím známych priestorovo reprezentovaných obrazových schém reálneho sveta (špeciálne pri dotykových obrazovkách – listovanie v dokumentoch, alebo multitouch gestúry). Gestúry sú v porovnaní s klávesovými skratkami navyše ľahšie zapamätateľné nakoľko sú kódované okrem motorickej pamätej stopy aj vizuálnou mnemotechnickou asociáciou 2D trajektórie „nakresleného“ symbolu akcie.

Po percepčnej stránke by sa teda dalo zhrnúť, že priestorová konceptualizácia má význam z pohľadu dlhodobej interakcie pre urýchlenie hľadania požadovanej informácie, resp. filtrovania aktuálne nepodstatných informácií (pomocou metafor a konvencií) a z pohľadu reaktívnej (temporálnej) úrovne interakcie má význam pre širší, efektívnejší

prietok konzumovania novej informácie (jej prehľadným usporiadaním alebo obohatením o implicitné súvislosti priestorových vzťahov).

Priestorová konceptualizácia motorickej stránky HCI ma význam hlavne pri mapovaní a prispôbovaní interakčných vstupov k človeku prirodzeným tendenciám jeho reaktívnych pohybov (obrazové schémy, gestikulácia), nakoľko mnohé kognitívne mechanizmy vychádzajú z princípov popisovaných teóriou stelesnenej kognície (embodied cognition).

Ideálne GUI prostredie by teda malo vychádzať z princípu ponúknutia iba takých affordancií vyplývajúcich z vizuálnych stimulov, pre ktoré existuje človeku prirodzená motorická reakcia, a to takým spôsobom, aby mohol užívateľ intuitívne, v reálnom čase dynamicky vnímať reakciu prostredia na svoje vstupy, bez potreby apriórnych skúseností.

7. Záver

Prudký rozvoj informačných technológií a výpočtovej techniky v posledných desiatkach rokov spôsobil, že interakcia človeka s počítačom vo virtuálnej forme grafických užívateľských prostredí sa dostala na poprednú pozíciu každodenného života veľkej časti produktívnej populácie civilizovaného sveta. Takáto interakcia, špeciálne pri použití bežných osobných počítačov, je pre človeka do veľkej miery neprirodzená, nielen z pohľadu klasickej, ale aj kognitívnej ergonómie. Častým problémom súvisiacim s neprirodzenosťou interakcie HCI je neprítomnosť prirodzených implicitných vodítok, navádzajúcich užívateľa ku kontextom situácie určenej, konkrétnej interakcii, popisovaných Gibsonom ako affordance. Tie sú v GUI kompenzované využívaním umelých užívateľských metafor, konceptualizovaním vizuálnych informácií použitím Gestaltových princípov pre vizuálne štruktúrovanie dokumentov, alebo inými implicitnými prostriedkami vyzorovanými prevažne zo štúdií vizuálnej percepcie. Prílišná komplikovanosť, interferencie navigačných a sémantických prvkov alebo neprehľadné zoznamy často spôsobujú užívateľovi interakčné komplikácie a produkujú tak nežiaducu kognitívnu záťaž.

Produkcia kognitívnej záťaže v GUI sa prejavuje najmä prehltením exekutívnej zložky pracovnej pamäte, ktorej dôsledkom je zúženie pozornosťných mechanizmov na konfliktný element a zníženie pozornosti voči vonkajším vnemom z prostredia. Okrem exponenciálnej náchylnosti ku kognitívnemu preťaženiu pri užívateľsky neriešiteľných interakčných problémoch je pravdepodobnosť záťaže v GUI zvyšovaná aj neprirodzenosťou vnímania informácie takmer výhradne prostredníctvom vizuálneho kanálu. Na zvýšenie efektívnosti pamäťového spracovania viacerými perцепčnými kanálmi poukazujú napr. štúdie multimodálnych teórií.

Engelkampova multimodálna teória poukazuje na fakt, že pamäťový kanál môže byť tvorený aj inými modalitami, napríklad motorickou akciou. Engelkamp navyše popisuje

špecifické kombinácie senzomotorických akcií, ktorých spracovanie nutne prechádza konceptuálnou - sémantickou vrstvou (verbálna reprodukcia vizuálneho vnemu a motorická reprodukcia sluchového), a tých ktorých spracovanie cez konceptuálnu vrstvu nutne prechádzať nemusí (vizuomotorická koordinácia pohybov a verbálna reprodukcia sluchového vnemu). Senzomotorický typ nekonceptuálnej interakcie teda poukazuje na možnosť ich nízkoúrovňového kognitívneho prepojenia, s vyšším potenciálom automatizovateľných reakcií.

Tuto hypotézu podporuje aj teória delenia dorzálnej a ventrálnej neurálnej dráhy, kde dorzálnej dráhe je prisudzované spracovanie priestorových a pohybových vizuálnych informácií ako aj spoluúčasť na senzomotorických koordináciách a ventrálnej dráhe mentálne procesy ako identifikácia objektov, rečové funkcie ako aj niektoré ďalšie sémantický orientované funkcie. Samostatnosť týchto neurálnych dráh, resp. neuropsychologická príbuznosť špecificky orientovaných susediacich oblastí v týchto dvoch dráhach, silne korelujú s behaviorálnymi výsledkami poukázanými Engelkampom.

Pomocou rozšírenia Engelkampovho modelu o hypotézu reaktívnych foriem senzomotorických kombinácií interakcie a prepojením s teóriou dorzálnej a ventrálnej dráhy som chcel v mojej práci poukázať na význam priestorovej konceptualizácie a to špeciálne v problematike kognitívnej záťaže v HCI a GUI. Ďalším argumentom podporujúcim načrtnutý model, je efektívnosť interakčných riešení využívajúcich implicitné priestorové vodítka ako aplikovanie Gestalt princípov, typografie, dynamicky reagujúcich (pohyblivých) grafických prvkov, či z hľadiska motoriky gesturálnou interakciou v post-WIMP zariadeniach. Všetky spomenuté riešenia sú z neuropsychologického hľadiska v rámci modelu interpretovateľné ako istá forma rozkladu potenciálnej kognitívnej záťaže v GUI, realizovaná distribúciou jej podliehajúcich interakčných procesov medzi nezávislé modalities, paralelne pracujúcich neurálnych oblastí.

Ďalšou prácou je teoreticky možné načrtnutý model rozšíriť o popísanie iných, paralelne pracujúcich vstupno-výstupných, nezávislých, interakčných kanálov ako je Engelkampom popísaný vizuomotorický a verbálny kanál (skúmaním prepojení somatosenzorických, či olfaktorických vstupov atď.). Podobne ma teoretickú perspektívu v rámci neuropsychologického prístupu k problematike pokúsiť sa analyzovať a špecifikovať podliehajúce kognitívne mechanizmy už popísaných nezávislých modalít a pokúsiť sa tak identifikovať a rozdeliť ich na skryté „submodality“, podobne ako je zdanlivo jednotný, vizuálny kanál rozdeliteľný na svoju modalitne nezávislú, priestorovú a sémantickú zložku.

Bibliografia

- [1] Accot, J. & Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces, ACM Conference on Human Factors in Computing Systems, Minneapolis, Minnesota, p 73-80.
- [2] Alibali, M. W., Heath, D. C. & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169–188
- [3] Alternative 3D GUI. [online].[cit.25.5. 2012].URL: <http://www.spheresite.com/>
- [4] Alternative 3D GUI. [online].[cit.25.5. 2012].URL: <http://www.tactile3d.com/>
- [5] Alternative 3D Window-Manager. [online].[cit.25.5. 2012].URL: <http://javadesktop.org/articles/LookingGlass/>
- [6] Atkinson, R.C.; Shiffrin, R.M. (1968). Chapter: Human memory: A proposed system and its control processes. In Spence, K.W.; Spence, J.T. *The psychology of learning and motivation* (Vol. 2). New York: Academic Press. pp. 89–195.
- [7] Augmented reality concept. [online].[cit.25.5. 2012].URL: <http://eyetap.org/>
- [8] Badcock, J. C., Whitworth, F. A., Badcock, R. D. & Lovegrove, W. J. (1990). Low-frequency filtering and the processing of local–global stimuli. *Perception*, 19, 617–629.
- [9] Baddeley, A .D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation. Advances in research and theory* (Vol. 8), New York: Academic Press, pp. 47–89.

- [10] Baddeley, A. D. & Wilson, B. A. (2002). Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia* 40 (10), 1737–43.
- [11] Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci. (Regul. Ed.)* 4 (11): 417–423.
- [12] Ball W., Tronick E. (1971). Infant responses to impending collision: Optical and real. *Science*; 171, 818-820.
- [13] Borthwick, C. (1996). Racism, IQ and Down's Syndrome. In *Disability & Society*, Vol 11, No. 3, 403-410.
- [14] Brooks, L. (1968). Spatial and verbal components of act of recall. *Canadian Journal of Psychology*, 22(5), 349-368.
- [15] Chandler, P. & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction* 8 (4), 293–332.
- [16] Chase, W.G. & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology* 4 (1), 55–81.
- [17] Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*, New York: Academic Press, 27-63.
- [18] Collin, C.A. & McMullen, P.A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception & Psychophysics*, 67 (2), 354–364.
- [19] Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology* 79 (4), 347–362.

- [20] Cowan, N. (1995). Attention and memory: An integrated framework. New York: Oxford University Press.
- [21] Desktop with physical laws in 3D. [online].[cit.1.1. 2010].URL: <http://www.bumptop.com/>
- [22] Desktop solution FileSpace. [online].[cit.25.5. 2012].URL: <http://www.marcmoini.com/>
- [23] Dvorak, A, et al. (1936). Typewriting Behavior. American Book Company.
- [24] EagleMode - Zoomable User Interface [online].[cit.25.5. 2012].URL: <http://eaglemode.sourceforge.net/>
- [25] Engelkamp, J. (1998). Memory for actions. Hove, England: Psychology Press.
- [26] Ettlinger, G. (1990). "Object vision" and "spatial vision": the neuropsychological evidence for the distinction. Cortex 26 (3), 319–41.
- [27] Example of NUI. [online].[cit.25.5. 2012].URL: <http://www.perceptivepixel.com/>
- [28] Farrington, J. (2011). Seven plus or minus two. Performance Improvement Quarterly 23 (4), 113–6.
- [29] Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 121(3), 262–269
- [30] Franklin, N. and Tversky, B. (1990). Searching imagined environments. Journal of Experimental Psychology: General, 119, 63-76.
- [31] Friedman-Hill, S. R., Robertson, L. C., Treisman, A. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions.

- [32] Gibson E.J., Walk R.D. (1960). The “visual cliff.” *Scientific American*. 64-71.
- [33] Gibson, J. J. (1977). The Theory of Affordances. In *Perceiving, Acting, and Knowing*, Eds. Robert Shaw and John Bransford, ISBN 0-470-99014-7.
- [34] Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum.
- [35] Goldberg T. (1998). *Psychological Medicine*, 28: 665–673
- [36] Golledge, R.G. (1995). Primitives of spatial knowledge. In T. L. Nyerges, D. M. Mark, R.Laurini & M. J. Egenhofer (Eds.), *Cognitive aspects of human-computer interaction for geographic information systems*, pp. 29–44.
- [37] Gomi, H. (2008). Implicit online corrections in reaching movements. *Current Opinion in Neurobiology* 18 (6), 558–564.
- [38] Goodale, M.A, Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience* 15 (1), 20–5.
- [39] Google Maps. [online].[cit.25.5. 2012].URL: <http://maps.google.com/>
- [40] Haberlandt K. (1999). *Human Memory: Exploration and Application*. Allyn and Bacon a Viacom Company.
- [41] Hadjikhani, N., Liu, A.K., Dale, A.M., Cavanagh, P., Tootell, R.B. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Natural Neuroscience* (3), 235-41.
- [42] Hemsley D. (1996). Schizophrenia. *Behavior Modification*.20: 139–150.

- [43] Hess, E.H.& Polt, J.M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143, 1190-1192.
- [44] Hotchkiss, G., Alston, S. & Edwards, G. (2005). Google Eye Tracking Report. Enquiro Search Solutions Inc.
- [45] Huberle E. & Karnath H.O. (2011). The role of temporo-parietal junction (TPJ) in global Gestalt perception. *Brain structure & function* (12), 1-12.
- [46] Humphrey, G. (1924). The psychology of the gestalt. *Journal of Educational Psychology*, 15 (7), 401–412.
- [47] Iqbal, S. T., Zheng, X. S. & Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. Extended abstract of the ACM Conference on Human Factors in Computing Systems, Vienna.
- [48] Interactive NUI music toy. [online].[cit.25.5. 2012].URL: www.reactable.com/
- [49] International Organization for Standardization: ISO 9241-11. [online].[cit.25.5.]. URL:http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=16873
- [50] Jones, C.M., Dlay, S.S. (1999). The Face as an Interface: The New Paradigm for HCI. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, vol. 1, 774–779.
- [51] Just, M.A. & Carpenter, P.A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology* 8, 441–480.
- [52] Kahneman, D. (1973). *Attention and effort*. Englewood cliffs, NJ: Prentice-Hall.
- [53] Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583-1585.

- [54] Kinchla, R. A., & Wolfe, J. (1979). The order of visual processing: "Top down", "bottom up", or "middle out." *Perception and Psychophysics*, 25, 225–231.

- [55] King, P.E., Behnke, R. R. (2000). Effects of communication load, affect, and anxiety on the performance of information processing tasks. *Communication Quarterly*, vol. 48 is. 1, 74-85.

- [56] Kintsch, W., Patel, V. L., Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia* 42 (4), 186–98.

- [57] Köhler, W. (1929). *Gestalt Psychology*. New York: Liveright.

- [58] Kurtenbach, G., Sellen, A. & Buxton, W. (1993). An empirical evaluation of some articulatory and cognitive aspects of "marking menus." *Human Computer Interaction*, 8(1), 1-23.

- [59] Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *Journal of Neuroscience*, 27 (48), 13232–13240.

- [60] Lakoff, G. & Turner, M. (1989). *More than Cool Reason - A Field Guide to Poetic Metaphor*. The University of Chicago Press, 221-223.

- [61] Lakoff, G., Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press. Chicago.

- [62] Lakoff, G., Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books

- [63] Lebow D. (1993). Constructivist values for instructional systems design: Five principles toward a new mindset. *Educational Technology Research and Development*, Vol. 41, No. 3., pp. 4-16.

- [64] Lee, Ch., Federmeier, K. D. (2008). To watch, to see, and to differ: An event-related potential study of concreteness effects as a function of word class and lexical ambiguity. *Brain and Language* 104, 145–158.
- [65] Lewis, J.R. (1994). Sample sizes for usability studies: Additional considerations. *Human Factors* 36, 368-378.
- [66] Liebowitz, S., Margolis, S. (1990). The Fable of the Keys. *Journal of Law and Economics* 33 (1), 1–26
- [67] Luria, A.R. (1959) Disorders of „simultaneous perception“ in a case of bilateral occipitoparietal brain injury. *Brain* 82, 437–449.
- [68] Luria, A.R., (1987). The mind of a mnemonist: a little book about a vast memory. Cambridge: Harvard University Press.
- [69] Mansell, W., Clark, D.M., Ehlers, A. (2003). Internal versus external attention in social anxiety: An investigation using a novel paradigm. *Behaviour research and therapy* 41 (5), 555–572.
- [70] Marr, D., (1982). *Vision*. New York: Freeman and Company.
- [71] McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- [72] Miall R. C., Reckess G. Z. & Imamizu H. (2001). The cerebellum coordinates eye and hand tracking movements. *Nature Neuroscience*, Vol. 4, No. 6, 638-644.
- [73] Microsoft NUI solution [online].[cit.25.5. 2012].URL: <http://www.microsoft.com/surface/>
- [74] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63 (2), 81–97.

- [75] Mistry, P., Maes, P.(2009) SixthSense: a wearable gestural interface In ACM SIGGRAPH ASIA 2009 Sketches.
- [76] Navon, D. (1977). Forest before the trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.
- [77] Nielsen, J. & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of ACM INTERCHI'93 Conference (Amsterdam, The Netherlands, 24-29, 206-213.*
- [78] Nielsen, J. (1994). *Usability Engineering*, Academic Press Inc, 165.
- [79] Nievergelt, J. & Weidert, J. (1980). Sites, Modes and Trails: Telling the User of an Interactive System where he is, what he can do, and how to get to places. In *Methodology of Interaction*, ed. R.A. Guedj et al. 327-338.
- [80] Norman, D. (1990) *The Design of Everyday Things*, ISBN 0-465-06710-7.
- [81] Norman, D. (1999). Affordances, Conventions and Design. *Interactions* 6(3), ACM Press, 38-43.
- [82] Oberauer K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 28 (3), 411–21.
- [83] Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 3, 241-263.
- [84] Paivio, A. (1986). *Mental representations: a dual coding approach*. Oxford: Oxford University Press.
- [85] Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology*, 45, 255-287.

- [86] Parkinson, B. (1996). Emotions are social. *British Journal of Psychology*, 87, 663–83.
- [87] Pecchinenda, A., Smith, C.A. (1996). The affective significance of skin conductance activity during a difficult problem-solving task. *Cognition and Emotion* 10(5), 481–504.
- [88] Podolsky, E. (1938). *The doctor prescribes colors*. New York: National Library Press.
- [89] Pogue, D. (2011). *Mac OS X Lion: The Missing Manual*. O'Reilly Media, 176. ISBN 9781449397494.
- [90] Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. (1994). Human-Computer Interaction: Concepts And Design (Ics S.). *Science* 269, 853–855.
- [91] Prendinger, H., Mori, J., Ishizuka, M. (2005). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International Journal of Human-Computer Studies* 62, 231–245.
- [92] Ramachandran V. S. (2004). *A Brief Tour of Human Consciousness*. Pi Press, 60–82.
- [93] Ramachandran V.S., Hubbard, E.M. (2001). Synaesthesia – A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8, No. 12, p. 3–34.
- [94] Ramachandran, V.S. (2004). *A Brief Tour of Human Consciousness* Pearson Education, New York, p 73.
- [95] Rashevsky, N. (1959). Mathematical biophysics of automobile driving. *Bulletin of Mathematical Biophysics*, 21, 375–385.

- [96] Sailer, U., Eggert, T., Ditterich, J. & Straube, A., (2000). Spatial and temporal aspects of eye-hand coordination across different tasks. *Exp Brain Res* (134), 163–173
- [97] Scandura, J.M. (1971). Deterministic theorizing in structural learning: Three levels of empiricism. *Journal of Structural Learning* 3, 21–53.
- [98] Shepard, R.N., (1990). *Mind Sights: Original Visual Illusions, Ambiguities, and other Anomalies*, New York: WH Freeman and Company.
- [99] Shyns, P.G. & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5 (4), 195–200.
- [100] Sierra-Vazquez, V., Serrano-Pedraza, I., & Luna, D. (2006). The effect of spatial-frequency filtering on the visual processing of global structure. *Perception*, 35, 1583–1609.
- [101] Skinner, B. F. (1960). Pigeons in a pelican. *American Psychologist*. Vol. 15, 28-37. (Reprinted in: Skinner, B. F. (1972). *Cumulative record* (3rd ed.). New York: Appleton-Century-Crofts, 574-591.
- [102] Smilek, D., Dixon, M. J., Cudahy, C. & Merikle, P. M. (2001). Synaesthetic Photisms Influence Visual Perception. *Journal of Cognitive Neuroscience* 13:7, 930–936.
- [103] Steffens, M. C., Buchner, A., & Wender, K. F. (2003). Quite ordinary retrieval cues may determine free recall of actions. *Journal of Memory and Language*, 48, 399–415.
- [104] Steinhauer, S. R., Boller, F., Zubin, J. & Pearlman, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology*, 20, 472.
- [105] Stern, R., Ray, W.J. & Quigley, K.S. (2001). *Psychophysiological Recording*, (2nd ed.) Oxford University Press, New York.

- [106] Sternberg, R. J. (2002). Kognitivní psychologie. Praha: Portál.
- [107] Stevenson, H. (2012). Emergence: The Gestalt Approach to Change. Unleashing Executive and Orzanizational Potential.
- [108] Svebak, S. (1982). The effect of task difficulty and threat of aversive electric shock upon tonic physiological changes. *Biological Psychology* 14(1-2), 113–128.
- [109] Sweetser, E. E. (1990). From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure. *Cambridge Studies in Linguistics*, ed. B. Comrie et al. Cambridge University Press.
- [110] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12 (2), 257–285.
- [111] Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312.
- [112] Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43), 215–266
- [113] Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction* 2 (1), 59–89.
- [114] Taglialatela, J. P., Russell, J. L., Schaeffer, J. A. & Hopkins, W. D. (2008). Communicative Signaling Activates 'Broca's' Homolog in Chimpanzees. *Current Biology* 18 (5), 343–348.
- [115] The U.S. National Library of Medicine. (2010). Williams syndrome – Genetics Home Reference.

- [116] Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S.(2006). The relationship between Spatial Transformations and Iconic Gestures. *Spatial Cognition and Computation*. 6(1), 1-29.

- [117] Tversky B. (2001). Spatial schemas in depictions, In M.Gattis (Ed.), *Spatial schemas and abstract thought*, 79-111.

- [118] Types of synesthesia. (2009) [online].[cit.18.2. 2009]. URL: <http://home.comcast.net/~sean.day/html/types.htm>

- [119] Valenza, N., Murray. M.M., Ptak, R., Viulleumier, P. (2004). The space of senses: impaired crossmodal interactions in a patient with Balint syndrome after bilateral parietal damage. *Neuropsychologia* 42,1737–1748.

- [120] Virzi, R. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors* 34, 457-468.

- [121] Voorhies, D. & Scandura, J.M. (1977). "7". Determination of memory load in information processing. pp. 299–316.

- [122] Ward, R.D., Marsden, P.H. (2003). Physiological responses to different WEB page designs. *International Journal of Human-Computer Studies* 59, 199–212.

- [123] Website demonstration of Zoomable User Interface [online].[cit.25.5.2012]. URL: <http://zoomism.com/>

- [124] Wilson, Margaret (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review* 9 (4), 625–636.

- [125] Winograd, T. & Flores, F. (1986). *Understanding Computers and Cognition - A New Foundation for Design*. Ablex.