

MASARYKOVA UNIVERZITA
Fakulta sociálních studií
Katedra psychologie

DISERTAČNÍ PRÁCE

Brno 2008

Michal Vavrečka

MASARYKOVA UNIVERZITA
Fakulta sociálních studií
Katedra psychologie

Mgr. Michal Vavrečka

Kognitivní sémantika a její aplikace v modelu reprezentace prostorových vztahů

Disertační práce

Školitel: doc. Lubomír Kostroň, Ph.D.

Brno 2008

Poděkování

Za možnost realizace: Lubomíru Kostroňovi

Za pomoc při realizaci: Igoru Farkašovi

Za inspiraci: Martinu Takáčovi

Za korektury: Josefu Vavrečkovi, Mileně Vavrečkové, Pavlu Kroupovi, Mirku Krečmerovi, Janě Krečmerové, Jiřímu Helánovi, Petru Jiráskovi, Mojmíru Snopkovi, Hikmet Vavrečkové-Salihové, Tomáši Krečmerovi, Radimu Markovi, Osmanu Salihovi, Martinu Krečmerovi, Zuzaně Vavrečkové, Daniele Urbancové, Pavlu Šlosarčíkovi, Štěpánu Borkovci, Jakubovi Vavrečkovi, Lukáši Polívkovi.

Za počítač: Vojtěchu Lebduškovi

Za celoživotní podporu: rodičům

Za lásku a trpělivost: Hikmet

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně a uvedl v ní veškerou literaturu a jiné prameny, které jsem použil

1 V Brně 21.4.2008

.....
Podpis

Abstrakt

V disertační práci se zabývám možnostmi aplikace principů kognitivní sémantiky do oblasti modelování kognitivních procesů. Nejprve je provedena podrobná analýza symbolických a konceptuálních přístupů k reprezentaci významu, pomocí které jsou identifikovány nedostatky formálních sémantik a na základě rozboru problematiky ukotvení symbolů navržen systém, který dokáže uvedené obtíže překonat. Jeho činnost je založena na perceptuálních a konceptuálních přístupech ke kognici, které jsou úzce svázány s oblastí kognitivní sémantiky. V praktické části je realizován model, založený na zmíněných poznatcích, sloužící k reprezentaci statických a dynamických prostorových vztahů a pracující na principu samoorganizujících se sítí (SOM). Na základě jeho činnosti byl odhalen princip privilegovanosti jednoznačného vstupu definující novým způsob vztah mezi verbální a neverbální vrstvou reprezentace významu.

Klíčová slova: kognitivní sémantika, reprezentace významu, ukotvení symbolů, SOM, vnímání prostoru

Obsah

1	Úvod	8
1.1	Modelování jako metoda	8
1.2	Kognitivní sémantika	13
2	Základní pojmy	17
2.1	Znak	17
2.2	Symbol	19
2.2.1	Arbitrárnost symbolů	20
2.2.2	Sémiotický symbol	22
2.3	Koncept	22
3	Teoretické přístupy	25
3.1	Konceptualismus	25
3.1.1	Konceptuální empirismus	25
3.2	Analytická filosofie	27
3.2.1	Úvod	28
3.2.2	Pojmy	28
3.2.3	Reprezentace významu	29
3.2.4	Teorie pravdy	30
3.2.5	Logické operátory	32
3.2.6	Pravdivostní hodnoty	32
3.2.7	Formální symbolický systém	33
3.2.8	Kritika formálních přístupů	35
3.2.9	Intenzionální vágnost	35
3.3	Perceptuální teorie kognice	37
3.3.1	Spor o existenci symbolů	38
3.3.2	Perceptuální symbol	39
3.3.2.1	Definice perceptuálního symbolu	39
3.3.2.2	Multimodální reprezentace	40
3.3.2.3	Mentální simulátory	41
3.3.2.4	Kritika simulátorů	43
3.4	Mentální představy	43
3.4.1	Depiktivní a deskriptivní teorie	44
3.5	Duální kódování	47
4	Ukotvení symbolů	50
4.1	Obecné řešení ukotvení symbolů	53
4.1.1	Dělení přístupů	55
4.1.2	Reprezentacionismus a nonreprezentacionismus	56
4.1.3	Nulový sémantický závazek	57
4.1.4	Sociální symbol grounding	60
4.2	Proces ukotvení	61
4.2.1	Diskriminace a identifikace	64
4.2.2	Rozšířený princip ukotvení	65
4.3	Modely ukotvení symbolů	67
4.3.1	Funkcionální model	67
4.3.2	Intencionální model	68
4.3.3	Aplikovaný enaktivní model	70
4.3.4	Model přenosu ukotvení	72
4.3.5	Rozšířený model přenosu ukotvení	73
4.3.6	Konekcionistický model rozpoznávání objektů	76
4.3.7	Model ukotvení prostorových vztahů	80

5	Vnímání prostoru.....	84
5.1	Neurobiologické poznatky	84
5.2	Psychologické poznatky	86
5.2.1	Psychopatologie vnímání prostoru	86
5.2.2	Osové systémy.....	88
5.2.3	Blízkost objektů.....	89
5.2.4	Centrum deixie	90
5.3	Lingvistické poznatky.....	94
5.4	Jednotný model reprezentace prostoru.....	96
5.4.1	Geometrické vztahy	97
5.4.1.1	Symbolická reprezentace prostoru	97
5.4.1.2	Víceúrovňová reprezentace prostoru.....	99
5.4.1.3	Region connection calculus.....	101
5.4.1.4	Vizuální rutiny.....	102
5.4.1.5	Teorie prostorových vzorů	102
5.4.1.6	Model sum vektoru pozornosti.....	103
5.4.2	Extra-geometrické vztahy.....	104
5.4.3	Znalost objektu	106
6	Shrnutí	109
II.	Praktická část	111
7	Východiska pro návrh modelu.....	111
7.1	Inspirační zdroje	111
7.2	Návrh systému	113
7.3	Způsob reprezentace	114
7.3.1	Auditivní vstup	115
7.3.2	Vizuální vstup.....	116
7.3.3	Multimodální vrstva	117
7.4	Učitelský přístup	117
8	Model 1 – statické prostorové vztahy.....	119
8.1	Zadání úlohy	120
8.2	Metoda	120
8.3	Realizace.....	122
8.4	Výsledky	129
8.5	Diskuze	134
9	Model 2 - dynamické prostorové vztahy	138
9.1	Zadání úlohy	139
9.2	Metoda	140
9.3	Realizace.....	141
9.4	Výsledky	144
9.5	Diskuze	145
10	Model 3 – rostoucí síť	147
10.1	Realizace	147
10.2	Výsledek.....	149
10.3	Diskuze.....	150
11	Závěr.....	151
11.1	Diskriminace v symbolické úrovni	152
11.2	Identifikace.....	157
11.3	Kritika navrženého přístupu.....	158
11.4	Sémantika modelu ukotvení	163
12	Literatura	168

Seznam ilustrací a tabulek

Obr. 1 <i>Sémantika symbolického systému</i>	52
Obr. 2. <i>Propojení obsahu a formy</i>	63
Obr. 3 <i>Saussurovo propojení obsahu a formy</i>	63
Obr. 4 <i>Ukázka fungování systému COSPAL</i>	70
Obr. 5 <i>Vizuální vstupy Rigova modelu ukotvení</i>	74
Obr. 6 <i>Schéma Rigova modelu ukotvení</i>	75
Obr. 7 <i>Schéma Dorffnerova modelu</i>	77
Obr. 8 <i>Základní schéma Dorffnerova modelu</i>	78
Obr. 9 <i>Návrh rozšíření Dorffnerova modelu</i>	79
Obr. 10 <i>Schéma modelu ukotvení prostorových vztahů</i>	81
Obr. 11 <i>Vliv referenčních rámců na prostorové označení</i>	91
Obr. 12 <i>Absolutní referenční rámec</i>	92
Obr. 13 <i>Prostor (c) či i). (Coventry, 2004)</i>	98
Obr. 14 <i>Ukázka konvexního obalu</i>	101
Obr. 15 <i>Vizuální rutiny</i>	102
Obr. 16 <i>Prostorové vzory</i>	103
Obr. 17 <i>Model vektoru sum pozornosti</i>	104
Obr. 18 <i>Experiment referenčních rámců</i>	107
Obr. 19 <i>Návrh architektury pro ukotvení prostorových vztahů</i>	119
Obr. 20 <i>Okolí neuronu</i>	120
Obr. 21 <i>Trénovací vstupy</i>	123
Obr. 22 <i>Funkce okolí</i>	124
Obr. 23 <i>Vzdálenost sousedních neuronů dle typu mřížky</i>	125
Obr. 24 <i>Závislost rychlosti učení</i>	126
Obr. 25 <i>Trénování šumem</i>	127
Obr. 26 <i>Natrénovaná SOM mřížka</i>	129
Obr. 27 <i>Ukázka reálných vstupů</i>	131
Obr. 28 <i>Testování velikostí auditivních shluků</i>	132
Obr. 29 <i>Analýza efektivity sítě</i>	133
Obr. 30 <i>Schéma celého modelu</i>	139
Obr. 31 <i>Schéma RecSOM</i>	140
Obr. 32 <i>Ukázka vstupů nad a pod</i>	141
Obr. 33 <i>Ukázka vstupu okolo</i>	142
Obr. 34 <i>Ukázka činnosti rostoucích sítí</i>	149
Tab. 1 <i>Referenční rámce</i>	92
Tab. 2 <i>Ukázka ideálních významů a typu užití</i>	100
Tab. 3 <i>Tabulky použitých přístupů</i>	110
Tab. 4 <i>Přehled jednotlivých částí modelů a způsobů zpracování</i>	122
Tab. 5 <i>Přehled jednotlivých částí modelů a způsobů zpracování</i>	134
Tab. 6 <i>Přehled jednotlivých částí modelů a způsobů zpracování</i>	145
Tab. 7 <i>Závislost počtu přidávaných neuronů na citlivosti algoritmu učení</i>	148

1 Úvod

V disertační práci navazuji na témata, která se objevila již v diplomové práci. Zabýval jsem se možnostmi napodobení lidské inteligence a otázkami souvisejícími s daným tématem. Jako teoretický rámec mi sloužily obory konstituující kognitivní vědy, v jejichž kontextu jsem zkoumal základní architektury (konekcionismus a komputacionismus) jako nástroje pro kognitivní modelování. Cílem práce bylo nalezení hranic použití těchto architektur s přesnější identifikací možnosti jejich aplikace v oblasti napodobování lidských kognitivních schopností. Kritériem při posuzování jednotlivých systémů byla jejich psychologická plausibilita, přičemž jsem využíval také znalosti z příbuzných oblastí, např. neurověd či filosofie mysli. Výsledky práce mě vedly k závěru, že klíčovou podmínkou pro tvorbu inteligentního systému (agenta) je nutnost nalézt způsob tvorby jeho vnitřních reprezentací, což tvoří výchozí bod disertační práce. Pozornost přesouvám do oblasti kognitivní sémantiky. Jelikož má být zachována kontinuita s předchozí prací, je nutné hovořit o reprezentaci významu (sémantice) jako o způsobu zpracování informace (kognitivně).

Jelikož praktickou část disertace tvoří konkrétní model reprezentace prostorových vztahů, je první kapitola věnována zdůvodnění volby kognitivního modelování jako použité metody. V oblasti psychologických výzkumů není tento přístup příliš používán.

1.1 Modelování jako metoda

Podíváme-li se na současný vývoj v oblasti kognitivní psychologie, vidíme mladý rozvíjející se obor, ve kterém utěšeně narůstají výsledky výzkumné činnosti. I přes uspokojivá fakta nacházíme nedostatky, které nám neumožňují plně využít potenciál daného přístupu. Pokusím se provést stručnou analýzu příčin a navrhnout možná řešení. Pozornost bude soustředěna na metody výzkumu a možnosti jejich využitelnosti. Detailněji se budu zabývat otázkou kumulace a integrace výsledků výzkumů. Konkrétně otázkou, zda stávající metody umožňují propojovat výsledky do kvalitativně nových celků.

Podobné otázky nacházíme již v Brunswikově knize *The Conceptual Framework of Psychology* (Brunswik, 1952), poukazující na nedostatky v oblasti metodologie. Kognitivní psychologie zdědila mnoho metod klasické psychologie, např. pozorování, experiment, rozhovor, dotazníky. Jejich společným znakem je analytický charakter výzkumné činnosti, přičemž analytičnost spočívá ve způsobu práce ve zkoumané oblasti. V případě takto realizovaných výzkumů lze postup zobecnit do následující posloupnosti:

1. *Zvolíme oblast psychiky a teorii, ze které vycházíme.*
2. *Pomocí zvolené teorie stanovíme klíčové pojmy pro zkoumaný vztah.*
3. *Měříme empirické jevy, jejichž zpracování a analýza nás vedou k identifikaci a ověření procesů, které v dané oblasti předpokládáme.*
4. *Pomocí ověřených vztahů můžeme popisovat jevy ve zvoleném stupni obecnosti.*

Což přesně vystihuje proces analýzy, tedy rozboru zkoumaného jevu a jeho deskripce v abstraktní úrovni. Z praktického hlediska je daná metoda velmi efektivní, jelikož nám umožní abstraktně (tedy i stručně) zachytit jevy, a vytvořit o nich obecnou znalost vyjádřenou jako potvrzenou hypotézu, tzn. ověřené pravidlo. Pro elementární a stabilní jevy je metoda dostačující.

Problémy nastávají, pokud je zkoumaná doména komplexní a dynamická. Čistě analytický přístup přestává být dostačující. Množství výzkumů, produkujících abstraktní znalosti roste, ale chybí zde možnost propojovat a kombinovat tyto znalosti do kvalitativně nových celků (Brunswik, 1952). Dochází ke kumulaci výsledků, ale chybí integrace do společného rámce, přestože je tento požadavek považován v teorii vědy za klíčový. Obdobný požadavek nacházíme v Lemově knize *Summa Technologiae* (1964). Ve vědě musí dojít k nakupení takového množství informací, aby byla překonána bariéra, za kterou se znalost posune do kvalitativně nové dimenze. Samotný princip integrace není v psychologii něčím novým. O její nutnosti hovoří již průkopník kognitivní psychologie Allan Newell v článku starém 35 let (Newell, 1973) a podobné názory se objevují i v oblasti filosofie vědy (Kuhn, 1996). Podívejme se podrobněji, co může procesu integrace bránit.

V praxi dochází velmi často ke tvorbě teorií vztahujících se k identické doméně, jejichž závěry jsou protichůdné. Mezi příčiny patří (kromě nedostatku znalosti o zkoumané doméně) i použití odlišných metod, popřípadě přílišná redukce výzkumné otázky. Jelikož je většina psychologických jevů komplexní a dynamické povahy, mohou být získané fragmenty nepropojitelné díky výše zmíněným problémům. Výsledkem je množství „mikroteorií“, které nejsou schopny dostatečně podchytit komplexitu psychických procesů a množství ověřených hypotéz, jejichž závěry nelze komparovat či integrovat. V obecné rovině se o těchto problémech hovoří již zmiňovaný E. Brunswik (1952). V současnosti někteří autoři (Hiles, 2001; Polkinghorne, 1983) navrhují řešení problému pomocí užití kvalitativní metodologie, postihující zkoumané aspekty komplexněji. Tato metoda nám však opět nabízí pouhou deskripci či reinterpretaci zkoumané domény, bez možnosti zajistit integraci výzkumů do společného rámce.

Jako konkrétní ukázka metodologických nedostatků analytického přístupu může posloužit oblast psychologie vnímání, rozpracovaná v této disertační práci. Jedná se o procesy obtížně přístupné analytickým metodám. Důvodem je nedostatek informací o obecných mechanismech, které se na vnímání podílejí, tzn. nepřítomnost jednotné teorie, umožňující apriorně volit klíčové faktory. Proto se ve výsledcích experimentů týkajících se vnímání prostoru objevují protichůdné názory v mnoha oblastech. Neexistuje jednotný názor na privilegovanost osových systému (Hayward a Tarr, 1995; cf. Huttenlocher, 1991), privilegovanost lingvistické a nelingvistické informace (Landau, Hoffman, 2005), potřeby absolutních a referenčních systémů (Talmy, 1983; cf. Carlson-Radvansky & Irwin, 1993) či privilegovanosti eukleidovských vzdáleností (Mc Namara, 1986). Pokud nás navíc zajímá, jakým způsobem je interní reprezentace kódována, setkáváme se s dalšími obtížemi. Diskuse o povaze mentálních představ a formě kódování reprezentací je předmětem sporu již přes 30 let. Zastánci jednotlivých stran (Kosslyn, 2005; cf. Pylyshyn, 2003) nejsou schopni najít konsensus, jelikož zastávají odlišná východiska a snaží se potvrdit své teorie použitím odlišných metod (neurobiologické výzkumy versus obecná teorie formálních systémů). Tyto výsledky lze interpretovat jako klasický vývoj soupeřících vědeckých teorií. Na druhou stranu, pokud tento proces trvá delší dobu a jednotlivé výzkumy již nevnašejí do problematiky nové světlo, je třeba začít hledat příčiny stagnace.

Nabízí se možnost, identifikovat současné metody používané v dané oblasti jako nedostačující. Procesy vnímání jsou obecně brány jako velmi komplexní způsob mnohaúrovňového zpracování a reprezentace informace z prostředí od primitivní až po abstraktní úroveň. Dané procesy je nemožné rozčlenit na jednotlivé části a ty pak odděleně analyzovat. Navíc nemáme přístup ke specifickým procesům vnímání. Použitím klasických metod (experiment, pozorování, dotazník, rozhovor) získáváme data až z úrovně chování či jeho reflexe a nejsme schopni přesně vydělit procesy vnímání, jelikož výsledky jsou „obohaceny“ o myšlení, rozhodování apod. Z toho plyne, že používané nástroje a metody, nejsou (pro oblast vnímání) validní a výsledky nás mohou díky nepřesnosti vést k rozporuplným závěrům.

Pokud shrneme výše uvedená fakta, dostáváme se do situace, kdy je třeba nalézt takové metody výzkumu, které by nám zajistily odstranění zmíněných nedostatků. Jedná se o použití metod postavených na univerzálním principu, který umožní jednotnou interpretaci a lze je kumulovat a integrovat do společného rámce. V případě vnímání navíc potřebujeme získat možnost, zkoumat jednotlivé části systému, aniž bychom se dostali do problémů (technická realizace či etické principy související s výzkumy na živých subjektech). Možným řešením je použití metod syntetického modelování. Jejich vznik souvisí s vývojem nástrojů na testování a simulaci kognitivních procesů, tedy s rozvojem poznání v oblasti informatiky a neurověd.

Díky současné výpočetní kapacitě a znalostem funkčních propojení mozku, dochází k možnosti dostatečně komplexně napodobovat lidskou psychiku, vytvářet kognitivní systémy či jejich části, tzn. modelovat kognitivní procesy.

Pokusím se v krátkosti shrnout, v čem tato metoda překonává problémy svých předchůdců.

Její počátky souvisí s pojmem syntetická psychologie, který se poprvé objevil v názvu knihy Valentina Braitenberga (1984) „Vehicles: Experiments in Synthetic Psychology“. Oproti tradiční analytické psychologii navrhuje autor metodologický posun. Výsledkem zkoumání kognitivních procesů již nejsou pouze abstraktní deskripce na úrovni ověřování hypotéz (bod 4. z uvedeného přehledu realizace výzkumu), ale přímá aplikace poznatků ze zkoumané domény do úrovně modelu. Získané poznatky můžeme testovat v „nezávislém“ médiu. Syntetičnost spočívá v převodu poznatků z abstraktní roviny do konkrétní domény. Kromě možnosti modelovat jednotlivé části kognitivního systému, což není u živých subjektů možné, získáváme ještě další výhody.

Současné množství informací o jednotlivých výzkumných doménách je natolik rozsáhlé, že přestává být v možnostech výzkumníků, aby dokázali v dostatečné šíři obsáhnout výsledky předchozích výzkumů a na jejich základě účelně rozšiřovat výzkumnou činnost v dané oblasti. Narážíme na bariéru související s omezením kognitivních schopností samotného výzkumníka. Velikost či hloubka výzkumné domény již překročila lidské možnosti (kvantitou) a proto je třeba využít nové nástroje a metody, které nám poslouží ke zvládnutí úloh s přílišnou komplexitou či velikostí zkoumané domény. Potřebujeme nástroje, které nám umožní kumulovat a integrovat výzkum pomocí jednotné metody a do jisté míry „automatizovaně“. Vhodným nástrojem je modelování, jehož hlavní výhodou je konsistence a rozšiřitelnost zjištěných poznatků, přičemž médiem sloužícím k přenosu informací je samotná fyzická realizace modelu.

Pro začátek si ukážeme dva případy, při kterých používáme modelování pouze jako podpůrný nástroj a doplněk, který nám zpřehlední některé části výzkumu prováděné pomocí klasických metod. V případě základního výzkumu či pilotních studií slouží model zkoumané domény jako nástroj při tvorbě hypotéz či plánování výzkumných metod. Můžeme simulovat některé varianty bez nutnosti jejich realizace a získat náhled, jak na základě jejich výsledků upřesnit výzkumnou otázku, eliminovat sporné varianty a vylepšit tím výzkumný design. Jiným příkladem je možnost testování již dosažených výsledků, popřípadě jejich verifikace. V takovém případě se snažíme ověřit výsledky již dokončených výzkumů. V úrovni konkrétního modelu můžeme testovat, zda dané hypotézy či celé teorie platí i pro nerealizované alternativy.

To jsou pouze triviální příklady, při kterých používáme model jako pomůcku. Následující příklad ukazuje plný potenciál modelování, při kterém jsou překonávány problémy, související s užitím klasických psychologických metod. Jedná se o tvorbu modelu, který je postupným

rozšiřováním schopen produkovat výsledky, vedoucí ke tvorbě nových teorií, čímž se dostáváme se do kvalitativně vyšší úrovně. Oproti analytickým přístupům je model svou existencí (funkčností, či konkrétně způsobem realizace) řešením dané problematiky. Je vybudován z elementárních částí, přičemž celková architektura je díky své komplexnosti za hranicemi výzkumníkovy odhadu a proto může produkovat, dle principu emergence, kvalitativně nové výsledky. Zde překonáváme první omezení klasických psychologických metod – neschopnost zachytit komplexitu jevů. Možnost rozšiřitelnosti modelu a zvyšování funkčnosti při zachování jednoduchosti základních stavebních prvků dává metodám simulace a modelování potenciál, který u analytických metod nenacházíme.

Další výhoda spočívá ve velmi snadné ověřitelnosti teorie, kterou pomocí modelu realizujeme. Oproti riziku rozdílné interpretace výsledků u klasických metod, je v oblasti modelování verifikace přesunuta do roviny empirie. Validita a reliabilita souvisí s fungováním modelu. Pokud model neprodukuje očekávané výsledky, nemůžeme jej použít jako důkaz pro správnost našich předpokladů. Právě tyto vlastnosti činí metodu modelování velmi transparentní a odolnou vůči skrytým chybám či desinterpretacím. To v žádném případě neznamená, že je metoda odolná proti chybám způsobeným špatným zadáním výzkumné otázky, popřípadě chybnou realizací. Přesto pomocí přímé kontroly funkčnosti modelu můžeme takové chyby dříve odhalit.

Tvorbou modelů můžeme překonat i obtíže spojené s kumulativností a integrací výsledků. Elementární modely jsou postupně rozšiřovány, přičemž se stoupajícím počtem funkčních prvků stoupá schopnost realizace komplexnějších úloh. V oblasti výzkumů to znamená schopnost analyzovat model, aniž bychom měli detailní znalost o předchozích částech. Ty jsou již realizovány a jejich funkčnost ověřena. Výzkumník může testovat a analyzovat jednotlivé části i systém jako celek, přičemž je mu k dispozici nejaktuálnější fungující komplexní model. Proces rozšiřování je možné provádět stupňovitě až do úrovně, kdy je samotný model schopen efektivně fungovat na podobné kvalitativní úrovni jako originál. V praxi tomu odpovídá plnohodnotné napodobení kognitivních procesů. Kumulativnost výsledků je zaručena jednotnou architekturou, kterou pro modelování používáme. Modely samostatných kognitivních funkcí je možné integrovat do společných celků, protože jsou postaveny na totožné architektuře, čímž opět překonáváme omezení klasických metod. Kromě výše zmíněných výhod nám modelování umožňuje podrobné zkoumání oddělených psychických „modulů“, jejichž výzkum na živých bytostech je nerealizovatelný, popřípadě v rozporu s etickým kodexem.

Právě pro zmíněné výhody jsem si vybral metodu modelování jako hlavní nástroj při tvorbě praktické části své disertační práce. Během realizace modelu pro vnímání a ukotvování prostorových vztahů jsem díky nutnosti precizní formalizace jednotlivých operací získal daleko

větší vhléd do problematiky. Právě nutnost rozboru problému na elementární části, a také možnost následně vytvářet komplexní celky, činí tuto metodu efektivní již z pohledu informačního přínosu pro výzkumníka.

1.2 Kognitivní sémantika

Pokud se zajímáme o oblast kognitivní sémantiky, je třeba nejprve provést její začlenění do širšího kontextu, který je tvořen obecněji definovanou kognitivní gramatikou a lingvistikou. Kognitivní gramatiku již od konce sedmdesátých let rozvíjí americký lingvista Ronald Langacker (1987, 1991, 1999). Důvody vzniku můžeme hledat v potřebě vymezení se k tehdy převládajícímu generativistickému hnutí.

Převládajícím rysem je popis gramatiky jazyka jako „strukturovaného inventáře konvenčních symbolických jednotek“ (Langacker 1988, s. 11). Podle Langackera existuje v gramatice pouze sémantická a fonologická struktura. Význam se chápe ve smyslu experiencialistické kognitivní sémantiky. Základní jednotky jsou sémantické, fonologické, popřípadě symbolické. Gramatický systém pak obsahuje 1) gramatické jednotky, 2) významová „schémata“ těchto jednotek, 3) kategorizační vztahy mezi jednotkami. V kognitivní gramatice nejsou žádné pomocné struktury bez fonologického nebo sémantického obsahu, jejichž existence by byla nutná pouze z formálních důvodů (např. transformační pravidla).

Kognitivní gramatika podle Langackera neprodukuje věty na základě algoritmu jako gramatiky generativní, ale pouze tvoří strukturovaný systém („inventář symbolických zdrojů“), který je součástí obecného lidského kognitivního systému. Gramatika není modulárně oddělena od zbytku lidského kognitivního systému a jsou jí k dispozici stejné procesy (např. kategorizace, metaforizace, schematizace atp.), které používáme při rozhodnutí o volbě barvy tapet, ale i volbě gramatického pádu.

Kognitivní sémantiku považujeme jako oblast kognitivní lingvistiky. Výchozí principy jsou shodné, přičemž je podrobněji rozpracována část věnovaná významu. Jako reprezentační mechanismus slouží koncepty, tvořící centrální pojem přístupu. Základní východiska můžeme shrnout do tří základních bodů:

1. Gramatika je konceptualizace.

2. Konceptuální struktura je vtělená a závislá na motivovaném užití.

3. Schopnost používat jazyk je založena na obecných kognitivních schopnostech a nikoliv na specializovaném jazykovém modulu.

Kognitivní sémantika (podobně jako kognitivní lingvistika) oproti klasickým teoriím nepoužívá v oblasti analýzy jazyka dělení na syntax, sémantiku a pragmatiku. Nahlíží rozdělení syntax – sémantika jako kontinuum a nepředpokládá samostatnou lexikální složku. Všechny lingvistické vrstvy fungují společně v jednom systému a nelze je interpretovat samostatně. Oblast sémantiky se dělí na konstrukci významu a reprezentaci znalostí. Na rozdíl od analytických přístupů (Tarski, 1944), je kognitivní sémantika postavena na konceptuálním přístupu k významu. Teorie založená na pravdivostních hodnotách má totiž omezenou explanační sílu a nedokáže nám zajistit plné pochopení vět. Takový způsob reprezentace významu není v kognitivní sémantice odmítán, zároveň není brán jako absolutní a jediný. Významy jsou budovány jako koncepty procesem učení a následně propojeny se symbolickou reprezentační vrstvou.

Teorie dynamických konstruktů

Extrémním příkladem reprezentace významu je *teorie dynamických konstruktů*, ve které slova nemají žádný význam, v nejlepším případě mají „základní konstrukty“, které slouží pouze jako způsob užití slov. Což vytváří odlišný způsob přístupu ke kompozicionalitě než u analytických přístupů. Kladení důrazu na subjektivitu činí kompozicionalitu závislou na pragmatických elementech jako je kontext či intence.

Zároveň se *teorie dynamických konstruktů* vymezuje proti klasickému přístupu ke kategorizaci či prototypovému přístupu. Kategorie jsou vytvářeny přímo během procesu myšlení či promluvy a proto nemají strukturu mimo kontext svého použití. Takové užití je extrémní variantou pragmatismu. Samotná tvorba kategorií, která je prováděna okamžitě, nemusí být v kontrastu s prototypovou teorií. Pouze způsob užití kategorií je odlišný.

Kognitivní sémantika jde ve svých základech hlouběji než pojem dostatečných a nutných podmínek, se kterým se setkáme v souvislosti s teorií pravdivosti. Význam slov je chápán v kontextu naší tělesné a kulturní zkušenosti. Například výraz restaurace je brán jako série konceptů jídlo, obsluha, stolů, placení apod. Způsob jakým jsou zachyceny asociační vazby mezi těmito koncepty, nemůže být založen na analýze nutných a dostačujících podmínek.

Pravdivostní hodnoty nedokážou zachytit asymetrie ve způsobu užití slov. Například při analýze významů slov chlapec a dívka, přestože je můžeme definovat na základě pohlaví a věku, nalezneme v oblasti jazykového užití značné rozpory, které nedokážeme pomocí nutných a dostatečných podmínek vyřešit. Podobně je tomu v případě latentních významových rámců, vytvořených na základě kulturních postojů, očekávání a přesvědčení, jako součástí slovního významu. Opět se nacházíme za hranicemi možností teorie pravdivostních hodnot. Další podrobnosti o rozdílech mezi analytickými a konceptuálními přístupy jsou rozpracovány v kapitole 2.

Základní principy kognitivní sémantiky shrnuje kompendium Leonarda Talmyho (2003) prezentující obor jako celistvý přístup k jazyku z hlediska pojmových struktur. Talmyho práci lze jejím významem považovat za paralelní Langackerově knize „Foundations of cognitive grammar“ (1987), přistupující se stejnou motivací k problému vztahu významu a formy, s důrazem na reprezentaci významu. Cílem zmíněného oboru je nalezení jednotného organizačního principu, který by spojoval aspekty sémantické s formálními. Pro upřesnění uvádím jednotlivé přístupy (formální, psychologický, lingvistický a počítačový) a jejich východiska.

Formální přístupy se soustřeďují na strukturální vzorce lingvistických forem, které jsou abstraktní, tzn. nezávislé na asociovaných významech. Základem je studium syntaxe a morfologie. Klasickým příkladem může být generativní gramatika (Chomsky, 1965). Sémantika takových systémů je odlišná od kognitivní lingvistiky či psychologického pojetí významu.

Psychologické přístupy zkoumají jazyk z hlediska obecných kognitivních systémů jako je vnímání, paměť, pozornost a uvažování. Psychologie nazírá jazyk a tvorbu významů z hlediska konceptuálního. Centrem zájmu je sémantická paměť (Tulving, 1983), asociace mezi koncepty, struktura kategorií, tvorba inference apod. Je však nutné zdůraznit, že v psychologii je dosud nedostatečně zpracován integrovaný strukturální systém, jehož pomocí jazyk organizuje konceptuální obsah. Možným řešením je užití kognitivního modelování, které nám umožní lépe napodobit procesy související s reprezentací významu. Příkladem je model prezentovaný v této práci.

V kognitivní lingvistice je pozornost zaměřena na strukturu a procesy, kterými je konceptuální obsah organizován v jazyce. Jedná se o lingvistickou strukturaci základních konceptuálních kategorií (čas a prostor, scéna a událost, objekty a procesy, pohyb a umístění, dynamika a kauzace apod.). Pozornost je zaměřena na sémantickou i syntaktickou strukturu. Svou obecností zastřešuje předchozí přístupy a integruje je do společného rámce.

Z hlediska počítačového přístupu k reprezentaci významu jsme se v oblasti umělé inteligence setkávali se způsobem, který je založen na čistě formálních principech (Newell, Simon, 1972;

Pylyshyn, 1984). Jak bylo zmíněno v souvislosti se sémantikou formálních systémů, je takový způsob reprezentace redukcionistický. Soustřeďuje svůj zájem pouze na formální sémantiku pravdivostních hodnot a zcela pomíjí korespondenci symbolické roviny s externím prostředím. Řešením je systém obsahující konceptuální úroveň pro dostatečnou sémantickou reprezentaci.

Tuto možnost nabízí kognitivní sémantika, která je zaměřená na proces a tvorbu konceptů a konceptuálního systému. Při zkoumání geneze konceptů je nejprve nutné odlišit pojmy, které se v této souvislosti vyskytují a vytýčit jejich definiční oblast. Jelikož se ve formálních a konceptuálních přístupech objevuje mnoho nepřesností v definici základních pojmů *znak*, *koncept* a *symbol*, bude následující kapitola věnována právě jim. Cílem je nalézt rozdíly jejich významů a upřesnit vzájemnou hierarchii. V dalších kapitolách se budeme věnovat podrobnější analýze formálních přístupů k sémantice a jejich nedostatkům v oblasti reprezentace významu.

2 Základní pojmy

2.1 Znak

Přestože se disertační práce zaměřuje na oblast zpracování symbolů, rozhodl jsem se nejprve začít s upřesněním pojmu znak. Důvodem jsou nejasnosti související s jeho užíváním. Některé nesrovnalosti následně vedou k chybnému užití a interpretaci v oblasti kognitivních věd.

Dvě nejznámější definice znaku bývají spojeny se jmény Saussure (1965) a Peirce (1931-35). V souvislosti s prvním jmenovaným hovoříme o dyadické definici znaku. Základními elementy jsou *signifiant* a *signifié*, což z našeho pohledu můžeme interpretovat jako forma a koncept. Znak je výsledkem jejich vzájemné asociace. Dle Saussura nemohou oba elementy existovat odděleně, nelze označovat neexistující koncept a naopak. Samotnou formou je u Saussura poněkud redukcionisticky pouze zvukový vzorek označovaného. Ten je brán jako materiální entita (pokud je prezentován naším sensorům), přičemž koncept je považován za entitu abstraktní. Dnes je na formu i koncept nahlíženo jako na materiální entity, zatímco v původním saussurovském podání se hovoří o psychologických úrovních, jelikož signifiant nereferuje k externímu světu, ale ke konceptu (signifié), tzn. k mentálnímu objektu (Chandler, 2007). Je potřeba zdůraznit, že výraz mentální (psychologické) entity v Saussurově pojetí neznamena mentalistické označení, ale pouze psychologický termín pro koncept. Je realizován v mozku, a proto o něm hovoříme jako o mentálním.

Jelikož se Saussure řadí mezi strukturalisty, je pro znakový systém důležité, jakým způsobem jsou znaky propojeny mezi sebou. Externí prostředí není zohledněno, neboť význam je odvozen na základě vzájemných vztahů mezi znaky (v lingvistice gramatická pravidla, v logice pravidla vyplývání) a také vztahem znaku a konceptu (v lingvistice o nich hovoří konceptuální přístupy, v logice se pojem konceptu neobjevuje).

Pro napodobení tvorby konceptu můžeme použít i umělé systémy, které dokážou reprezentovat koncept jako abstraktní entitu referující ke třídě objektů či jevů. Přesto Saussurův přístup nezmiňuje způsob, jakým dochází ke tvorbě konceptů. Procesy učení a tvorby konceptuální úrovně jsou zcela opomenuty. Koncept je apriorně předpokládán a z toho důvodu není potřeba externího prostředí, jelikož konceptuální úroveň jej plně nahrazuje. Pokud nás ovšem zajímá způsob reprezentace významu, potřebujeme definovat konceptuální úroveň přesněji, než jako pouhý apriorní předpoklad. Proto rozšíříme pojetí znaku o způsob propojení s externím prostředím.

S takovým přístupem se setkáme u Peirce (1931-35). Jedná se o zástupce pragmatismu (na rozdíl od strukturalisty Saussura), proto v jeho přístupu můžeme vidět posun k subjektivitě, kontextualismu a pragmatickým principům při reprezentaci významu.

V základní úrovni je znak definován triadicky. Nacházíme obdobu *signifiant* nazývané *representamen*, tedy forma, kterou znak nabývá. *Signifié* je *interpretant*, což je smysl (význam) vytvořený znakem. Jak jsme se již zmínili, většinou pod tímto pojmem chápeme koncept. Novým prvkem, proti Saussureovi je *objekt*. Jedná se o entitu v externím prostředí, ke které *representamen* skrz *interpretant* referuje. Právě rozšíření o referenci činí tuto teorii použitelnou v oblasti kognitivních věd. Klade důraz na nutnost propojení inteligentního systému (schopného manipulovat se znaky) s externím prostředím. Tvorba konceptu (v Peircově pojetí *interpretant*) je zásadním požadavkem při vývoji umělých systémů. Až do současnosti byly používány přístupy, které konceptuální úroveň zcela pomíjejí (logika), popřípadě jí vytvářejí pouze na základě arbitrárních symbolů (např. sémantické sítě). Z toho plynou nedostatky při reprezentaci významu, jelikož nepřítomnost či neadekvátní forma konceptu omezují schopnosti systému. Proto je převod *objektu* do úrovně *interpretantu* hlavním tématem disertační práce. Souvisí s problematikou ukotvení symbolů (viz níže).

Dalším posunem oproti Saussurovu přístupu je zdůraznění znaku jako relativního a kontextově závislého. Jedná se o příklon k subjektivismu, tedy zdůraznění individuálního konceptuálního systému jedince a z něho odvozeného způsobu užití znakového systému. Nutnost reprezentovat *interpretant* jako mentální obsah se objevuje již v základní definici znaku.

A sign is something which stands to somebody for something in some respect or capacity.

Znak, je něco, co pro někoho něco zastupuje z nějakého hlediska nebo v nějaké úloze.

(Peirce, 1931-36, 2.228)

Důraz na subjekt (někoho), který je interpretuje vztah (v jeho hlavě existuje *koncept* - *interpretant*) činí definici použitelnou jak z hlediska psychologie tak oblasti reprezentace a modelování významu v rámci kognitivních věd. Obecně lze shrnout, že pro oblast kognitivní sémantiky je bližší přístup CH. S. Peirce, jelikož v sobě zahrnuje propojení konceptu s externím prostředím.

Jak již bylo zmíněno, zajímají nás z hlediska kognitivní sémantiky také nelingvistické (sub-symbolické) principy odhalující způsob tvorby konceptů na základě vstupů z prostředí. Samotný pojem znaku je příliš této oblasti příliš obecný a neumožní nám detailněji studovat konceptuální úroveň. Znak je definován jako soubor více komponent (forma, obsah a popří-

padě objekt, ke kterému referuje), což je pro naše účely nedostačující a proto budeme muset analyzovat jednotlivé úrovně samostatně.

Shrnutí: Již v začátku jsme zjistili, že jednotlivé teorie se od sebe liší počtem použitých úrovní znaku. Peircova obsahuje externí prostředí, interní koncept a symbolické označení. Saussure považuje za dostačující konceptuální úroveň a následné symbolické označení. Analytické přístupy pak využívají pouze symbolickou rovinu. Právě tyto rozdíly se ukážou v pozdějším výkladu jako klíčové a považuji za nutné je zmínit již v počátku. Jelikož se v minulém odstavci objevilo místo pojmu *signifiant* či *representamen* pojem symbol, přikročíme nyní k analýze právě tohoto pojmu.

2.2 Symbol

S definicí symbolu se setkáváme již v pracích zmíněných autorů. Saussure (1965) nejprve používal symbol jako nadřazený pojem znaku, následně ale dochází k redefinici symbolu jako specifického typu znaku. Totožnou hierarchii zastával i Peirce (1931-35), přestože se jejich pojetí symbolu lišilo. Saussure se vyhýbal označení znaku jako symbolu. Problematickou se mu jevila arbitrárnost symbolů, které v případě lingvistiky nebyly totožné se symboly v matematice. Jelikož matematika nereferuje k externímu prostředí, může používat symboly zcela arbitrárně, pouze na základě relací mezi nimi (podobně jako logika), což si v oblasti lingvistiky nemůžeme dovolit (Langer, 1951). Důvodem je zmiňovaná vazba na koncept.

Peirce přichází s odlišnou definicí, podle které jsou symboly zcela arbitrární. Vytváří propracovanou formu typologie znaku, ve které se mohou vyskytovat kromě arbitrárních symbolů také ikony a indexy. Podle Peirce je možné rozlišit tři typy znaků z hlediska úrovně denotace (*ikon*, *index* a *symbol*). *Ikon* na sebe bere podobu zobrazovaného objektu (socha či fotografie konkrétní osobnosti je *ikonem* své předlohy). *Index* je k objektu, k němuž odkazuje, vztažen nějakou věcnou souvislostí (kouř je indexem ohně, kýchnutí je indexem nachlazení či alergie). *Symbol* je s objektem či konceptem, k němuž odkazuje, vztažen konvencí, společenskou dohodou. Vztah mezi označovaným a označujícím je nahodilý a arbitrární. Zároveň nejsou typy znaků vzájemně výlučné, jelikož znak může být *symbol*, *ikon* nebo *index*, popřípadě jejich kombinace. Již podle Peirce je obtížné nalézt instance čistého *indexu* či *ikonu* (Peirce 1931-58, 2.306). Podle Hawkesa koexistují jednotlivé typy znaků ve vzájemné hierarchii, přičemž dominance nad ostatními je dána v závislosti na kontextu (Hawkes 1977, 129).

Díky Peircově typologii dostáváme více způsobu, jakým lze reprezentovat znak v úrovni *representamen*. Forma může mít podobu vnímatelného kauzálního spojení (*index*), totožné ča-

soprostorové uspořádání (*ikon*) nebo zcela arbitrární symbolické označení přiřazené na základě asociace. ***Symbol* můžeme definovat jako typ znaku (jeho *representamen*), který je *interpretantu* přidělen zcela arbitrárně.**

Je důležité zmínit, že Peirce (podobně jako Saussure) ve svých textech často zaměňuje pojem znak za *representamen* (*signifiant*), přestože, jak již bylo zmíněno, samotná forma (*representamen*) nemůže existovat. Proto je nutné odlišovat znak od formálního označení (*representamen*), jelikož znak v sobě obsahuje formu i obsah, zatímco *representamen* je pouze jednou jeho složkou. Podobné záměny mohou vést k chybným interpretacím.

Na druhou stranu Peirce oproti Saussureovi zdůrazňuje možnosti nelingvistických znaků. Ikonickou reprezentaci považoval za dostatečně primitivní, aby byla „originálním znakem“, tedy elementární formou. *Symbols* se podle něj vyvinuly z *ikonu* procesem zjednodušování, ve kterém lze nalézt i přechod od analogového k diskrétnímu (digitálnímu).

2.2.1 Arbitrárnost symbolů

Zaměřme se ještě podrobněji na pojem arbitrárnosti. Zajímá nás symbol jako entita arbitrární a diskrétní, referující k objektu nebo stavu světa (Dorffner, 1993). Přestože se nám známá slova nezdají arbitrární, je tomu naopak. V souvislosti se sebereferenčností symbolické roviny se objevuje pojem „hermeneutická zrcadlová hala“ (Harnad, 1990), který vystihuje arbitrárnost této úrovně, tzn. definování jednoho symbolu pomocí jiného.

Nutnost konceptuální úrovně

Objevují se také názory (Dorffner, 1993), že lidé mají schopnost manipulovat arbitrárními znaky, bez toho aby věděli, co znamenají (bez nutnosti konceptuální úrovně). Příkladem je věta „Love ever keeps trying“ - v myšlence vytáhneme každé druhé písmeno a dostaneme slovo „over“. Jedná se ale o zavádějící příklad. Proces provádíme na základě mentální představy (vytvářené z konceptuálně uložené reprezentace) a nemůžeme tedy říci, že manipulujeme symboly. Jelikož jsou znakové systémy nerozlišitelně propojeny (Peircova triáda), nemůžeme říci, že systém pracuje pouze na základě jedné vrstvy.

Je to právě arbitrárnost, kterou Saussure považuje za výhodu a nazývá jí prvním principem jazyka (Saussure 1983). Díky arbitrárnosti je možné vybudovat systém jazyka do efektivní a univerzální podoby. Striktně odmítá nativismus, jelikož v kontextu přirozeného jazyka neexistuje žádné vrozené, esenciální, jasné a evidentní spojení mezi *signifiant* a *signifié* - mezi zvukem (tvarem písma) a konceptem, ke kterému referuje (Saussure 1983). Libovolný symbol je možné přiřadit libovolnému konceptu. Daný princip se objevuje již u Aristotela, který tvrdí, že zvuk libovolného označení jazyka může být přiřazen libovolnému objektu či konceptu (Richards 1932). Platón pak myšlenku rozvíjí s poukazem na to, že každé jméno přiřazené objektu je správné. Pokud změníme jméno na jiné, je stejně správné jako to předchozí (Harris, 1987).

Přesto je náhodné přiřazení možné pouze při vytváření systému, nikoliv, pokud jsou strukturní vazby již vytvořeny. Jestliže budeme nahodile přiřazovat označení, můžeme užitím nového označení pro stávající koncept (vymažeme propojení na staré označení) tyto vazby porušit. Tvzení o arbitrárnosti lze definovat volně pouze pro nově budované systémy. Je možné vytvářet systém označení, které jsou ve formální úrovni přiřazovány nahodile či arbitrárně. Není však možné vytvářet náhodná označení do systému, který již má vytvořeny relační a strukturní vztahy. Pokud změníme (například zaměníme) dvě označení, je stávající jazykový systém narušen, jelikož staré vztahy (vytvořené relačně a kontextuálně) přestávají platit. Proto je nutné princip arbitrárnosti limitovat pouze na genezi nových označení, při zachování struktury a konceptuálního ukotvení významu již existujících. V praktické rovině je náhodnost symbolů v úrovni jazyka pouze zdánlivá, jelikož většina symbolických označení je ke konceptům přiřazena na základě konvence, podmíněné dlouhodobým vývojem jazyka.

Na druhou stranu je princip arbitrárnosti výhodný v procesu vývoje libovolného znakového systému. Efektivita symbolů je zřejmá v oblastech, kde ikonická forma ztrácí svou potenci (a naopak). Jelikož jsou symboly arbitrární a nejsou podobné (tvarově) s objekty v reálném světě, mohou od nich abstrahovat a být používány jako zástupce kategorií, které v sobě integrují ikonicky odlišné objekty. Příkladem je neschopnost ikonické reprezentace abstraktních slov jako láska, pohoda, ale i jednodušších kategorií např. luštěnin nebo primátů. Ve zmíněných případech arbitrárnost zaručuje, že dokážeme vyjádřit třídu, která je ikonicky nekonzistentní, popřípadě hierarchicky abstrahovat od shodných znaků (například postupujeme-li od ryb k živočichům, přecházíme od základních shod ploutve, šupiny, hlava, tělo, přes ploutve, hlava tělo až do vyšší úrovně živočichů, kde platí pouze shoda hlava, tělo). Což dává symbolům potenciál v abstraktní rovině a činí je použitelné v procesech kategorizace, ale i oblastech jako matematika či logika.

2.2.2 Sémiotický symbol

Přestože jsme v těchto dvou kapitolách podrobně rozebrali pojmy znak a symbol, velmi málo jsme se dozvěděli o významu a sémantice znakových systémů. Samotná reprezentace významu založená na teoriích kognitivní lingvistiky a sémantiky je spojená s procesem semiose.

Jedná se o interakce mezi formou (*representamen*), významem (*interpretant*) a objektem (*referent*). Význam je závislý na tom, jak a s jakou funkcí je konstruován *semiotický symbol*, tzn. symbol tvořený procesem semiose. Samotná sémiotika se věnuje studiu procesů společenské produkce významů ze znakových systémů (kódů). Jedná se o teoretický přístup, popř. metodu analýzy, která se soustřeďuje na tři základní oblasti: analýzu významů znaků (sémantika), analýzu pravidel pro kombinování znaků (sémiotický syntax) a analýzu vztahu mezi znaky/kódy a jejich uživateli (pragmatika). Semioze je brána jako trilaterální vztah mezi konstituenty (*representamen*, *interpretant*, *objekt*), který nelze redukovat na vztah mezi jednotlivými dvojicemi. (Peirce, 1931-35)

V konečném důsledku nás však ani samotná semiose nepřiblíží k základům reprezentace významu. Definuje význam znaku (či semiotického symbolu) na základě významu reprezentovaného v *interpretantu* a nevnaší do problematiky reprezentace významu nové poznatky. Z hlediska analýzy musíme postoupit o úroveň níže a ptát se, jakým způsobem je vytvářen koncept (*interpretant*), jestliže ten je označován jako nositel významu.

2.3 Koncept

Podle obecné definice je koncept abstraktní myšlenkou, popřípadě získanou dispozicí rozpoznávat objekty, které mají společné vlastnosti. Bývá někdy ztotožňován s procesem kategorizace, tzn. schopností diskriminovat objekty či jevy na základě jejich podobnosti a následně vytvářet třídy, které sdružují objekty a jevy podle společných vlastností. Koncepty (někdy nazývané reprezentace) abstrahují mezi rozdíly uvnitř třídy.

Často se setkáváme se zaměňováním pojmu symbol a koncept a proto uvádím přesnější rozlišení. V předchozí kapitole jsme definovali symbol z hlediska Peircovy triadické teorie, tzn. rozlišení *representamen*, *interpretant* a *objekt*. Jestliže symbol je arbitrární označení, konceptem jsou kategoriální mentální stavy, tvořené na základě adaptivního kategorizačního procesu.

Rozdíl si můžeme ukázat na příkladu pojmu židle a rozlišením mezi jeho konceptem a symbolickým záznamem podle kritérií arbitrárnosti, reference a diskrétnosti. Slovo židle je arbitrár-

ní, ale koncept židle nikoliv, jelikož je svázán s našimi mentálními stavy a tělesnou zkušeností. Slovo židle referuje ke třídě židlí, ale kategorie židlí nereferuje k ničemu, je kauzálně spojena se stimulem či situací. Je potřeba externího pozorovatele, abychom tento vztah mezi stavy světa a stavy mysli zachytili. Slovo židle je diskrétní, jelikož se objevuje odděleně v akustické i psané formě. Koncept je také určitým způsobem diskrétní, ale konekcionisté většinou tvrdí (Dorffner, 1993), že právě jeho nediskrétnost je podstatná (židle podobná křeslu). Pravděpodobně žádná diskrétní entita nebo akce nekonstituje koncept. Konceptuální funkce emerguje z komplexní konfigurace mechanismů ve světě i mozku (Barsalou, 2003).

V oblasti kognitivní lingvistiky se setkáváme s pojmem abstraktních konceptů, které vznikají transformacemi z konceptů konkrétních. Jsou derivovány ze zkušenosti, např. pomocí strukturního mapování, ve kterém jsou vlastnosti ze dvou a více zdrojů selektivně mapovány do *smíšených prostor* (Turner, Fauconnier, 2002). Výchozím principem je teorie konceptuálního (pojmového) míšení (autoři této teorie v angličtině většinou hovoří o *conceptual blending*, tj. pojmovém míšení, ale používá se také termín *conceptual integration*).

Tento proces umožňuje zpracování vět jako „*Ve Francii by aféra Watergate Nixonovi neškodila*“. Samotná věta je složitě integrovaným pojmem. Dva prostory se zkombinují a vytvoří prostor třetí (integrovaný), který zdědí část struktury původních prostorů, ale zároveň vytváří strukturu vlastní. Integrované prostory splňují tyto čtyři podmínky:

- 1) *Částečná projekce mezi prostory (například prezidenti, volební systémy atd. jsou vzájemně propojeny)*
- 2) *Obecný prostor, který se částečně promítá do obou prostorů, v jehož rámci můžeme na jejich integraci nahlížet (například systém západní demokracie)*
- 3) *Integrace, v níž se oba zdrojové prostory částečně promítnou na nový prostor (například prostor USA promítne Nixona, jeho osobní charakteristiky a chování, ale ne to, že je americkým prezidentem, a prostor Francie promítne to, co víme (nebo se dozvídáme) o volebních preferencích francouzských voličů)*
- 4) *Nová struktura integrovaného prostoru, která nepochází ani z jednoho ze zdrojových prostorů (například Nixon je francouzským prezidentem, nemusí rezignovat a neztratí prestiž) (Fauconnier 1997)*

Teorie *pojmového míšení* slouží jako inspirační zdroj pro návrh modelu, realizovaného v disertační práci. Některé poznatky z oblasti abstraktních konceptů a způsobu jejich integrace, byly využity při řešení ukotvení symbolů pomocí multimodálních reprezentací (viz níže).

Podle Fregeho (1892) a stoupenců analytického přístupu dělíme reprezentaci významu na smysl a referenci (Sinn a Bedeutung nebo Sense a Reference). Právě smysl (není v analytické filosofii podrobněji rozpracován) popisuje stavy světa a způsoby, jakými je objekt prezentován. Proto je možné smysl vnímat jako identický s konceptem. V minulé kapitole jsem uvedl, že koncept není v logice používán, přesto jej zde zmiňuji. Rozdíl spočívá v užití pojmu koncept ve formálních systémech (kalkulech). Koncepty v kalkulech nejsou založeny na percepci. Jsou užívány pouze abstraktní koncepty (např. sčítání, derivace), které se nezakládají na časových a prostorových vlastnostech a jsou zcela autonomní, abstraktní, univerzální a nezávislé na kontextu.

V kapitole věnované analytickým přístupům se podíváme na rozdíl mezi pojmáním konceptů u konceptuálního a analytického přístupu. Již nyní je nutné zdůraznit, že právě pojmání smyslu jako symbolického popisu stavu světa nás dovede k problematice ukotvení symbolů (Harnad, 1990), jelikož způsob tvorby konceptu je založen na subsymbolických mechanismech. ***Zaměňování konceptu se symbolickým popisem vede k jeho chybné reprezentaci v oblasti umělé inteligence a k neschopnosti takových systémů „porozumět“ svým (konceptuálním) reprezentacím.***

Často bývá také chybně užíváno pojetí konceptu jako mentální a abstraktní entity, díky tendenci přiřazovat jim neadekvátní vlastnosti. Pojem mentální pouze lokalizuje, v jakém médiu jsou reprezentovány (neuronová síť našeho mozku). Pojem abstraktní hovoří o procesech kategorizace. ***Při tvorbě konceptu jsou abstrahovány konstantní znaky, které určují náležitost objektu do určité kategorie. To neznamena, že pouhý abstraktní (symbolický) popis těchto znaků je dostačující způsob reprezentace. V elementární úrovni je konceptuální úroveň vytvářena na základě percepce, a proto jsou některé vlastnosti nesymbolické a nelingvistické.*** Podívejme se podrobněji na filosofické pojmání procesů kategorizace a tvorby konceptů. V první fázi nás zajímá, zda lze všechny koncepty získat empiricky (perceptuálně) nebo potřebujeme apriorní zkušenost.

3 Teoretické přístupy

3.1 Konceptualismus

Jak bylo uvedeno dříve, nepředpokládáme u racionalistických (analytických) přístupů konceptuální úroveň ve formě perceptuálně získaných kategorií. Racionalisté zdůrazňují vrozenost konceptů sloužících procesům myšlení. Jak si řekneme později v souvislosti s formálním symbolickým systémem, nejedná se o konceptuální úroveň v pravém slova smyslu. Naproti tomu empiristé považují konceptuální úroveň jako perceptuální. Tento rozpor mezi racionalisty a empiristy se objevuje již od období Aristotela a Platóna. Okolo 17. a 18. století dochází k opětovnému vymezování jednotlivých přístupů. Racionalismus je spojován se jmény Descartes a Leibnitz a empirismus s britskou školou Locka a Huma.

V počáteční fázi je oblast kognitivních věd ztotožnitelná s racionalistickým přístupem. S nástupem konekcionismu začíná převládat příklon k empirismu a perceptuálním teoriím kognice. Oproti racionalistům je na koncept nahlíženo odlišně:

- 1. Koncepty jsou založeny na percepci.**
- 2. Koncepty jsou naučené.**
- 3. Koncepty referují ke kategoriím objektů ve světě skrze spolehlivé kauzální vztahy.**
- 4. Koncepty jsou kontextuální.**

Pokud považujeme koncepty za naučené a perceptuálně získané, nevyhneme se otázce, jakým způsobem je můžeme apriorně vytvářet a jaké schéma vnímání musíme aplikovat, chceme-li je z perceptuálních vstupů získat bez předchozí zkušenosti.

3.1.1 Konceptuální empirismus

S problémem týkajícím se schémat použitých při vnímání pro následnou kategorizaci se setkáváme již v oblasti filosofie mysli. Dotýká se ještě specifičtějších témat než problematika ukotvení symbolů a nazývá se konceptuální (konceptový) empirismus. Jedná se o aplikaci klasického empirismu (Locke, 1690) do oblasti tvorby konceptů, které jsou tvořeny na základě interakce s prostředím. Problematická je čistě empirická tvorba konceptů, jelikož potřebujeme apriorní schémata pro jejich získání.

Tento fakt vedl Kanta (cca. 1781) k odmítnutí teorie konceptového empirismu. Nelze získat koncepty pouze ze zkušenosti, protože mít zkušenost znamená použití konceptů (např. koncepty časoprostorových vztahů nebo koncept příčinnosti). Proto musejí být koncepty vrozené. Chomsky (1965) došel ke stejnému zjištění v oblasti učení se jazyka, přestože se detaily jeho teorie později změnily. Již Hume (1739) tvrdil, že veškerá znalost je buď analytická (pravdivá definicí a esenciálně prázdná) nebo empirická (vyžaduje experiment a pozorování pro ověření.). Příklady apriorních schémat nacházíme v oblastech, které nejsou přímo spojeny s percepcí (abstraktní principy matematiky a logiky) nebo v mechanismech, které kategorizaci umožňují. Kant tvrdil, že existují, zvláště v matematických znalostech, příklady syntetické povahy znalosti. Jedná se o neempirickou (apriorní) tvorbu znalostí, která je imunní empirickým námitkám. Teorie postavené na čistě empirických základech nazýval Kant *noogonie*. Podle něj nelze získat koncepty pouze ze zkušenosti, protože mít zkušenost znamená použití konceptů. Rozděluje tedy koncepty na apriorní a aposteriorní, přičemž apriorní umožňují vznik aposteriorních, které souvisí s empirickým způsobem geneze konceptů.

Jestliže je způsob tvorby konceptů součástí řešení problematiky ukotvení symbolů, musíme vzít tyto připomínky v potaz, přestože při jejich analýze narážíme na metodologické problémy. Pokud budeme předpokládat vrozené mechanismy, dostáváme se do oblasti nativistických teorií, u kterých je velmi obtížné zjistit, zda jsou použité mechanismy vrozené nebo se vyvíjejí až po narození. Kritika vyčítá nativismu obtížnou falzifikovatelnost.

Pokud se přesuneme do úrovně architektury systémů, nejsou apriorní koncepty obtížně realizovatelné. V úrovni živých systémů nemusíme při řešení problematiky apriorních konceptů využívat nativistické teorie. Mozek je složitou neuronovou sítí, která dokáže procesy kategorizace primárně provádět a vzniklé koncepty velmi komplexně reprezentovat. Apriorní koncepty můžeme v rovině živých systémů ztotožnit s architekturou mozku a jejím funkčním propojením. Kritika empirismu také často zapomíná vzít v potaz „bohatost“ struktury prostředí. Informace, které přijímáme, jsou natolik komplexní, že umožňují vytvořit koncepty a jejich hierarchie. Ze základních konceptů je následně možné odvodit koncepty vysoce abstraktní. Systém musí obsahovat dostatečnou senzorickou výbavu pro příjem informací z prostředí a schopnost reprezentovat koncepty, jejich hierarchie a vzájemné vazby.

V rovině umělých systémů můžeme problematiku apriorních konceptů řešit obdobně. Pomocí neuronových sítí založených na principu samoorganizace (Kohonen, 1989), můžeme vytvořit koncepty, které zachycují vztahy bez předchozí znalosti. Na základě dostatečné senzorické výbavy, schopné zachytit informace z prostředí a vhodného kategorizačního a reprezentačního mechanismu (samoorganizace), vytvoříme systém s empiricky získanou konceptuální úrovní. Princip samoorganizace umožní struktuře vnitřní reprezentace přizpůsobení se struktuře vy-

skytující se v externím prostředí. *Pokud vycházíme z předpokladu, že okolní prostředí nám manifestuje své vlastnosti, ze kterých postupnou kategorizací můžeme dospět až k velmi abstraktním znalostem (konceptům), považujeme princip samoorganizace jako odpověď na kritiku konceptového empirismu a potvrzení východisek perceptuální teorie kognice a kognitivní sémantiky.*

Někteří autoři (Sloman, 2007) argumentují proti konceptovému empirismu s poukazem na fakt, že některá zvířata jsou po svém narození schopná následovat matku, nalézt potravu apod. Pokud mají zvířata takové apriorní koncepty, není možné připustit tvrzení konceptového empirismu o získávání všech znalostí ze zkušenosti. Argument poukazuje na přílišnou striktnost konceptového empirismu, který považuje *veškeré* koncepty jako empiricky odvozené.

Rozeberme tento argument podrobněji. Přestože je úroveň reflexů a pudů apriorní a organismu vlastní od narození, jsou reflexivní a pudové reakce natolik predefinované a neměnné, že je obtížné je začlenit do oblasti konceptů. Jedná se o autonomní a uzavřené mechanismy, které je možné začleňovat do komplexních struktur jedině jako specifickou komponentu, která funguje deterministicky. Bližší poznatky o dané úrovni souvisí s adekvátní teorií emocí, která dosud nebyla prezentována v oblasti psychologie, natož kognitivních věd.

Závěrem lze říci, že problematika apriorních konceptů není v úrovni realizace obtížně řešitelná. Přestože kladu ve svém přístupu důraz na nutnost empirických poznatků při formování konceptů, používám apriorní koncepty ve formě samoorganizujících se sítí, čímž se odchyluji od striktní definice empirismu (která ve své čisté podobě není technicky realizovatelná). Z hlediska probíraných argumentů bych svůj přístup zahrnul do oblasti konceptualismu, s důrazem na empirickou zkušenost.

3.2 Analytická filosofie

Jestliže je základním principem racionalismu tvorba znalostí pomocí abstraktních procedur, přičemž pravdivost není založena na senzorické zkušenosti, ale na základě dedukce, potřebujeme rozřešit otázku týkající se povahy a způsobu reprezentace těchto abstraktních procedur, které používáme k reprezentaci a transformaci znalostí.

3.2.1 Úvod

Již v Saussurově teorii znaku (1965) nacházíme obdobné názory, jako u zastánců analytické filosofie (Frege, 1892; Tarski, 1944). Podle něj jsou znaky smysluplné pouze jako součást formálního, abstraktního systému. Na podobných principech vystavěl Frege a jeho následovníci logické formální systémy. Saussure klade důraz na strukturální a relační vztahy. Reference s externím prostředím pro něj není primární. Konceptuální úroveň dostačuje jako nositel významu, protože obsahuje podstatné vlastnosti objektů a jevů v externím prostředí. Jakým způsobem korespondence s externím prostředím dosáhneme, Saussure neupřesňuje, jelikož se soustřeďuje pouze na strukturální vztahy. „*Význam znaků spočívá v jejich systematické souvztáhnosti (znaku ke znaku), nikoliv v jejich derivaci z neměnných vlastností formy (signifiant) či reference k materiální věci*“ (Saussure 1983, s. 121. Definice se velmi blíží základním předpokladům formálních logik, ve kterých je význam definován v úrovni symbolů (arbitrární *representamen*). Saussurovo pojetí je odlišné v definici znaku, složeného z formy a konceptu, zatímco klasické logiky konceptuální vrstvu neobsahují nebo definují symbolicky.

3.2.2 Pojmy

Jestliže byl centrálním pojmem předchozích přístupů *znak*, přičemž jsme se soustředili na konceptuální úroveň, jako způsob reprezentace významu, v analytické filosofii a logice hovoříme v této souvislosti o *pojmu*. Pojmy jsou definovány jako objektivní abstraktní (ideální) entity, které mohou být reprezentovány výrazy jazyka (Materna, 1995). Odklonem od definice znaku (pojmu) v závislosti na objektu, ke kterému referuje a konceptu, který jej reprezentuje, se dostáváme k odlišnému způsobu reprezentace znalostí a významu.

Zástupci analytického přístupu (Materna, 1995) kritizují empirické a konceptualistické přístupy, pro způsob reprezentace konceptu. Chybně jej přirovnávají k (mentálním) představám místo k (mentálním) reprezentacím. Hlavním argumentem proti užití představ pro reprezentaci pojmů souvisí s neschopností lokalizovat abstraktní objekty v prostoru a čase, na rozdíl od představ, které jsou lokalizovatelné v čase a prostoru a mají konkrétní podobu. Pomocí představ nedokážeme reprezentovat obecninu, což vede logiky k odsouzení psychologických způsobů reprezentace významu. Nepřípustný je také fakt, že mentální představa je založena na smyslovém vjemu, což je interpretováno jako příklon k empirismu. Poslední obtíž spatřují stoupenci analytické filosofie (Materna, 1995) v subjektivním charakteru představ.

Zkusme tyto námitky interpretovat z hlediska konceptuálního přístupu v psychologii. Nejprve je třeba zdůraznit, že pojmy jsou tvořeny mentálními reprezentacemi, nikoliv mentálními představami. Představy jsou pouze konkrétní manifestace možností, které potencionálně obsahují mentální reprezentace. Samotné reprezentace nejsou pouze konkrétní a lokalizovatelné časoprostorově, jelikož obsahují vlastnosti, které mohou dosahovat vysokého stupně abstrakce. Subjektivita reprezentací je logickým důsledkem způsobu jejich akvizice. Subjektivita je základní předpoklad pro vznik intencionality a významu obecně. Pokud postulujeme pouze objektivní svět, není potřeba jeho interpretace či abstrakce vlastností, jelikož neexistují systémy, které by byly takových operací schopny. ***Abstraktní rovina mimo čas a prostor nemůže existovat, pokud neexistují uživatelé takového systému. Bez subjektu neexistuje abstraktní objekt.***

Námitky týkající se subjektivních významů a neschopnosti shody (Wittgensteinova námitka privátního jazyka) nejsou v souladu s našimi intuicemi a jedná se o extrémní variantu objektivismu. Jestliže sdílíme společné externí prostředí, jsou reprezentace, které z tohoto prostředí odvozujeme subjektivní, ale struktura prostředí je natolik stabilní, že zaručuje možnost intersubjektivní shody.

Dostáváme se k poslední námitce týkající se smyslové povahy představ/reprezentací. Při použití vhodné architektury, nám není znemožněno abstrahovat informace z prostředí do libovolného stupně obecnosti. Senzorický příjem informací navíc umožňuje kontakt s empirickou rovinou a tedy možnost verifikovat význam a pravdivost na základě interakce s prostředím či využitím senzoričky získaných reprezentací. Takové možnosti systémy založené na objektivních pojmech neumožní. Vycházejí z odlišné teorie pravdy a způsob práce s reprezentací významu je také odlišný. Podrobnosti o rozdílech jsou popsány v následujících kapitolách.

3.2.3 Reprezentace významu

Jestliže jsou významy objektivní a ideální abstraktní entity, máme k nim (jako subjektivní bytosti propojené s okolním světem skrze senzoričku výbavu) omezený přístup. Sémantika formálních systémů je odlišná od konceptualismu, jelikož způsob definice pojmu v takových teoriích není deskriptivní, ale normativní.

Jediným systémem, kde lze obdobné principy beze zbytku aplikovat, je matematika. Je založena na číslech, což jsou entity, které splňují podmínku primární abstraktnosti, tzn. nelze je identifikovat v čase a prostoru. Pokud se snažíme teorii aplikovat na veškeré entity, setkáváme se s problémy, jakým způsobem nalézt jejich objektivní a ideální význam.

Logikové vytýkají konceptualismu, že je založen na Kantově principu fixace podstatných vlastností, což koresponduje s procesy kategorizace. Argumentují příklady z matematiky, kdy pojmy jako mocnina či prvočíslo, nejsou entitami, které by fixovaly podstatné znaky, jelikož odkazují pouze k jediné funkci či entitě. Tyto pojmy jsou však matematické, jedná se tedy o primárně abstraktní entity. Pokud nás ale zajímají také empirické výrazy, potřebujeme při tvorbě pojmů použít princip zmíněný Kantem.

3.2.4 Teorie pravdy

Základní rozdíly v přístupu k sémantice nacházíme již v oblasti teorie pravdy. Ta určuje, jakým způsobem je v daném přístupu vyjadřována pravdivost o objektech a jevech. V logice byla zpočátku upřednostňována *koherenční teorie pravdy*. Takový způsob se objevuje u průkopníka moderní logiky Gottloba Frega. Základem je soubor symbolů, pomocí kterých formalizujeme jazyk jako uzavřený a konsistentní systém. Koherenční teorie se zabývá pouze syntaktickou stránkou výpovědí. Systém dokáže produkovat výroky dle formálních pravidel vyplývání a přiřazuje jim pravdivost. Pokud v systému výroků neexistuje rozpor, můžeme je označit za pravdivé. Právě konzistentnost a uzavřenost nám zaručí, že se v systému neobjeví výrok, který by byl pravdivý i nepravdivý.

Prozatím se soustředíme na odlišnosti oproti předchozím systémům. Logický systém založený na koherenci není propojen s externím prostředím. Má význam pro logickou syntaxi, ale stěží může sloužit k poznávání vnějšího světa, jelikož opomíjí sémantickou i pragmatickou rovinu. Manipuluje pouze s abstraktními entitami, které referují k externímu prostředí, ale ***propojení s empirickou rovinou je ve zmíněných systémech možné pouze pomocí externího, sémanticky znalého pozorovatele***. Znaky (pojmy) obsahují pouze symbolickou úroveň, konceptuální úroveň není zastoupena. Formální sémantika platí jen „uvnitř systému“ a vyžaduje jeho bezrozpornost a úplnost, což může být v aplikované oblasti kognitivních věd nedostačující.

Nová koncepce teorie pravdy se objevuje v souvislosti s logikou u Tarského (1944). Vychází z *teorie korespondence*, podle které výrazy jazyka referují k objektům v externím světě. Vztah je však pouze jednosměrný. Změny v externím světě znamenají změny v popisu (symbolické abstrakci). Významy mají svou kauzální příčinnost v externím světě. Podle logiků význam souvisí s jazykovými výroky a jejich pravdivostí. V Tarského pracích se objevuje kritika koncepce pravdivostní hodnoty přímo v jazyce. Postuluje metajazyk, jehož pomocí je možno pravdivost vyjádřit. Metajazyk je postaven na *T-schématech* propojujících jazyk a metajazyk (aktuální stav věcí ve světě). Přestože se nabízí závěr, že metajazyk zastupuje konceptuální

úroveň, která propojuje symbolickou rovinu s externím prostředím, nelze jej přijmout, jelikož pravdivostní hodnota definovaná v libovolném symbolickém systému nezajišťuje propojení s rovinou empirie. V oblasti vtělené kognitivní vědy je propojení zajištěno skrze agentovu senzorickou výbavu, ale Tarského přístup propojení s empirickou rovinou pomocí percepce nepředpokládá, přestože je to bezpodmínečně nutné. Jím definovaná sémantika platí pouze pro formální jazyky, protože přirozený jazyk je příliš komplexní a obsahuje mnoho výjimek z pravidel. Podle zastánců intenzionálních logik (Tichý, 1996) navíc není v možnosti žádného formálního systému odvodit *aktuální stav* ve světě. U klasických formálních systémů je proto definice významu redukcionistická a nedostačující pro oblast kognitivních věd.

Rozšíření formální sémantiky na přirozený jazyk provedl Davidson (1984), který chápe pravdivost jako primitivum. Jím postulovaná pravdivost je tautologická (s touto definicí pravdivosti se setkáváme ve všech formálních systémech), protože přisuzuje význam větě její pravdivostní hodnotě. Kromě toho používá jako atom příliš komplexní strukturu (větu), jejíž pravdivostní hodnota je kompozicí částí a neodhaluje nic o významu jednotlivých slov.

Můžeme shrnout, že nedostatkem formální logiky a sémantiky je jejich základ v korespondenční teorii pravdy. V okamžiku, kdy se změní externí svět (např. v úrovni změny prostředí vnímaného agentem), musí to být reflektováno i v symbolické rovině, pokud má být pravdivostní hodnota zachována. To by ale znamenalo schopnost formálního systému reflektovat změny v doméně, do které nemá žádný přístup, protože je jejím abstraktním a *arbitrárním* popisem. ***Je nemožné postulovat agenta jako „všemohoucí bytost“, která je schopna ze svých univerzálně platných axiomů inferovat v každém okamžiku aktuální stav světa. Takový agent navíc empirickou rovinu nepotřebuje.***

I pokud bychom použili *epistémickou teorii pravdy*, která je založena na verifikaci symbolických výrazů v rovině empirie, opět nejsme schopni tento proces zajistit pouze pomocí formálního systému. Verifikaci musí provést externí pozorovatel, který je obeznámen se světem na subsymbolické úrovni a který je schopen reflektovat změny na nearbitrární úrovni, založené na znalosti *prostoru* a *tvaru*, jejíž reprezentace je bez adekvátní konceptuální úrovně nemožná. Problém tedy nespočívá v korespondenční teorii pravdy, ale ve formálním symbolickém systému, který tuto korespondenci nedokáže zajistit, díky nepřítomnosti konceptuální úrovně. Ukazuje se, že mezi důležité aspekty při zjednání korespondence mezi externím prostředím a reprezentacemi (např. ve formě peircovského znaku) patří schopnost učení se a také schopnost kategorizovat informace z prostředí a vytvářet konceptuální úroveň, kterou u formálních systémů postrádáme.

3.2.5 Logické operátory

Podívejme se ještě podrobněji na některé vlastnosti formálních systémů založených na koherenční a korespondenční teorii pravdy. Základními pojmy, které tvoří jádro formálních systémů, jsou logické operátory a pravdivostní hodnoty. Přestože Tarski definoval sémantiku na základě pravidel vyvozování, nenacházíme přesnější definici samotných logických operátorů. V systému jsou obsaženy apriorně.

U predikátové logiky nejsou výrokové spojky (operátory) v rámci vlastní interpretace spojeny s pravdivostními funkcemi, neboť je interpretován celý kontext, nikoliv pouze logická spojka. Samotný symbol spojky je prohlášen za *nevlastní* (Materna, 2000). Kdyby interpretace v predikátovém kalkulu byla postavena na funkcionálním přístupu, byly by logické spojky i kvantifikátory brány jako sémanticky svébytné výrazy a interpretace složených výrazů (včetně atomických vět) by byla odvozena na základě operace aplikace funkce na argument. Z toho vyplývá, že logické operátory jsou v extenzionálních logikách brány pouze jako *nevlastní symboly*. S rozvojem intenzionálního přístupu (Tichý, 1996) byly prohlášeny za funkce, které přiřazují pravdivostní hodnotu.

3.2.6 Pravdivostní hodnoty

Pokud hovoříme o pravdivostní hodnotě, velmi zřídka si uvědomujeme, že ji nelze aplikovat univerzálně na všechny typy výroků, jak předpokládali logikové. V prvé řadě je nutné rozdělit výroky, kterým přisuzujeme pravdivostní hodnotu, na analytické a empirické, přičemž mezi analytické patří i výroky matematické, které platí za všech podmínek a můžeme je vyjádřit pomocí extenze, tzn. rozsahu pojmu (nepotřebujeme použít intenzi, která je smyslem pojmu). V extenzionálně pojaté logice se rozlišení neprovádí a při její aplikaci do oblasti přirozeného jazyka vznikají obtíže.

Příkladem je ztotožňování významu označení přímo s objektem objevující se u Fregeho, v případě analýzy výrazu Jitřenka a Večernice. Abychom správně identifikovali objekt Venuše (o kterém oba výrazy referují), potřebujeme vytvořit intenzi (nikoliv pouze extenzi jak tvrdí Frege), abychom dokázali odlišit Jitřenku od Večernice. Dostaneme dvě propozice, jejichž pravdivostní hodnoty přiřadí pouze astronom na základě pozorování a nikoliv logik, díky nepřístupnosti aktuálního světa. Jelikož se jedná o empirická, nikoliv analytická tvrzení, nelze problém v logické úrovni vyřešit (Materna, 2000). Frege přiřazuje objektům rovnou pravdivostní hodnoty, čímž nerespektuje jejich časovou a prostorovou proměnlivost.

Při srovnání pravdivostních hodnot matematických a empirických docházíme k závěru, že u matematických je pravdivostní hodnota nutná, nezávislá na empirických faktech a věčná. Naproti tomu empirická věta nemůže označovat pravdivostní hodnotu, ale pouze *propozici*, tedy funkci nabývající jednu pravdivostní hodnotu (Materna, 2000). Pokud potřebujeme určit pravdivostní hodnotu propozice, musíme použít externího, sémanticky vybaveného pozorovatele, popřípadě rozšířit formální systém o konceptuální úroveň. Druhou možnost budeme probírat v souvislosti s realizovaným modelem.

Zatím se přesněji podíváme na vztah logiky a empirie. Jak již bylo řečeno, analytické přístupy od empirické roviny abstrahují, jelikož ji považují za nepřesný zdroj informací o světě a raději pracují s ideálními a abstraktními entitami. V případě přirozeného jazyka se ovšem bez empirických výrazů a výroků neobejdeme. Přitom žádné tvrzení logiky není empirické, závislé pravdivostní hodnotou na stavu světa.

Při užití logiky v oblasti empirických výroků a snaze zjistit či odvodit aktuální stav světa, narážíme na paradox *vševědoucnosti*. Z hlediska formálního systému je nemožné vědět, který svět je aktuální, jelikož bychom museli znát všechna fakta a být vševědoucí. Proto nutně docházíme k závěru, že to, co závisí na empirických faktech, nemůžeme znát apriorně, a nelze tedy zjistit aktuální hodnotu. Ke zjištění pravdivostní hodnoty nutně potřebujeme mimologické prostředky. V opačném případě bychom verifikovali empirické věty pouhou logickou analýzou (Materna, 2000).

Proto je potřeba mechanismů, které převedou vstupy z externího prostředí do úrovně interní reprezentace. V případě, že systém obsahuje reprezentace, které jsou v korespondenci s aktuálním světem, můžeme verifikovat pravdivost empirických výrazů na základě jejich korespondence s interní reprezentací (*interpretant* neboli koncept) či aktuálním prostředím vnímaným skrze senzorickou výbavu (objekt), popřípadě verifikovat empirické výroky formou přiřazení pravdivostních hodnot propozicím, na základě znalostí z konceptuální úrovně.

3.2.7 Formální symbolický systém

Několikrát byl v souvislosti s analytickými přístupy zmíněn formální symbolický systém, jako jádro, které slouží ke tvorbě inferencí, založených na syntaktických pravidlech. Podívejme se tedy blíže na jeho základní principy a způsoby jakým pracuje s reprezentací významu. Definici vystihující fungování formálního systému, nacházíme u Harnada (1990). Podle něj je formální symbolický systém:

1. Soubor arbitrárních „fyzických znaků“ zaznamenaných na papíře, pásce či jiném médiu, se kterými je
2. manipulováno na základě „explicitních pravidel“, která jsou
3. podobná znakům, či řetězcům znaků. Pravidly řízená symbolická manipulace je zaměřena
4. pouze na tvar symbolických znaků (ne jejich význam), tzn., je čistě syntaktická a obsahuje
5. „pravidly řízené kombinace“ a rekombinace symbolických znaků. Existují
6. primitivní atomické symbolové znaky a
7. kombinované (složené) řetězce. Celý systém a všechny jeho části - atomické znaky, složené znaky, syntaktické manipulace, aktuální i možná pravidla jsou
8. „sémanticky interpretovatelné“: Syntax může být systematicky přiřazován významu, např. zastupovat objekty, jevy atp.

Harnad dále dodává „Sémantická interpretovatelnost, aby byla symbolická, musí být propojena s explicitní reprezentací (2), syntaktickou manipulovatelností (4) a systematicitou (8). Žádné z těchto kritérií není arbitrární, a pokud je zlehčíme, ztratíme možnost pracovat s pojmy jako přírodní kategorie a také možnost operovat na základě formální teorie výpočtu, čímž se stává ze slova „symbolický“ nevysvětlená metafora (která se liší interindividuálně)“ (Harnad, 1990).

Autor poukazuje na hlavní výhody symbolického systému a dodává, že neexistuje alternativa k formálnímu systému, kterou bychom mohli použít jako podklad pro myšlenkové operace. Přesto si ukážeme, že formální systém a jeho teorie významu není univerzálně použitelná pro napodobování kognitivního systému. Problémy jsou zřejmé již z definice. V bodě 4. se hovoří, že pravidly řízená manipulace je zaměřena pouze na tvar, nikoliv význam symbolického znaku. V potaz není brána konceptuální úroveň (kterou může sémanticky interpretovat konceptu znalý pozorovatel), sémantika je redukována na soustavu univerzálních pravidel, které sdílejí význam na základě principů inference. Systém manipuluje pouze v syntaktické rovině, udržuje konsistentnost svých výroků, ale bez sémanticky vybaveného pozorovatele jsou jím prováděné operace pouhou mechanickou procedurou zachovávání pravdivosti. Systém konceptuální úroveň během činnosti nebere v potaz.

Ukazuje se, že pouhá symbolická úroveň a syntaktická (logická) pravidla, pomocí kterých se symboly manipulujeme, se jeví jako nedostatečná pro popis či napodobení lidských kognitivních schopností. Již v oblasti kognitivní psychologie se setkáváme s kritikou formálních systému, právě pro jejich důraz na syntaktičnost. Argumentace vychází z psychologických poznatků o fungování kognitivního aparátu. Lidské usuzování (odvozování, inference) se odehrává na intrapsychickém sémantickém základě. Sémantická inference člověka není založena

na principu operací se symboly. Pro každou entitu reálného světa, které věnujeme pozornost, si ve vědomí tvoříme odpovídající přirozený model. Víme, že dotyčnou entitu můžeme chápat z nesčetné řady hledisek a s různou hloubkou a šíří, s rozpoznáním a uvažováním různých souvislostí a vztahů (souhrnně řekneme kontextů).

3.2.8 Kritika formálních přístupů

Také v oblasti umělé inteligence se setkáváme se způsobem reprezentace znalostí, který je založen na čistě formálních principech. Jak bylo zmíněno v souvislosti s formální sémantikou, je takový způsob reprezentace redukcionistický. Přesto bývají čistě symbolické přístupy používány v reprezentačních systémech založených na seznamech, rámcích, schématech či sémantických sítích (Minsky, 1974; Quillian, 1966). Na takto definovaných reprezentacích lze vytvořit plně funkční symbolický systém s kombinatorickou syntaxí a sémantikou, který podporuje vyšší kognitivní funkce, jako jazyk, paměť, znalosti. Zastánci daného přístupu jsou např. Newell a Simon (1972), Fodor (1975), Pylyshyn (1984) apod. Soustřeďují svůj zájem pouze na formální sémantiku pravdivostních hodnot a zcela pomíjejí způsob propojení symbolické roviny s rovinou empirickou.

Problematické je, že amodální teorie (není založena na znalostech z jednotlivých senzoric-
kých modalit, ale je abstraktní – symbolická) dokáže vysvětlit modální specificitu, protože dokáže v principu vysvětlit cokoli. Jak bývá často zmiňováno, amodální teorie mají sílu Turingova stroje, dokážou tedy vyjádřit libovolný popsatelný vzor. Síla napodobit každý empirický výsledek je ale na úkor její nefalzifikovatelnosti. Protože takové teorie vysvětlují cokoli a žádný výsledek je nemůže vyvrátit (Barsalou, 2003).

3.2.9 Intenzionální vágnost

Velmi neotřelý pohled na rozdíl mezi formálními systémy a lidskými kognitivními schopnostmi přináší Křemen (2006). Z filosofického hlediska hodnotí možnosti formálních systémů jako nástrojů napodobení mysli. Základním pojmem, od kterého odvíjí svou teorii, je intenzionální vágnost. Naše vnitřní reprezentace, které vytváříme na základě informací, jsou intrapsychické modely, které jsou subjektivní a nepřesně vymezené. Právě toto vedlo logiky k odklonu od empirismu a psychologismu, jelikož není možné dospět k abstraktním a ideálním reprezentacím. Podle Křemena jsou naše mentální reprezentace zatíženy intenzionální

vágností. Důvodem je způsob akvizice znalostí, jelikož z empirických dat nedokážeme odvodit obecný abstraktní význam. Také subjektivní způsob tvorby konceptů přispívá ke zvýšení vágnosti. Na druhou stranu formální jazyky, které využíváme v logice, mají minimální či nulovou intenzionální vágnost. Jelikož naše reprezentace obsahují nekomunikovatelné a sub-symbolické aspekty, nedokážeme nalézt transformaci, která by umožnila odstranit vágnost a pracovat s takovými reprezentacemi ve formálních systémech. Proto je při tvorbě formálních reprezentací nutné od těchto vlastností abstrahovat. Pokud však chceme pomocí modelu, který je založen na takovém způsobu reprezentací, aby napodobil lidské kognitivní schopnosti, nedosáhne této úrovně.

Na rozdíl od zástupců analytických přístupů vidí autor ve vágnosti základní předpoklad při tvorbě inteligentních systémů. Vágnost je prostředkem samopohybu, který umožňuje procesy myšlení a tvořivosti. Formální inference s nulovou intenzionální vágností tento princip postrádají a potřebují externí řízení. Proto autor nepředpokládá, že by mohly formální systémy myslet. Pro napodobování inteligence je potřeba systémů, které kopírují biologické neuronové sítě a umožňují reprezentovat znalost s dostatečnou vágností (Křemen, 2006). Autor předpokládá kvalitativní rozdíl mezi činností vědomí a fungování počítače. Vědomí pracuje s volnými (nefixovanými) modely takřka libovolně vágními, zatímco počítač pracuje s modelem fixovaným. Fixace spočívá v pevné jazykové reprezentaci modelu. Nemáme žádné jazykové prostředky pro popis nefixovaných modelů. Ty zůstávají tajemstvím principu činnosti lidského vědomí. Rovněž nedovedeme konstruktivně popsat zrod nového nápadu (Křemen, 2006).

Rozdíl mezi logickým a kognitivním pojmáním významu můžeme demonstrovat i na Dennetově dělení typů postojů (1987), které zaujímáme k externímu prostředí. Entitám ve světě se snažíme porozumět tak, že k nim zaujímáme různé druhy postojů. Nejjednodušší jsou fyzikální postoje, kdy entitu interpretujeme prostřednictvím kauzálních zákonů, definovaných přírodními vědami. V tomto bodě se neliší logické a kognitivní přístupy, jelikož vysvětlovací principy jsou objektivní, společné všem objektům. Další typy postojů (projektový a intencionální) leží mimo rámec objektivistické teorie logiky. Předpokládají subjekt přistupující k určitým entitám světa na základě znalosti, uložené v jeho mysli. Postoje jsou obohaceny o víry, přání či intence, které je činí neanalyzovatelné z hlediska logického.

Postoje obsahující subjektivitu jsou pro objektivní přístup k významům velmi obtížně interpretovatelné. Problematika intencionálních postojů bývá řešena pomocí speciálních konstrukcí v intenzionální logice, přesto však není možné definovat veškeré možnosti. Narážíme také na nepřístupnost intrapsychických procesů subjektu. Psychické procesy typu *věřit*, *chtít*, *spoléhát* apod. v sobě obsahují emoční složky, které jsou formalizaci pro svou neverbální povahu ještě méně přístupné. Hranice užití formální sémantiky tedy nedovoluje, abychom pomocí

formální teorie významu vyvozovali závěry o entitách, které jsou nám nepřístupné, nebo u kterých neznáme transformaci do úrovně formální symbolické reprezentace.

Donald Davidson (2001) upozorňuje, že zatímco porozumění přírodě směřuje jednoznačně k deterministickým kauzálním zákonům, porozumění našim bližním směřuje ke stanovení mnohem méně deterministických principů, na základě kterých víme, co je od koho rozumné očekávat. Proto také potřebujeme jiný druh pojmů, než jaké jsou potřeba pro vědecké zvládnutí přírody.

Takovým pojmem je například pravda. Ta má podobně fundamentální povahu, jako pojem čas, tzn. všichni mu dobře rozumí, ale není možné ho nějak definovat či vysvětlit, protože už schází pojmy, se kterými bychom se dokázali dostat „pod něj“ (Peregrin, 2007). Dostáváme se až k otázkám, jakým způsobem je reprezentována pravda a pravdivost a ve které doméně (konceptuální či symbolické) bychom ji měli hledat. Díky obtížím souvisejícím se symbolickými reprezentacemi se podíváme na možnost reprezentace významu pomocí perceptuálních a konceptuálních teorií.

3.3 Perceptuální teorie kognice

Perceptuální teorie kognice má za sebou dlouhou historii výzkumné činnosti, neboť do 20. století byly teorie tvorby znalostí z velké části perceptuální. V kognitivní lingvistice se perceptuální charakter znalosti objevuje u mnoha badatelů (Fauconnier, 1985, 1997; Jackendoff, 1987; Johnson, 1987; Lakoff, 1987, 1988; Lakoff & Johnson, 1980; Lakoff & Turner, 1989; Langacker, 1987, 1991, 1997; Talmy, 1988). V psychologii patří mezi zastánce Paivio (1986), Miller, Johnson-Laird (1976), Huttenlocher (1973, 1976), J. Mandler (1992), Glenberg (1997), Goldstone (Goldstone, 1994;), Wu (1995), MacWhinney (1998), (Barsalou, 1993; Barsalou, Yeh, Luka, Olseth, Mix, & Wu, 1993; Barsalou & Prinz, 1997). Ve filosofii Nersessian (1992), Peacocke (1992), Thagard (1992), Clark (1997) a Prinz (1997). V oblasti umělé inteligence (Glasgow 1993; Bailey, Feldman, Narayanan, & Lakoff, 1997; Rosenstein & Cohen, 1998).

Změna nastala s rozvojem logiky, statistiky a programovacích jazyků, tedy amodálních (formálních) teorií, které jsou postavené na jiných než perceptuálních základech. Perceptuální hledisko bylo popíráno, protože jako základ byl považován „nahrávací systém“. Jelikož se nejedná o běžný pojem, bude nutné jeho vyjasnění. Systémy zpracování reprezentací můžeme dělit v základě na „nahrávací“ (zaznamenávací) systémy a konceptuální systémy. Perceptuálně založené teorie bývají většinou nazírány jako „nahrávací“ systémy. Nahrávací systémy

vytvářejí záznamy fyzických informací, které nejsou zcela přesné, přibližně jako fotografie, videokazety apod. Takový nahrávací systém informace nezpracovává, pouze je přijímá a ukládá (Barsalou, 1999). Pokud budeme perceptuální teorie reprezentace znalostí vidět jako nahrávací/zaznamenávací systémy, nebudou nikdy plausibilní či konkurenceschopné. Takové budou jedině, pokud získají vlastnosti konceptuálních systémů. Nazírání perceptuální teorie jako „nahrávací“ místo konceptuální znamenal nepochopení jejich výchozích principů.

Zkoumání kognice je v oblasti způsobu reprezentace značně rozpolcené. Podle jedněch zastánců perceptuální systém shromažďuje vjemy z okolí a odesílá je odlišnému systému, který s nimi pracuje na úrovni jazyka, paměti a procesů myšlení. Naopak podle Barsaloua (1999) je kognice zcela perceptuální a není založena na kognitivních reprezentacích neperceptuálního charakteru, které nazýváme amodální.

Pokud se podíváme pozorněji na perceptuální teorie, zjistíme, že jdou hlouběji než zjednodušující stereotypy. Mnoho filosofů věřilo, že reprezentace jsou komponentové a užívané produktivně (např. Locke, Russell, Price), že existují nevědomé reprezentace ve formě „dispozic“ a „schémat“, které následně produkují vědomé představy (např. Locke, Kant, Reid, Price), že představy mohou odrážet nesenzorickou zkušenost (např. Locke, Hume, Kant, Reid), že přirozené mechanismy interpretují a organizují představy (Kant, Reid), a že představy mohou být dynamické (Barsalou, 1999). Proto je cílem autorů zabývajících se perceptuální teorií kognice (Barsalou, 1999; Prinz, 2003; Pecher, Zwaan, 2005), vybudovat teorii perceptuálního symbolického systému, který v sobě bude obsahovat výhody amodálních (formálních) systémů, ale bude zohledňovat kognitivní a neurální aspekty.

3.3.1 Spor o existenci symbolů

V souvislosti s perceptuálními teoriemi nacházíme zcela opačné názory na reprezentaci znalostí, než tomu bylo u formálního přístupu. Zpochybňují samotnou existenci symbolu jako centrálního reprezentačního mechanismu. Obecně se dá říci, že neexistuje přímý důkaz o existenci amodálních symbolů. Pouze instrumentálně jsou zachytitelné v konceptuálních systémech, u kterých předpokládáme existenci amodálních symbolů. Nejvíce problémů s užitím amodálních symbolů je v reprezentaci časoprostorových znalostí (Barsalou, 1999).

Objevují se však opačné názory (Pylyshyn, 2003), podle kterých zastánci perceptuálních teorií podcenili obtížnost nalezení důkazů proti amodálnímu přístupu ke konceptům a vyšší kognici. Je obtížné určit rozdíl mezi amodálními a perceptuálními reprezentacemi na základě rozlišování analogových (podle povahy perceptuálních vstupů) a symbolických reprezentací. Přesto-

že nacházíme důkazy analogových reprezentací ve vizuálním systému, nelze prokázat, že perceptuální reprezentace jsou analogové. Naopak, mnoho perceptuálních reprezentací může být lingvistické povahy (Pylyshyn, 2003). Takový argument je ale nedostačující. Z hlediska teorie jsou oba přístupy postavené na jiných principech, jelikož obsahují jinou definici sémantiky. Proto jejich vzájemné porovnávání není možné.

3.3.2 Perceptuální symbol

Primárním cílem v oblasti perceptuálních teorií je nalezení způsobu převodu perceptu do konceptuální úrovně a následné přiřazení symbolického označení konceptu. Tento převod není triviální a jednotlivé teorie často redukuje část procesu, která je v rozporu s jejich principy.

Reprezentační přístupy zdůrazňující symbolickou vrstvu, vidí kognitivní a perceptuální reprezentace jako oddělené systémy, které pracují na odlišných principech. Po vytvoření perceptuálního stavu dochází k jeho *transdukci* do symbolického, amodálního systému, kódu, který je zcela neperceptuální. (Barsalou, 1999). Po tomto převodu jsou amodální symboly používány v reprezentačních systémech založených na seznamech, rámcích, schématech, sémantických sítích či produkčních systémech. Tím je konstituován plně funkční symbolický systém s kombinatorickou syntaxí a sémantikou, který podporuje vyšší kognitivní funkce, jako jazyk, paměť a znalosti. Zastánci obdobných principů jsou např. Newell a Simon (1972), Fodor (1975), Pylyshyn (1984) apod. Je jasné, že takové systémy jsou amodální a arbitrární. Amodálnost je způsobena nulovou korespondencí interní struktury reprezentace s perceptuálními stavy, které je produkují. Způsob převodu znehodnotí některé informace a způsob uložení neumožňuje jejich přímé použití při opětovné percepci (Barsalou, 1999).

Abychom daný převod lépe reprezentovali, potřebujeme najít jiný způsob propojení perceptů s koncepty a následně s jazykovou úrovní či s úrovní chování. Základní intuice nás směřují k závěrům, že jazyk musí být propojen se světem stejně jako při dětském vývoji (Roy, 2005a).

3.3.2.1 Definice perceptuálního symbolu

Z výše uvedených důvodů je třeba vytvořit alternativu k symbolickému přístupu. Jestliže je symbolický systém založen na arbitrarnosti a nepřítomnost konceptuální úrovně, je možné postulovat symbol, který by měl odlišné vlastnosti a více se přiblížil klasickému pojetí znaku podle Peirce, ve kterém je symbol (*ikon* či *index*) propojen s konceptem (*interpretant*) a refe-

ruje k objektu reálného světa. Jedna z teorií postulující nearbitrární symboly, se objevuje v článku Lawrence Barsaloua „Perceptual Symbol System“ z roku 1999. Přichází s dělením symbolů na modální a amodální a členěním reprezentací na modální a multimodální. Navíc postuluje pojmy perceptuální symbol a perceptuální simulátor. Pokusme se rekapitulovat jeho teorii.

Na rozdíl od klasických symbolů formálních systémů jsou perceptuální symboly modální a analogové. Modálnost zaručuje, že jsou informace zpracovávány i reprezentovány identickým systémem, tzn. probíhá v jedné modalitě. Například neurální systém pro rozpoznávání barvy je odpovědný také za její reprezentaci, je tedy používán pro zpracování i vybavení (Barsalou, 1999). Perceptuální symboly jsou neurální reprezentace v senzomotorických oblastech mozku. Autor klade důraz na senzorické modalitty, které mohou vést k tvorbě ukotvených reprezentací.

Perceptuální symboly jsou dynamické, nikoliv diskrétní. Pokud je perceptuální symbol uložen, nefunguje rigidně jako diskrétní symbol. Protože je perceptuální symbol asociovaný vzorec neuronů, jeho následná aktivace má dynamické vlastnosti. Nabízí se nám škála možných modifikací. Jestliže je vnímán perceptuální symbol, který je asociován ve stejné oblasti, může změnit vzorec neuronů. Jiný kontext může modifikovat originální vzorec. Perceptuální symbol lze nazírat jako atraktor v neuronových sítích, který se v průběhu změn také modifikuje (Barsalou, 1999).

Proces formace symbolů neplatí jenom pro vidění, ale i pro ostatní senzorické modalitty. V každé modalitě vytvoří selektivní pozornost záznam o jejich činnosti, který je uložen v dlouhodobé paměti, a který poté slouží jako symbol. Tak vzniká pestrá škála symbolů různým modalit. (Barsalou, 1999).

3.3.2.2 Multimodální reprezentace

Dalším pojmem, který se v souvislosti s teorií perceptuálního symbolického systému objevuje, jsou multimodální reprezentace. Samotný pojem multimodální znamená, že jsou rozprostřené ve specifických senzorických modalitách. Reprezentace jsou multimodální, protože vznikají na základě senzorických modalit, propriorecepce a introrecepce. Barsalou (1999) hovoří o specifických oblastech, ale nezmiňuje se o možnosti jejich distribuovaného uložení v jedné oblasti kombinující více modalit dohromady. Právě tato modifikovaná varianta multimodálních reprezentací tvoří jádro modelů realizovaných v praktické části disertace.

Integrace modalit se objevuje již u Damasia (1989), který hovoří o zónách konvergence integrujících informace ze senzomotorických map, čímž dochází k jejich reprezentaci. Konvergenční zóny pomocí hierarchických množin asociačních oblastí integrují specifické modalit do společného rámce. Damasio (1989) dále tvrdí, že není rozdíl mezi perceptuální a kognitivní pamětí. Hovoří o gradientu z posteriorní do anteriorních asociačních oblastí, jež vyjadřuje míru komplexity a specifity paměťových obsahů vzhledem k síle aktivace senzomotorických oblastí. Paměťové oblasti jak uvnitř, tak vně senzomotorického systému se podílí na tvorbě kognitivních reprezentací (Barsalou, 1999).

3.3.2.3 Mentální simulátory

Dalším pojmem, se kterým se v perceptuální teorii setkáváme, je mentální simulátor. Jak bylo zmíněno výše, nemůže nám symbolický systém v rovině sémantiky nabídnout dostačující způsob reprezentace. Proto je třeba vytvořit model, schopný perceptuální vlastnosti původních objektů zachytit a reprezentovat s ohledem na jejich časové a prostorové vlastnosti a zajistit tak verifikovatelnost jejich významu, tzn. propojit empirickou a abstraktní (např. symbolickou) rovinu.

Je-li základním elementem perceptuální symbol, potřebujeme také mechanismy, jak tyto jednotky propojovat do celků a provádět s nimi operace podobné inferenci v oblasti logiky.

V amodálních (formálních) teoriích je způsobem reprezentace rámec či sémantická síť, se kterými provádíme logické operace. V teorii perceptuálního symbolu tuto úlohu přebírá mentální simulátor. Jeho základní struktura obsahuje dvě úrovně: 1) Podkladový rámec integrující perceptuální symboly dané kategorie a 2) nekonečnou množinu simulací, které mohou být vytvořeny z daného rámce. Tyto dvě úrovně slouží jako základ pro tvorbu konceptuálních funkcí (Barsalou, 1999).

Právě druhá vrstva odlišuje simulátor od rámce. Z hlediska klasické umělé inteligence se jedná o spojení rámce a inferenčního mechanismu. Systém by produkoval a poté ohodnocoval možné kombinace řešení, což by odpovídalo činnosti simulátoru. Přesto bude operovat pouze na symbolické úrovni, z čehož plynou již zmíněná omezení. Naopak činnost simulátoru, jak jej navrhuje Barsalou, je založena na operaci s perceptuálními symboly, které překonávají možnosti klasických rámců. Systémy jsou pak nazývány konceptuálními.

Během své činnosti využívají simulátory základních vlastností formálních symbolických systémů, tzn. mohou být používány kombinatoricky a rekurzivně, což zajišťuje produktivitu.

Jelikož jsou perceptuální symboly ukotvené v sensomotorickém systému, vykazují také prvky vtělenosti (Barsalou, 1999). Navíc v případě komplexních simulací fyzických a introspektivních událostí reprezentují perceptuální symboly abstraktní koncepty.

Mentální simulátory produkují simulace, které jsou jen částečné a schematické, nikdy kompletní, jelikož selektivní pozornost tvoří z percepce perceptuální symboly, ale neextrahuje veškeré dostupné informace. Právě tento převod se podílí na výsledné síle a možnostech simulátorů. Simulace jsou zkreslené i pomocí vnitřních procesů, což je také činí odlišným proti originálu (jak potvrzují gestalt principy a zákony). Simulátory nejsou pouze sbírkou empirických dat ze smyslů. Mozek, jako mechanismus pracující na principu mnoha genetických omezení, ovlivňuje práci simulátorů. Jedná se o propojení empirie a racionality, přičemž racionalita je právě ta část, která je silně ovlivněna geneticky (Barsalou, 1999). Přestože nám intuice říká, že evoluční vývoj našeho mozku je ovlivněn principy efektivity (darwinismus), nelze tato tvrzení přímo ztotožnit se závěry, že náš mozek funguje racionálně. V závěru uvádím přehled základních vlastností simulátorů, které je odlišují od ostatních reprezentačních mechanismů.

1. *Simulace nejsou v žádné úrovni kompletní (podobně hovoří Kosslyn o mentálních představách).*
2. *Simulace nejsou přesné (obsahují chyby apod.).*
3. *Nejsme si vědomi všech simulací, které současně probíhají.*
4. *Nejen lidé mohou provádět simulace.*
5. *Lidé neprovádějí ty samé simulace (poukaz na subjektivitu).*
6. *Nevíme, jak tento systém implementovat do umělého vizuálního systému.*
7. *Teorie není kompatibilní se současnými teoriemi učení.*
8. *Teorie není kompatibilní se současnými poznatky o fungování mozku.*
9. *Simulace nejsou isomorfní s tím, co simulují.*
10. *Ne všechny simulace prostorových procesů musí být prostorové. Někdy je jednodušší použít výpočet. Nelze vždy zjistit způsob implementace, pokud nemáme informace o designu systému.*

3.3.2.4 Kritika simulátorů

Přestože Barsalou (1999) definuje základní vlastnosti simulátoru, je obtížné přesněji formalizovat jeho činnost. Samotný princip perceptuálního symbolického systému je nastíněn pouze v úrovni teorie a pro ověření je třeba vytvořit jeho počítačovou implementaci. Protože je proces simulace založen na metodách asociace a učení, nabízí se užití neuronových sítí jako přirozená volba. Základní prvky architektury jsou načrtnuty, ale schází implementace modálně specifických obvodů a mechanismů pozornosti, časoprostorový rámec atd. (Barsalou, 2003). Řešení celého systému, tvořeného pozorností, paměťovými moduly, asociacemi více modalit a samotnými simulátory, je v současné době pro realizaci příliš složité. Přesto jsou některé principy teorie použity v modelu realizovaném v této práci. Jedná se multimodální reprezentace, které integrují auditivní a vizuální reprezentace. Bohužel tvorba mentálních simulátorů, které by dokázaly produkovat kombinatorické varianty možných reprezentací je nad rámec této práce. Přínos Barsalouovy teorie je v navržení způsobu integrace symbolických a konceptuálních (perceptuálních) přístupů. Realizovaný model je malým krokem vytyčeným směrem.

3.4 Mentální představy

Při zkoumání povahy reprezentací je nutné zmínit oblast mentálních představ, jelikož se jedná o bohatý inspirační zdroj, který nám poskytuje mnoho empirických faktů použitelných při modelování reprezentačních systémů podobných lidským. Právě mentální představy, které jsou introspektivně evidovatelné, nám mohou pomoci při odhalování mechanismů podílejících se na operacích s reprezentacemi a také způsobu, jakými jsou reprezentace uloženy.

V samotném začátku je nutné detailněji odlišit mentální představy od reprezentací. Představy bývají nazírány jako zážitek (někdy nazývaný kvaziperceptuální), při kterém subjekt vnímá nepřítomný objekt nebo událost. Jedná se tedy o subjektivně zakoušený fenomén. Mentální reprezentace bývá definována jako informace, jež zastupuje v mysli určitý aspekt vnitřního či vnějšího světa, se kterou je možné dále operovat. Na základě těchto operací organismus konstruuje model světa.

Ze zmíněných definic nejsou příliš zřetelné dva důležité fakty. Reprezentace je nadřazena představě. Tvoří základ, ze kterého je možné představy konstruovat. Dále je nutné zdůraznit, že představy jsou vždy konkrétní a aktuální, zatímco reprezentace je připravenost systému představy tvořit, a proto je daleko robustnější, neboť musí obsahovat veškeré důležité aspekty

ke tvorbě rozličných představ. Jak již bylo zmíněno, některé logické směry chybně identifikují představu jako psychologický reprezentační mechanismus.

Představy jsou z výzkumného hlediska oproti reprezentacím ve výhodě pro svou konkrétnost a evidovatelnost. Existuje mnoho výzkumů, které z faktů týkajících se představ vyvozují závěry, jež se vyjadřují k povaze (obtížně zkoumatelných) mentálních reprezentací. Přesto narážíme i v oblasti představ na rozpor dvou typů reprezentace znalostí a dvou teorií významu, se kterým jsme se setkali v předchozích kapitolách. Místo symbolických a perceptuálních teorií kognice se setkáváme s deskriptivní a depiktivní teorií představ.

3.4.1 Depiktivní a deskriptivní teorie

V psychologii velmi známý spor o způsobu mentálního kódování probíhá do dnešních dní.

Prapůvod hledejme v diskuzi mezi Kosslynem a Pylyshynem o povaze mentálních reprezentací (Kosslyn, 2003; cf. Pylyshyn, 2003). Přestože se jedná o velmi rozsáhlou diskuzi, lze říci, že jádrem sporu je otázka, zda existuje imaginativní (perceptuální) kód, či jsou veškerá perceptuální data převoditelná do symbolického kódu, který je pak používán v procesu myšlení či usuzování (Thomas, 1999). Paralela se symbolickými a perceptuálními přístupy je zřejmá. Pro přehlednost shrnuji základní informace o mentálních představách a jejich vlastnostech.

Co o nich víme?

- Jsou fenomenálně evidovatelné.
- Jsou neverbální.
- Jsou obtížně zkoumatelné.
- Bývají spojovány s termínem mentální reprezenace.
- Jejich potenciál využíváme při myšlenkových operacích.
- Dělíme je dle senzorických modalit na vizuální, auditivní, haptické, kinestetické, olfaktorické apod.

Co o nich nevíme?

- Nevíme mnoho o jejich povaze, formátu či kódování.
- Na jakém základě jsou tvořeny?
- Jakým způsobem jsou uloženy?
- Jakým způsobem interagují s jazykovými schopnostmi?

Jak je zřejmé, nejméně informací máme z oblasti, která je psychologickým experimentům nepřístupná, tzn. o způsobu kódování a následném užití. Oproti reprezentacím je možné představy experimentálně zkoumat, přestože je současný stav výzkumu velmi nepřehledný. Nevýhodou je jejich introspektivní povaha. I přes tyto obtíže byl navržen velký počet výzkumných designů, které měly nepřímo odhalit povahu a funkční komponenty tohoto subjektivně prožívaného fenoménu. Jelikož je rozdíl mezi depiktivním a propozičním formátem představ nejlépe identifikovatelný pro zrakovou senzorickou modalitou, využívá se často jako zkoumaných osob lidí, kteří jsou od narození slepí.

Deskriptivní teorie (Pylyshyn)

- Představy jsou tvořeny deskripcemi.
- Jsou uloženy v jazyku podobném výpočetnímu reprezentačnímu formátu.
- Jsou výsledkem operací se symboly.
- Vytváří je fixní symbolická kognitivní architektura.
- Je založena na principech „shora-dolů“.
- Pracuje na principech formální sémantiky.

Depiktivní teorie (Kosslyn)

- Představy mají depiktivní (znázorňující) charakter.
- Jsou identifikovatelné topologickým uspořádáním neuronálních vzruchů.
- Jsou prokazatelné metodami z oblasti neurověd.
- Teorie je problematická postulací „homunkula“.
- Je založena na principech „zdola-nahoru“.
- Pracuje na principech kognitivní sémantiky.

Použití slepých subjektů při experimentech je možným způsobem, jak potvrdit či vyvrátit Kosslynovu kvazi-obrazovou a později depiktivní teorii (Kosslyn, 1994). Základ výzkumů tvoří předpoklad, že od narození slepí lidé nemají zkušenost s vizuální informací, což by se mělo projevit v produkci, prožívání a operacích s mentálními představami.

Obecně lze výsledky studií se slepými subjekty rozdělit do dvou skupin. První skupina výzkumů (Zimler, Keenan, 1983; De Beni, Cornoldi, 1988; Haber, Haber, Levin, Hollyfield, 1993; Cornoldi, Bertuccelli, Rocchi, Sbrana, 1993) dochází ke zjištění, že je vizuální předsta-

vivost během mentálních operací shodná u slepých i vidoucích, na základě podobných výsledků v experimentální a kontrolní skupině. Druhá část (Heller, Calcaterra, Tyler, Burson, 1996; Vecchi, 1998; Aleman, van Lee, Mantione, Verkoijen, de Haan, 2001; Vanlierde, Wagnet-Defalque, 2004) nachází naopak rozdíly ve prospěch vidoucích ZO.

Příčiny odlišných výsledků lze interpretovat vzhledem ke způsobu prezentace úlohy během experimentu. U první skupiny převažuje zadání, kdy lze postup při řešení transformovat do úrovně jazyka a nepotřebujeme využít potenciál vizuálních představ. To se může promítnout do výsledků. Dalším faktorem je rozdílný způsob operací s reprezentacemi, díky užití odlišných senzoričkových modalit u slepých ZO. Jedná se o zastoupení vizuálních představ haptickými, tzn. ZO využívá během testových úloh znalostí získaných hmatem. Jako důkaz demonstrující selhávání slepých ZO slouží experimenty zaměřené na představování si prostoru např. posuzování perspektivy, při kterých nelze použít jako substituci haptické představy (Arditi, Holtzman, & Kosslyn, 1988).

Přestože výzkumy se slepými subjekty odhalily mnoho cenných detailů o povaze představ, nepodařilo se jim dostatečně prokázat nebo naopak vyvrátit tvrzení depiktivní či deskriptivní teorie. Podíváme-li se na současný stav, není situace o mnoho přehlednější. Kosslyn se snaží potvrdit depiktivní teorii využitím metod z oblasti neurovědy, pomocí identifikace topologického uspořádání neuronálních vzruchů odpovídající povaze vnímaného předmětu (Slotnick, Thompson, Kosslyn, 2005). Naopak Pylyshynova teorie je založena na principech klasické počítačové teorie. Jím prezentovaná *nulová hypotéza* je argumentem pro počítačové zpracování libovolné kognitivní domény. Východiska pro její formulaci můžeme hledat již v principech analytické filosofie, teorii fyzického symbolického systému (Simon, Newell, 1963) nebo v argumentech silné umělé inteligence (Searle, 1980).

Právě kvůli odlišným východiskům (empirickým versus teoretickým) a odlišné metodologii nemohou zmiňovaní autoři dosáhnout konsensu. Jelikož bývá empirická oblast v psychologických disciplínách klíčová pro ověřování teoretických východisek, je žádoucí hledat experimentální design umožňující *nulovou hypotézu* testovat. Takovým pokusem je realizovaná pilotní fáze experimentu (Vavrečka, 2006), ve kterém subjekt trávil 4 dny v čistě symbolickém prostředí, přičemž byly zaznamenávány změny v jeho produkci a prožívání vizuálních mentálních představ. Díky nedostatečným metodám zachycení změny, se do problematiky nepodařilo vnést nové světlo.

Zaměříme se přesto podrobněji na některé detaily zmíněných teorií. Jádrem Pylyshynovy deskriptivní teorie (souvisí také s propoziční reprezentací či tacitní znalostí) spočívá ve tvrzení, že představy jsou tvořeny deskripcemi, uloženými v jazyku podobném výpočetnímu reprezentativnímu formátu (Thomas, 2003). Zmíněný reprezentativní systém je podle Pylyshyna natolik

fundamentální, že je možné potvrdit jeho platnost pomocí *nulové hypotézy*, považujíc mentální představy za výsledek operací se symboly. V dané koncepci existuje za fenomenologickou zkušeností fixní kognitivní architektura, která je symbolická. Můžeme ji ztotožnit s fundamentálními principy komputace vedoucí až k Turingovu stroji, jako univerzálnímu výpočetnímu zařízení.

Kosslyn (2003) naopak považuje pojem *nulová hypotéza* za zavádějící. Navrhovaná hypotéza je podle něj nihilistická, protože nás nevede k testovatelným predikcím. Právě netestovatelnost *nulové hypotézy* z ní činí tvrzení mimo vědecký rámec. Pylyshyn se na základě chybných interpretací hypotézy rozhodl upřesnit spornou definici.

„Důvodem pro použití „nulové hypotézy“ není přesvědčení, že představy musí být symbolické, ale že mnoho fenoménů, o které se výzkumníci v oblasti představ zajímají, může být vysvětleno pomocí libovolné teorie schopné adekvátně vysvětlit procesy myšlení, tedy i symbolické teorie implementující „jazyk myšlení“.“ (Pylyshyn, 2002, s. 223)

Zmíněné tvrzení je velmi vágní a proto dále upřesňuje, že *„(n)ulová hypotéza není formálním popisem čehokoliv; jedná se o návrh ... toho, jaký formát reprezentace můžeme použít, pokud zakoušíme vizuální představy. Důvod, proč nazývám symbolickou formu „základní“, je ten, že o ní něco víme (protože ji obsahují různé formální jazyky ... a protože máme formální sémantiku, která nám říká, jak je význam komplexních struktur tvořen z významů jejich částí), a protože splňuje základní minimální nároky, které systém pro reprezentaci znalostí a uvažování musí mít. Také víme, že rekurzivní symbolický systém má vlastnosti jako produktivita, kompozicionalita a systematicklost, což jsou esenciální složky procesů uvažování a reprezentace znalostí.“* (Pylyshyn, 2002, s. 224)

Jak je zřejmé z citace, je jádrem deskriptivní teorie předpoklad formálního symbolického systému a formální sémantiky. Reprezenční systém, který produkuje představy, je tedy založen na přístupu, který jsme v předchozích kapitolách identifikovali jako sémanticky nedostatečný pro oblast modelování kognitivních systémů. Můžeme shrnout, že problematiku dvou druhů sémantik nacházíme i v oblasti mentálních představ.

3.5 Duální kódování

Na druhou stranu existují přístupy, které se snaží obě výše zmíněné teorie propojit. Kosslynova teorie o perceptuální povaze (či perceptuálním podkladu) reprezentací, byla rozvinuta v práci A. Paivia (1986). Autor přichází s teorií duálního kódování. Samotný pojem kódování je zde poněkud nešťastný, jelikož v oblasti kognitivních věd se tento termín spojuje více se

způsobem elementárních operací jednotlivých architektur než s oblastí reprezentační. Vhodnější by bylo hovořit o duálním systému reprezentací – obrazovém a verbálním (propozičním a analogovém, verbálním a neverbálním, symbolickém a subsymbolickém, *imagenu* a *logogenu*).

V teorii se hovoří o tom, že vizuální a verbální informace jsou zpracovávány odlišně, a že si lidská mysl vytváří odlišné reprezentace v jednotlivých typech kódů. *Analogový kód* reprezentuje fyzické stimuly, které pozorujeme v prostředí. Jedná se o formu znalosti, která uchovává perceptuální vlastnosti toho, co bylo pozorováno. *Symbolický kód* na druhou stranu reprezentuje arbitrárně, jako opositum perceptuálního. Reprezentuje v naší mysli symboly jako slova či jejich kombinace, které slouží k vyjádření myšlenek.

Podle Paivia tedy zpracováváme reprezentace ve dvou funkčně nezávislých, ale přesto propojených systémech, verbální a obrazové paměti. Pokud je aktivován jeden typ paměti, jsou zároveň asociovány odpovídající reprezentace paměti druhé. Paiviova teorie je podpořena mnoha výzkumnými studiemi, které hovoří v její prospěch. Nejčastěji bývají zmiňovány výzkumy, které se zabývají *selektivní interferencí*. Jestliže jsou ZO prezentovány současně dvě úlohy, které vyžadují využití stejného kódu (např. *imagenu*), bude ZO vykazovat horší výsledky než u dvou úloh využívajících rozdílný kód (*imagen* a *logogen*). Úlohy ve stejném kódu užívají stejné kognitivní zdroje a interferují spolu (Logie & Baddeley, 1990; De Beni & Moè, 2003). Další silnou argumentační základnu tvoří výsledky neurovědných studií za použití zobrazovacích metod (PET, fMRI). Během nich se prokázalo, že pacienti používají stejná mozková centra při vnímání či představě obrazových materiálů, ale rozdílná pro verbální materiál (Just, 2004). Zároveň bylo zjištěno, že si ZO daleko lépe zapamatovaly materiál, pokud jim byla prezentována informace v obou kódech a studie také poukázaly na zvýšenou mozkovou aktivitu při zpracování abstraktních úloh.

K teorii duálního kódování musíme při interpretaci přistupovat se zvláštní opatrností. Prolínají se nám zde totiž informace ze tří odlišných oborů (informatika, psychologie a neurovědy) a proto bývá často chybně interpretována. Samotné dělení kódu na vizuální a verbální je nepřesné. Pokud přistupujeme k lidskému kognitivnímu aparátu jako k neuronové síti, je třeba hovořit pouze o jediném neurálním kódu, který je dán způsobem komunikace neuronů. Je proto třeba hovořit o dvou reprezentačních systémech – vizuálním a verbálním, což koresponduje s neuroanatomickou strukturou mozku. Auditivní i vizuální kůra obsahuje primární reprezentační oblasti, ve kterých jsou uloženy elementární protoreprezentace samostatně. V těchto oblastech můžeme identifikovat vizuální a verbální centra. Následně jsou systémy integrovány do společných reprezentací, které využívají potenciál obou. Jelikož o tomto

funkčním propojení dosud mnoho nevíme, je možný pouze schematický popis. Pro účely popisu teorie je vysvětlení dostačující.

Paivioův přístup lze interpretovat v kontextu kognitivní sémantiky a integrovat ho do teoretického rámce, který jsme v předchozích kapitolách nastínili. Pokud se vrátíme k teorii znaku, tak jak ji definoval Peirce, můžeme ztotožnit *representamen* s verbálním kódem a *interpretant* s obrazovým kódem. Samotný znak je poté asociací *representamen* a *interpretant*. Reprezentační systém je tvořen vzájemně interagujícími samostatnými systémy. Proto je Paiviova teorie v korespondenci s přístupem kognitivní sémantiky zdůrazňující nutnost konceptuální úrovně (obrazového kódu). Pro vágní definici *imagenu* připodobňující jej spíše k mentální představě než reprezentaci (vyjadřuje konkrétní objekty a jejich perceptuální vlastnosti, nikoliv abstraktní vlastnosti), by mohl být obrazový kód napadnutelný z pozice analytických přístupů jako reprezentačně nedostatečný. Na Paiviovu teorii je však nutné nahlížet jako na teorii psychologickou, která definuje pojmy ve značné míře obecnosti. Mají sloužit pro lepší interpretaci psychologických fenoménů, tzn. pro oblast modelování jsou nedostačující. Používám je zde jako ukázkou psychologické teorie, která koresponduje s perceptuálními a konceptuálními přístupy ke kognici a zdůrazňuje nutnost konceptuální úrovně při budování sémantiky.

Oproti klasickému obrazovému kódu potřebujeme konceptuální úroveň reprezentovat jako abstraktní, neboť zachycuje vlastnosti prostředí v dostatečné míře obecnosti. Jak bylo zmíněno v souvislosti s mentálními simulátory, potřebujeme reprezentační systém, který není „nahrávací“ tzn. nezaznamenává okolní prostředí 1:1, ale zachycuje pouze podstatné a neměnné znaky. Právě tato modifikace Paiviovy teorie z ní činí model použitelný v oblasti kognitivní sémantiky. Verbální kód dokáže reprezentovat symbolickou úroveň a její syntaktické vztahy a modifikovaný obrazový kód tvoří konceptuální úroveň sloužící jako *interpretant* symbolické úrovně, podobně jako v Peircově pojetí znaku.

Přestože Paivio během dlouholeté experimentální činnosti prokázal, že jednotlivé systémy vzájemně spolupracují, povahu či přesnější formalizaci této součinnosti nikdy nerozvedl. Jedním z důvodů je fakt, že zásadně odmítal nazírání kognitivních procesů jako procesů výpočetních. Z dnešního pohledu, kdy je možné výpočetně modelovat i konceptuální úroveň, je jeho vymezení spíše chápáno jako nesouhlas s klasickými výpočetními teoriemi, které jsou založeny pouze na syntaxi a formální sémantice. Pro disertační práci je Paiviova teorie důležitá právě pro svůj důraz na konceptuální a perceptuální úroveň reprezentačního systému.

Následující kapitola tvoří teoretické jádro přístupu, který jsem zvolil pro realizaci svého modelu. Jedná se o způsob propojení dvou rozdílných reprezentačních systémů. Naším cílem bude přesněji formalizovat způsob, jakými jsou symboly (*representamen*, verbální kód), propojeny s koncepty (*interpretant*, neverbální kód).

4 Ukotvení symbolů

Obvyklým názorem symbolistů je, že význam je ustaven skrze „správné“ propojení symbolů se světem. Ono „správné“ propojení je jinak nazvaný problém kognice, který je velmi obtížně řešitelný. Mnoho symbolistů věří, že kognice je pouhou manipulací se symboly, která je uskutečňována skrze nezávislý funkční modul připojený na vstupní zařízení umožňující mu „vidět“ svět objektů, kterému odpovídají symboly. Definice vychází z podivného principu „správného propojení“ či „vidění“, které je pro ustavení symbolů nutné, aniž by bylo blíže vysvětleno. Jedná se o podcenění převodu informací z prostředí do formy, která je zastupitelná pomocí symbolů, tedy o zjednodušování problému ukotvení (Harnad, 1990).

Problém, který byl v počátcích historie kognitivní vědy brán jako marginální, ale který je pro simulaci inteligence podstatný, souvisí s ukotvením symbolů. Situace je následující. Jak již bylo zmíněno v této práci mnohokrát, současné stroje zpracovávají své informace na syntaktické úrovni. Pokud ale systém potřebuje interagovat s reálným světem a pomocí zpracování interních reprezentací „o něm“ provádět operace, potřebuje zvládnout i sémantickou úroveň symbolů. To znamená soustředit se na problematiku tvorby významu (Pfeifer a Scheier, 1999).

Symbol grounding problem (Harnad, 1987; 1990; Newton, 1996; Searle, 1980) je jiná forma vyjádření otázky transdukce, tedy převodu jednoho typu stimulů na jiný. Nejenže neznáme proces převodu z perceptuálních stavů do amodálních symbolů, ale ani zpětný proces převodu do perceptuálních stavů (Barsalou, 1999). Jak již bylo zmíněno, většina formálních teorií tyto otázky přehlíží.

V současné době patří problematika ukotvení symbolů k velmi diskutované oblasti kognitivních věd, přičemž návrhy způsobů řešení začínají v oblasti filosofie mysli a končí konkrétními realizovanými systémy. Jen v oblasti obecného způsobu řešení nalezneme několik návrhů, jak problematiku uchopit. Jelikož některé z nich budou v následujícím textu zmiňovány, uvádím pro přehled názvy přístupů, které jsou s otázkou ukotvení symbolů spojeny:

Argument čínského pokoje (Searle 1980), (Cole, 2004)

Problém ukotvení symbolů (Harnad, 1990)

Problém ukotvení znalostí (Chalmers, 1992)

Perceptuální ukotvení prostorové sémantiky (Regier, 1992)

Problém ukotvení konceptu (Dorffner a Prem, 1993)

Fyzické ukotvení (Brooks, 1993)

Past internalismu (Sharkey a Jackson, 1994)

Problém ukotvení (attachment/tethering) symbolů (Sloman, 2002)

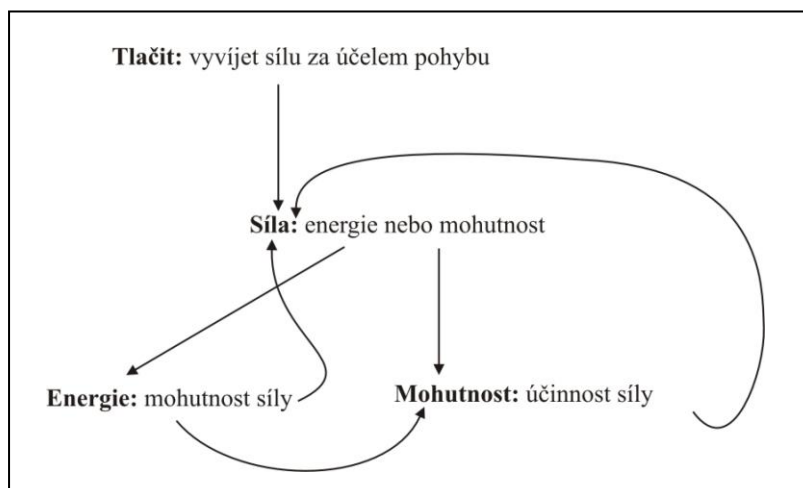
Problém zakotvení (anchoring) symbolů (Coradeschi a Saffioti, 2003)

Bereme-li jako výchozí moment Harnadův článek z roku 1990, můžeme říci, že následná odezva související s rozvojem přístupů k ukotvení symbolů byla značná (viz přehled výše). Pokuším se posoudit jednotlivé teorie a nalézt společné body, které mohou následně sloužit jako základní prvky pro návrh systému blížíící se svými sémantickými schopnostmi lidskému kognitivnímu aparátu.

Potřeba řešení otázky ukotvenosti symbolů je spojena zejména s rozvojem klasické počítačové architektury během druhé poloviny minulého století. Následkem tohoto rozvoje je prudký nárůst aplikací umělé inteligence, zaměřené na oblast zpracování symbolů. Ať už hovoříme o silné umělé inteligenci, fyzickém symbolickém systému či reprezentaci pomocí rámců, ve výsledku dochází ke směřování syntaxe a sémantiky. Lépe řečeno používání syntaktické struktury jako sémantické, bez užití konceptuální úrovně pro reprezentaci významu. Jedná se o celý proud autorů, kteří považují pouze symbolickou rovinu za dostačující pro tvorbu reprezentací (Fodor, Newell, Pylyshyn apod.).

V epistemologické rovině souvisí zmíněná oblast se Searlovým argumentem čínského pokoje, tedy problémem vlastní (intrinsic) sémantiky (vlastního významu nebo také intencionality). V hlavní námitce proti formálním přístupům se poukazuje na to, že pokud je význam symbolů v symbolickém systému nevlastní (oproti vlastnímu významu v našich hlavách), nejedná se o přípustný model sémantiky podobné lidské. Kognice nemůže být pouhá manipulace se symboly (Harnad, 1990). Základní otázkou groundingu je, jak vytvořit funkci a interní mechanismus (např. jako interní reprezentaci odkazující na externí objekt) tak, aby byla vlastní vzhledem k samotnému artefaktu, místo toho, aby byla závislá na interpretaci externím pozorovatelem (Ziemke, 1999). Searle (1980) a Harnad (1990) zkonstatovali, že čistě tradiční a čistě výpočetní způsob UI, založený na programování, nemůže vést ke vzniku inteligentního systému, protože znalost zůstane pro systém vždy nevlastní, a je znalostí pouze pro externího inteligentního pozorovatele (designéra). Proto je nutné, aby znalost vstupovala do systému z jeho prostředí v ukotvené formě.

Pro vytvoření úvodní metafory snad postačí následující ilustrace. Harnad (1990) popisuje čistě symbolický přístup metaforou učení se čínskému jazyku, jako svému prvnímu (mateřskému), pomocí čínsko-čínského slovníku. Podobně se k problematice vyjadřuje Roy (2004), který problém demonstruje na cyklické slovníkové definici jednotlivých pojmů.



Obr. 1 Sémantika symbolického systému (Roy, 2004)

Toto omezení nás nutí vytvořit systém, který nebude obsahovat pouze symbolickou rovinu. Nové požadavky na konstrukci inteligentních systémů (agentů) leží právě v oblasti propojení symbolického systému s rovinou konceptuální. Opuštění „internalistické pasti“ je klíčové pro tvorbu inteligentního systému (Law a Miikkulainen, 1994). Stroje musí být *napojeny* na externí svět, například pomocí kauzálních spojů, které **nejdou** zprostředkované externím pozorovatelem (Sharkey a Jackson, 1996).

Někteří autoři zastávají v oblasti reprezentace znalostí a tvorby inteligentního systému funkcionalistické hledisko (Dennet, 1987). Takový přístup předpokládá, že je možné myšlenkové procesy popsat jako sled abstraktních funkcí, které vyjadřují danou operaci, přičemž jejich fyzická realizace je nepodstatná. Problematické je, pokud abstraktní funkce nedokážeme určit ani odvodit. Jestliže neznáme funkce, pomocí kterých dochází k převodu perceptuálních vjemů na abstraktní kategorie, nemůžeme tyto schopnosti realizovat v umělém systému, díky neschopnosti vyjádřit algoritmickou posloupnost. Funkcionalismus začíná selhávat. Alternativou je užití metod učení, které nám mohou odhalit principy tvorby konceptů. Například použitím neuronových sítí a principu samoorganizace (Kohonen, 1989), popřípadě genetických algoritmů, můžeme vytvořit systém, schopný převádět percepty do konceptuální a následně symbolické úrovně. Odpovídající funkce není predefinována designérem, jelikož ten ji nezná. Funkce se systém učí na základě interakce z prostředí. Jsme nuceni použít odlišné strategie, abychom obdrželi kýžené mechanismy ve funkcionální podobě.

4.1 Obecné řešení ukotvení symbolů

Podle Dorffnera (1993) je užití neuronových sítí, a s ním spojených metod učení, ideální variantou realizace procesu ukotvení. Nejdříve se podíváme na původní Harnadovo řešení procesu ukotvení. Základní otázky, které Harnad ve svém článku (1990) řeší, se týkají následujících témat:

Jak učinit sémantickou interpretaci formálního symbolického systému vlastní systému?

Jak může být význam bezvýznamných symbolických prvků, se kterými je manipulováno pouze na základě jejich (arbitrárních) tvarů, ukotven pouze v dalších bezesmyslných symbolech?

Odpovědí je návrh systému, schopného problémy řešit. Je rozdělen na tři reprezentační úrovně, na jejichž základě dochází k ukotvení (Harnad, 1990):

- a) ***Ikonická reprezentace*** - vzniká převodem vzdálených objektů či událostí na blízké senzorické projekce.
- b) ***Kategoriální reprezentace*** – vzniká na základě vrozených a naučených mechanismů a dokáže ze senzorických projekcí abstrahovat společné vlastnosti objektů či událostí – kategorie či koncepty.
- c) ***Symbols*** - jsou názvy, přiřazené objektovým nebo událostním kategoriím

V Harnadově řešení můžeme opět identifikovat Peircovo schéma znaku. Ikonická reprezentace zprostředkuje přenos externího objektu na senzory systému, poté je tvořen *interpretant* jako kategoriální reprezentace a následně je konceptu přiřazen *representamen* ve formě symbolického označení. Navržený systém je hybridní, zčásti konekcionistický a zčásti symbolický. Systém neobsahuje pouze symbolický modul, protože symboly jsou ukotvené pomocí kategoriálních reprezentací a tedy vlastní pro daný systém (Harnad, 1990).

Již zmíněný Dorffner (1993) navrhuje extrémnější přístup, nazývaný *radikální konekcionismus*, založený na použití neuronových sítí v celém systému, tedy jak pro kategoriální reprezentace, tak pro symbolický modul. Takový systém již není hybridní, jelikož používá pouze jednu architekturu.

Radikální konekcionismus je cestou ke tvorbě komplexních konceptuálních struktur, ale není tak systematický jako kombinování externích symbolických znaků. Je subkonceptuální, jelikož stavební bloky leží o úroveň níže, a také nesymbolický, protože stavební prvky nejsou

arbitrární. *Symbol grounding* se v jeho případě stává *Koncept groundingem* - modelováním formací konceptů ukotvených ve zkušenosti (Dorffner, 1993). Dorffnerovo vysvětlení přikládá důležitost formování konceptů, ale nechává otázku následné manipulace v symbolické úrovni nevyřešenou. Přestože lze neuronové sítě použít jako architekturu vhodnou pro symbolické operace (Touretzky, Hinton, 1985), je užití klasické architektury vhodnější a efektivnější.

Symbolický systém pomocí neuronových sítí

Formálními principy, na kterých jsou založené neuronové sítě, jsou hlavně asociační a statistické struktury kauzálních vztahů v daných dynamických systémech. Podle Harnada (1990) jsou neuronové sítě dobrým podkladem pro implementaci dynamických systémů.

Existují však určité pochybnosti o tom, zda může být konekcionismus symbolický. Argumenty proti možnostem neuronových sítí v oblasti práce se symboly podali Fodor a Pylyshyn (1988). Podle nich sítě, jako symbolické systémy, nesplňují požadavek složenosti (7) a systematicity (8). Neuronové sítě nepracují na základě formální syntaxe, takže není možná systematická sémantická interpretace.

Přesto v roce 1985 Touretzky a Hinton vytvořili model, který měl ukázat schopnosti konekcionista architektury v oblasti zpracování symbolů. Jejich *konekcionista produkční systém* umožňoval reprezentovat a manipulovat s arbitrárními a diskrétními *symboly*. Přestože síť dokáže pracovat s kontinuálním stavovým prostorem (analogově), smysluplné jsou pouze diskrétní *symboly* (Touretzky a Hinton, 1985). Autoři museli takovou síť vytvořit pomocí speciálního nastavení vah a prahů. Mělo jít o důkaz, že lze pomocí neuronových sítí operovat se symboly či na symbolech. Z čehož vyvstává otázka, zda je subsymbolický přístup pouze alternativní implementací symbolického systému?

Objevují se názory pomíjející rozdíl mezi symbolickým a subsymbolickým, jelikož každá neuronová síť může být implementována do Turingova stroje, který je již z definice symbolický (Dorffner, 1993) a také lze vytvořit neuronovou síť napodobující svou činností Turingův stroj.

4.1.1 Dělení přístupů

Podle užití architektury či přístupu můžeme rozdělit řešení problematiky ukotvení na kognitivistické, konekcionistické, enaktivní a popř. hybridní. Z již zmíněných řešení je zástupcem konekcionistického přístupu Dorffnerův *radikální konekcionismus*, hybridního pak Harnadovo řešení.

Kognitivistické paradigma v zásadě rozlišuje kognici a percepci. Příkladem může být Fodorovo rozlišení na vstupní systémy (*low-level* percepce) a centrální systémy (myšlení, řešení problémů). Kognitivisté se většinou soustřeďují na ukotvení atomárních reprezentací vstupního systému pomocí senzorických nebo senzomotorických invariant. To znamená, že kauzální spojení mezi agentem a prostředím jsou tvořena pomocí propojení atomárních interních reprezentací s externími entitami či objektovými kategoriemi. Takto ukotvené atomy jsou následně použity jako stavební kameny pro tvorbu komplexních reprezentací, ze kterých je konstruován koherentní reprezentační model světa (Ziemke, 1999).

Enaktivní přístup se od kognitivního odlišuje nepřítomností centrálních systémů a také jinou formou reprezentací. Existují důkazy, že oblasti primárně spojené s motorickou kontrolou jsou použity při tak odlišných činnostech, jako je pochopení jazyka a pracovní paměť. Pokud je tento fakt kombinován s myšlenkou, že motorická kontrola je podstatou řízení sensomotorické zpětnovazebné smyčky, pak můžeme považovat motorickou kontrolu jako součást procesu pobídek (*affordance*). Protože *affordance*, tedy vnímaná dostupnost objektu z hlediska možné interakce, není pouze motorické povahy, je možné chápat motorický kontrolní systém jako primitivní sémantický (významový) procesor (Anderson, 2006).

Například u hmyzu spočívá reprezentace pobytu v prostředí ze sensomotorických závislostí (kontingencí), ve kterých je reprezentace percepce spojená s reprezentací motoriky. Jedná se o posun v chápání prostředí, jako závislostí jednání na důsledcích. V psychologii a biologii se objevují požadavky na vyšší propojenost sensomotoriky již v pracích Gibsona (1979) či Prinzze (2003), který zjistil interferenci senzorických a motorických procesů.

Podle enaktivního přístupu je znalost nazírána jako distribuovaně vtělená (rozprostřená a vtělená v senzorech, efektorech, nervovém systému apod.) a částečně také ležící v prostředí (Varela, 1991). V praxi se jedná o využití konekcionistické architektury pro realizaci systému. Neuronové sítě umožňují distribuovanou formu reprezentace, jak v senzorické, tak v motorické části systému. Jsou také vhodné při jejich vzájemném mapování. Díky paralelnímu způsobu příjmu a zpracování informace se hodí k reprezentaci senzorických vstupů z prostředí, které je dle enaktivního přístupu natolik strukturované a informačně bohaté, že

umožňuje inteligentnímu systému vytvářet reprezentace na základě interakce s ním. Enaktivní přístup je v podstatě rozšířenou variantou konekcionismu. Využívá neuronových sítí pro jejich biologickou plausibilitu a rozšiřuje způsob tvorby systémů o motorickou část, pomocí které může systém interagovat s prostředím.

Podle Searla (1980) je kognitivistický a enaktivní přístup spojen se vtělenými systémy (jsou kauzálně spojené se svým prostředím) než s nevtělenými programy. Oba přístupy vyžadují plně ukotvené systémy. V kognitivistickém požadujeme ukotvenost propojení vstupních a centrálních systémů. Enaktivní přístup vyžaduje vývoj celku v interakci s prostředím.

4.1.2 Reprezentacionismus a nonreprezentacionismus

Přestože jsou zmíněné přístupy odlišné ve způsobu použitých architektur a také způsobu, jakým jsou informace z prostředí reprezentovány a zpracovány, řadí se do společného rámce, který zastřešuje přístupy založené na předpokladu, že pro napodobení inteligentního systému a jeho kognitivních schopností, je nutné vytvářet interní reprezentace externího prostředí. Tento směr se obecně nazývá reprezentacionismus. Přestože zmíněné varianty se ve způsobu reprezentace liší, většinou jsou založeny na základním schématu, které se objevuje jak u Peircovy definice znaku, tak u Harnadova modelu ukotvení symbolů. Obecný postup je následující:

- a) Zachytit podstatné znaky, které jsou společné perceptuálním datům*
- b) Abstrahovat*
- c) Identifikovat abstrakci jako obsah kategoriální a konceptuální reprezentace*
- d) Použít takové reprezentace k ukotvení symbolů*

V oblasti kognitivních věd se vyskytují také přístupy, které na reprezentace zcela rezignují. Základem jsou Brooksovy práce z oblasti reaktivní robotiky. Podle hesla „Svět je svou nejlepší reprezentací“ nejsou reprezentace v takto navrženém systému nutné a nepotřebujeme řešit otázku ukotvení symbolů. Přesto při popisu robotických agentů hovoří Brooks o tom, že obsahují určitý stupeň „fyzického ukotvení“ kvůli tomu, že jsou propojeni s prostředím pomocí senzorů a efektorů (Brooks, 1993). Podle Ziemkeho (1999) je takové řešení klamné, jelikož nedochází k ukotvení chování ani interních mechanismů. Ukotvení vnitřních mechanismů není možné, jelikož jsou konstruovány jako vzájemně kooperující (z hlediska *subsumpce* lépe „potlačující se“) deterministické *finite state automata* a fungují na základě mechanismů, které

jsou předprogramované. Přesto jsou Brooksovy systémy schopné interagovat s prostředím. Pokud bychom systémy analyzovali, došli bychom ke zjištění, že systém je dostatečně senzory vybavený, aby dokázal přijímat informace z prostředí. Nedostatky se projevují při následném zpracování. Jestliže má systém všechny mechanismy předprogramovány a neobsahuje žádnou reprezentační úroveň, dokáže fungovat pouze reaktivně a není schopen inteligentní akce, která vyžaduje znalost o nepřítomném objektu. Právě neschopnost učení a následného přizpůsobení činí takové systémy efektivní pouze pro jednoduché motorické akce, které vyžadují reflexivní reakce či jejich kombinace. Systém je obtížné posoudit z hlediska ukotvení symbolů, jelikož u něj k žádným reprezentačním procesům nedochází. Přestože díky souhře svých elementárních funkcí dokáže inteligentní akce simulovat, nelze jej zařadit mezi způsoby řešení reprezentace významů.

4.1.3 Nulový sémantický závazek

Právě z těchto důvodů bude nutné, abychom si před detailnější analýzou procesu ukotvení a možných řešení, upřesnili kritéria a požadavky, podle kterých můžeme dané návrhy hodnotit. Takový pokus provedla dvojice autorů Tadeo a Floridi (2005), kteří přicházejí s podmínkou *nulového sémantického závazku*. Podmínku dále nazývají *Z kritériem* a snaží se pomocí něj posoudit současné systémy, které mají ambice být řešením *symbol grounding problemu*. Autoři vycházejí z Harnadovy předpokladu, že interpretace symbolů musí být vlastní vzhledem k systému (nemůže být nevlastní nebo parazitická). Pro návrh a tvorbu systému z toho plynou následující omezení (Tadeo, Floridi, 2005):

1. *Není povolena žádná forma vrozenosti, neexistují sémantické zdroje (virtus semantica), které bychom považovali jako předinstalované do autonomního agenta.*
2. *Není povolena žádná forma externalismu - významy nelze do systému vkládat zvenčí pomocí designéra, který má sémantickou znalost.*

Tyto dva body nemají být v rozporu se třetím:

3. *Agent má zdroje a schopnosti (výpočetní, syntaktické, procedurální, perceptuální, schopnosti učení) ukotvit symboly.*

Autorům se během analýzy současných přístupů podařilo prokázat, že ani po patnácti letech (1990-2005) neexistuje systém, který by byl schopen úspěšně vyřešit problém ukotvení symbolů. Jelikož je jejich způsob aplikace *Z kritéria* u některých systémů nepřesný, provedu podrobnější analýzu jejich závěrů.

Podle Tadeo nesplňuje již základní hybridní varianta systému (Harnad, 1990) *Z kritérium* (nulový sémantický závazek). Problém spočívá ve způsobu, kterým hybridní systém nachází neměnné vlastnosti senzorických projekcí umožňujících kategorizaci a identifikaci objektu. Problematickou se jeví neschopnost takových systémů dojít k označení třídy (například zvíře) bez předchozí znalosti. Pokud se jedná o učení s učitelem (*supervised*), dodá označení třídy designér. Pokud ale probíhá učení bez učitele (*unsupervised*), není jasné, jak symbolické označení třídy vznikne. Lze jej vysvětlit pouze sociálním kontaktem a možnostmi během komunikace. Autoři *Z kriteria* považují takový způsob za neadekvátní a odporující základním požadavkům. Na druhou stranu by při takto striktním vymezení, nedokázal úlohu zvládnout žádný systém (včetně člověka). Jedná se o podobný problém, se kterým jsme se setkali už v části věnované reprezentacionismu. Pokud hovoříme o vzniku symbolické roviny, musí mít tato rovina nějaký účel (procesy myšlení, intersubjektivní komunikace apod.). Jestliže neexistuje možnost sociální komunikace, je zbytečná symbolická úroveň, jelikož nenachází v systému opodstatnění a nemůže ani dojít k jejímu vzniku. Požadavek autorů je paradoxní: vyžadují emergenci symbolické roviny, přičemž se ale v prostředí, ve kterém se agent pohybuje, žádná forma symbolů neobjevuje. V této interpretaci nemá symbolická úroveň možnost vzniknout. Při vzniku kategorií během *učení bez učitele*, se jedná o podobný problém. Neexistuje-li komunikace, probíhá kategorizace na základě diskriminace (způsob tvorby kategorií je subjektivní), ale nemůže dojít k integraci se symbolickou vrstvou reprezentace, jelikož není přítomna.

Další problematické místo vidí autoři v neschopnosti systému rozpoznat, na které vlastnosti či vzorce se má zaměřit a které má považovat za invariantní. Podle nich musí být taková možnost dopředu napevno zakódovaná, což vyžaduje účast designéra. Daný postup je podle autorů *Z kritéria* nevlastní. S takovým názorem ale nelze souhlasit. Je jasné, že systém založený na *unsupervised učení* (učení bez učitele) a principech samoorganizace, dokáže zachytit strukturu prostředí a organizovat perceptuální kategorie na základě podobnosti. Je možné vytvořit systém ukotvování použitím dvou subsystémů, přičemž jeden operuje s vizuálními vstupy a druhý s jazykovým označením kategorií. V takovém případě můžeme integrovat systémy do společné multimodální vrstvy a vytvářet kategorie, které obsahují symbolickou komponentu. Takový přístup používám v navrhovaném modelu.

Autoři nulového sémantického závazku se mýlí, jestliže pokládají metody učení, sociální interakci a komunikaci za nepřípustné (podrobněji o problematice sociální interakce v kapitole

věnované sociálnímu symbol grounding). Bez těchto procesů lze ukotvovat symboly, pouze pokud předpokládáme apriorně formální symbolický systém „uvnitř“ systému. Jedná se o spor týkající se konceptového empirismu. Základem je požadavek apriorních konceptů, pomocí kterých lze kategorizovat informace z prostředí. Zde se otázka přesouvá do úrovně, zda jsou tyto základní mechanismy syntaktické či sémantické? Než zodpovíme otázku, shrneme si konkrétnější požadavky na tvorbu systémů, jak ji navrhuji autoři *Z kritéria*. Jedná se o:

1. *Bottom-up, senzomotorický přístup*
2. *Top-down zpětnou vazbu, která umožní sladění nejvyšší vrstvy ukotvených symbolů a spodní vrstvy, zajišťující senzomotorickou interakci s prostředím.*
3. *Schopnost agenta vytvářet reprezentace*
4. *Schopnost agenta kategorizovat a abstrahovat*
5. *Schopnost agenta komunikovat s ostatními, tedy ukotvovat symboly diachronicky, čímž předejdeme Wittgensteinovu problému „soukromého jazyka“*
6. *Evoluční přístup během vývoje 1-5*
7. *Uspokojení Z kritéria během vývoje 1-6*

Analýzou *Z kritéria* se prokázalo (Vavrečka, 2006b), že některé jeho aplikace nejsou zcela přiléhavé, a že argumentace pomocí evolučního přístupu může jít za hranice designérských možností.

Pokud shrneme uvedené argumenty, lze říci, že požadavek nulové účasti designéra při tvorbě systému by znamenal přílišnou restriktivitu. Ve svém důsledku popírá existenci vrozených (v případě umělého systému předdefinovaných) mechanismů, např. reflexivní výbavy či implicitních protomechanismů vnímání. Striktní aplikace kritéria vlastně popírá jeho třetí dodatek, který umožňuje systému mít zdroje (výpočetní, syntaktické, procedurální, perceptuální, učení) k ukotvení symbolů. Jelikož tyto zdroje nelze v případě umělých systémů získat dědičně, nemůžeme úlohu designéra zcela eliminovat. Musí mít možnost vytvořit architekturu systému a nastavit jeho základní parametry. Systém má poté výpočetní zdroje (například ve formě neuronové sítě) sloužící k reprezentaci znalosti. Architektura je schopná reprezentovat význam již svou strukturou, tzn. pokud má výpočetní zdroje, je schopna reprezentovat a vytvářet „významy“. Již dle autorů teorie znaku (Saussure, Peirce) či výchozích principů kognitivní sémantiky, nelze oddělit syntax od sémantiky, a proto je obtížné v systému odlišit, zda se jedná o syntaktické či sémantické zdroje. Z toho důvodu hodnotí autoři *Z kritéria* dosavadní principy ukotvení, i přestože jsou založeny na perceptuálních a konceptuálních principech, jako nedostačující.

4.1.4 Sociální symbol grounding

Podíváme-li se na *symbol grounding* coby badatelskou oblast, zjistíme, že množství výzkumu dosáhlo hranice, při kterém dochází k rozdělení problematiky na více částí. V současnosti probíhají výzkumy způsobu asociace symbolů s koncepty v kontextu jednotlivého agenta (individuálně) - *fyzický symbol grounding*, a také způsobu interakce agentů (skupinově), na jejichž základě mohou mezi sebou sdílet významy - *sociální symbol grounding*.

V případě *fyzického symbol grounding* se nejedná o Brooksův (1993) fyzický grounding, ale o proces ukotvení založený na derivaci kategorií z prostředí jedincem. *Fyzický symbol grounding* vychází z Harnadova (1990) ukotvení dle schématu *externí entity* \Leftrightarrow *interní reprezentace* \Leftrightarrow *symboly*. Propojení by dle Roye (2004) mělo být obousměrné.

Sociální symbol grounding (SSG) je definován jako interindividuální vyjednávání o výběru sdílených symbolů (slov) a jejich ukotvených významů. Zabývá se vývojem slovníku ukotvených symbolů v populaci kognitivních agentů (Cangelosi, 2006). Z hlediska evoluce se jedná o emergenci jazyka od prelingvistické formy tzn. bez explicitních symbolických a komunikačních významů, z hlediska ontogeneze o proces akvizice jazyka a kultury. V první fázi probíhá imitace okolí, která vede ke konstrukci lingvistické znalosti. SSG může fungovat také na základě dočasných interakčních her mezi individui. Zmíněným tématem se zabývalo mnoho badatelů, např. Steels (1996) v modelu *Talking Heads*, popřípadě Cangelosi (Cangelosi 2001; Cangelosi a Harnad 2000; Munroe a Cangelosi 2002). U nás se problematice podrobněji věnuje Takáč (2005, 2006, 2007).

Pro lepší pochopení si uvedeme, jakým způsobem experimenty probíhají. Například Steelsovy *Talking Heads* prozkoumávaly dva typy *sociálního ukotvení* - ontogenetické i evoluční. Tvorba významu je modelována pomocí diskriminační hry (Steels, 1996). Modelujeme diskriminační fázi, jak jí popsal Harnad a tvoříme kategorie ve třech fázích: vektory vlastností z kontextu jsou kategorizovány b) kategorie, které odlišují téma od dalších segmentů v kontextu jsou identifikovány c) ontologie kategorií jsou adaptovány. Každý robot hraje diskriminační hru pro získání (potencionálního) tématu individuálně. Téma je segment z konstruovaného kontextu (popsané vektorem vlastností) a je arbitrárně vybráno z kontextu. Agent, který je mluvčím dané téma komunikuje, přičemž naslouchající agent se snaží odhadnout, co mluvčí komunikuje. Bere v potaz všechny segmenty v kontextu jako potencionální téma a hledá, který segment má komunikující na mysli.

Podobný model realizoval Vogt (2003). Ve svém přístupu využívá k procesu ukotvení také komunikaci. Agenti hrají jazykové hry, čímž na základě mnoha interakcí vyladují označení

svých kategorií či objektu, které percipují ve společně sdíleném prostředí. Dále se snaží (Vogt, 2003) pomocí multiagentních systémů vytvořit simulaci ukotvení a vzniku jazyka. Ve velmi komplexním prostředí by mělo dojít k emergenci a následnému studiu jevů, které při menších simulacích nedosáhneme.

Cangelosi (2001) simuloval emergenci učení kompozičního jazyka a získal silnou evoluční tendenci vývoje struktury predikát-argument podobný pravidlu „sloveso-podstatné jméno“. V Cangelosi a Parisi (2001) je explicitní „sloveso-podstatné jméno jazyk“ vyvíjen v populaci simulovaných organismů. Model zdůrazňuje výhody jazyka nad nelingvistickým chováním a také asymetričnost evoluce podstatných jmen před slovesy. Analýza prokázala (Cangelosi, Harnad, 2000), že slovesa produkují více zřetelné interní reprezentace objektů a událostí, protože mají přímější vliv na motorické akce, kterými organismus reaguje na senzorické vstupy. Jelikož v experimentu není věnováno mnoho místa architektuře sítě a není identifikováno uložení interních reprezentací sloves a podstatných jmen v různých částech sítě, jsou výsledky diskutabilní.

V této kapitole jsem provedl pouze krátký nástin výzkumů sociálního symbol groundingu a jeho možností. V dalších částech se dané oblasti nebudeme věnovat, neboť pozornost soustředím pouze na fyzický symbol grounding. Důvodů je několik. Zaprvé jsou multiagentní systémy příliš náročné na realizaci, zadruhé při jejich činnosti působí mnoho faktorů ovlivňujících výsledek bez možnosti přesné analýzy příčin. Navíc neznáme přesné mechanismy v oblasti fyzického ukotvení, abychom mohli přejít do roviny sociální. Poslední argument je sporný, jelikož nelze posoudit, zda se v oblasti sociálního přenosu symbolu nevyskytují mechanismy, které by napomohly řešení ukotvení v úrovni jednotlivce. Jelikož nemáme přesnou znalost potřebných mechanismů, je optimálním způsobem provádět výzkum v úrovni fyzické i sociální.

4.2 Proces ukotvení

Zkusme se podívat na proces ukotvení z hlediska teorií znaků, které byly nastíněny v předchozích kapitolách. Z lingvistického pohledu se jedná o propojení mezi konceptem a zvukovým vzorcem, který je brán jako smyslový vjem. Saussure (1965) se soustředil pouze na zvukové vzorce, které jsou spojeny s přirozeným jazykem, tzn. základním elementem je slovo. V původní koncepci odlišoval slyšené od psaného a považoval psaný projev jako označení zvukových vzorců, později došlo k revizi, ve které jsou zvukové i psané označení brány za rovnocenné.

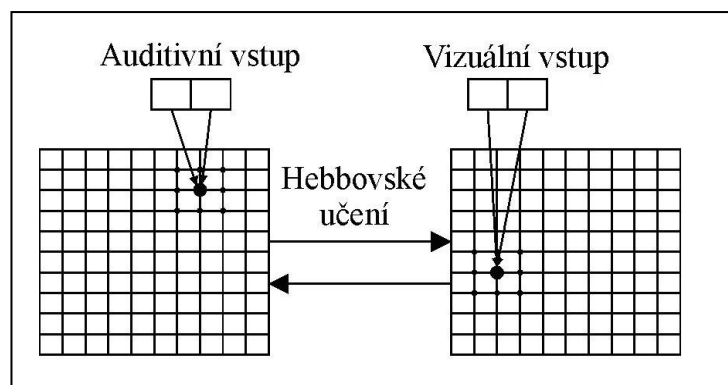
Způsob propojení formy a konceptu je již u Saussura (1965, s. 66) definován jako asociační vazba. Je zcela nezávislá a vytvořena dříve než koncept i forma. Obtížné je stanovit hierarchii či souvztažnost výše zmíněných. Objevují se názory, že forma (*signifiant*) je od konceptu (*signifié*) zcela separovaná a funguje na principech vlastní autonomie (Lechte, 1964). Takový názor vede až k postulaci samostatného symbolického systému, jehož nevýhody byly zmíněny. Z opačného tábora zaznívají hlasy, že koncept předchází formě (Carroll, 1960). V post-saussurovském období se badatelé přiklánějí spíše na stranu primárnosti formy (*signifiant*), z čehož vyplývají obtíže spojené s ukotvením symbolů.

Problematické je již samotné oddělení formy od konceptu či významu, což bývá spojováno se známým dualismem obsahu a formy. Metafora je ale v konečném důsledku chybná, jelikož posuzování formy jako kontejneru, do kterého umístíme význam, by znamenalo možnost existence významu jako samostatné entity. Takový význam může být extrahován bez aktivního procesu interpretace, tzn. forma není sama o sobě smysluplná (Chandler, 1995). Výsledkem je čistě konceptuální systém (v podobném duchu hovoří Barsalou (1999) nebo zastánci extrémního konceptualismu).

Dalším důležitým rysem Saussurovy teorie je, že nedefinuje znaky na základě jejich esenciální či intrinsické (vlastní) povahy. Znaky se definují vztahy mezi sebou. V rámci jazykového systému závisí vše na vztazích, žádný znak nemá smysl sám o sobě. (Saussure 1983, 121). Připomínám, že hovoříme o znacích (symbol+koncept), nikoliv pouze o symbolech. Důraz je kladen na strukturu (v logice systém axiomů), vztahy (pravidla vyplývání). Z toho vyplývá, že znak nemůže mít význam sám o sobě, ale pouze v určitém kontextu. Oproti formálním je Saussurova teorie znaku komplexnější. Obsahuje formální i konceptuální vrstvu.

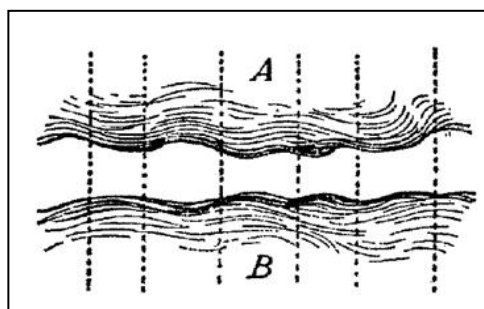
Právě konceptuální vrstva činí ze systému psychologicky plausibilní model. Jako u ostatních teorií se setkáváme s problémem, jak nadefinovat vzájemný vztah mezi formou a obsahem. Přestože Saussure postuluje preexistující asociaci, nevyhneme se rozdělení znaku na dvě strukturální úrovně. Saussure je nazýval rovinou zvukovou (forma či *signifiant*) a rovinou myšlenkovou (obsah či *signifié*).

O tom, že zvuková či psaná podoba znaku může obsahovat vlastnosti, pomocí kterých je lze primárně kategorizovat (na úrovni fyzické podobnosti) si více řekneme v části věnované návrhu vlastního modelu. Ve zkratce se jedná o primární reprezentace založené na fonetické či grafické podobnosti, které můžeme využít při reprezentaci jednotlivých vrstev znaků pomocí neuronových sítí (viz Obr. 2).



Obr. 2. Vztah mezi obsahem a formou je zajištěn pomocí metod učení.

Přestože Saussure postuluje vazby mezi formou a konceptem (zvukem a myšlenkou), znázorňuje je (viz Obr. 3) jako dva autonomní systémy (A a B), které nejsou propojeny, pouze arbitrárně strukturovány (tečkované čáry). Máme dva systémy, u kterých předpokládáme propojenost, přičemž je postulujeme jako samostatné, což je základní rozpor a problém, který tvoří jádro sporu o způsob ukotvení abstraktních a arbitrárních znaků s rovinou konceptů.



Obr. 3 Saussurovo propojení obsahu (A) a formy (B) pomocí arbitrární strukturace (Chandler, 2007)

Kromě formy a významu definuje Saussure také hodnotu znaku. Ta závisí na vzájemném vztahu mezi znaky, a je tedy závislá na kontextu. Znak je podle něj více než suma částí (obsahu a formy), ze kterých se skládá. Právě hodnota mu přiřazuje vztah k ostatním znakům a také systému jako celku. Opět nacházíme paralelu k logickým systémům. Hodnotu lze chápat jako pravdivostní hodnotu. V tomto případě je však hodnota závislá na kontextu (je v kontextu proměnlivá), což odpovídá systému přirozeného jazyka. V logice je pravdivostní hodnota fixně daná a v systému je manipulovaná dle pravidel vyplývání. Oba návrhy obsahují implicitní předpoklad, že systém je úplný a bezsporný. V logice je takový princip zaručen definicí axiomů a pravidel vyplývání, Saussure pouze implicitně předpokládá, že systém funguje na pravidlech holismu, přičemž ty jsou dány gramatikou. Systém přirozeného jazyka však nelze

definovat jako úplný a bezesporný, což činí jeho užití na základě hodnot znaků diskutabilní. Zároveň je nutné zdůraznit, že pro Saussura jsou složky znaku obsah a forma (symbol a koncept), a pokud o jednotlivých úrovních hovoří jako o samostatných, dochází k rozporu s jeho předchozími tvrzeními. Přes některé rozpory je pro nás Saussurův přístup inspirující, jelikož obsahuje myšlenky, které se týkají způsobu propojení konceptuální a symbolické roviny.

4.2.1 Diskriminace a identifikace

Pro lepší pochopení mechanismů podílejících se na ukotvování symbolů, se detailněji podíváme na jednotlivé procesy, o kterých jsme v souvislosti s tímto tématem hovořili. Kategorická percepce (percepce kategorií) byla prokázána u zvířat (Zentall, 1986) i lidí (Goldstone, 1994). V oblasti umělých systémů je zmíněný druh kategorizace v centru pozornosti, přestože je oblast percepce a vnímání stále nedostatečně prozkoumaná.

Funkce kategoriální percepce v oblasti symbol groundingu spočívá v interakci mezi diskriminací a identifikací. Jedná se o základní mechanismus pro tvorbu kompaktních reprezentací. Samotná kategorizace je prvním krokem v procesu ukotvení a její výsledek tvoří podklad pro následné mapování do symbolické roviny reprezentace. Pro studium nám nejlépe poslouží lidský kognitivní aparát. Lidé jsou schopni diskriminovat, manipulovat, identifikovat a popisovat objekty, události a stavy věcí a mohou produkovat deskripce a odpovídat na deskripce objektů, událostí a stavu věcí (Harnad, 1990). Cílem kognitivní teorie je zjistit, jak to lidé dokážou. Podívejme se tedy na proces kategorizace detailněji.

Diskriminace je posouzení, zda jsou dva vstupy stejné nebo rozdílné, popřípadě jak jsou rozdílné. Jedná se o relativní posouzení, tzn. hledání míry shody. Identifikace je přiřazení arbitrární odpovědi – „jména“ - klasifikující vstupy, které následně tvoří kategorii podle shodného prvku. Identifikace je absolutní posouzení, rozhodující, zda je vstup členem kategorie či nikoli (Harnad, 1990).

Schopnost diskriminace je zajištěna pomocí *ikonických reprezentací* (Harnad, 1987). Diskriminace je tedy podle Harnada nezávislá na kategorizaci, jelikož mohou diskriminovat, aniž bych věděli, co dané objekty znamenají. To platí pouze v případě, že předpokládáme jednosměrný proces (pasivní) při zpracovávání a kategorizaci podnětů (nebo v případě, že je podnět prezentován poprvé). Pokud předpokládáme aktivní anticipující subjekt, který diskriminuje na základě předchozí zkušenosti (a v případě existence apriorních schémat vnímání je tomu tak vždy), nemůžeme hovořit o nezávislosti diskriminace na identifikaci. Proto nelze Harnadovo tvrzení zobecňovat na všechny modely.

Objevují se i otázky týkající se procesu kategorizace jako celku. Může nám jediná kategorizace, popřípadě kategorizace podle jednoho kritéria, sloužit k vytvoření dostatečně robustní reprezentace entity ve světě? Dorffner (1996) navrhuje rozšíření kategorizačních vrstev a jejich následné zpětné propojení, čímž bude docházet ke kategorizaci podle více kritérií, popřípadě k tvorbě kategorií abstraktních. V nonsymbolické úrovni (bez identifikace se symboly) je možné s výslednou kategorií dále pracovat a znovu na ní aplikovat mechanismy, které umožní její diskriminaci podle více atributů. Problematické se jeví vytvoření dostatečně důmyslných mechanismů víceúrovňové, paralelní, heterarchické a zpětnovazebné kategorizace, které dokážou vytvořit robustní reprezentace, schopné plně napodobit proces vnímání a následně myšlení.

Druhá část procesu ukotvení je obtížnější, neboť potřebujeme nalézt způsob propojení mezi kategoriální reprezentací (*interpretant*) a symbolickou rovinou (*representamen*). Potřebujeme najít ekvivalenci mezi obrazovou reprezentací a označením, přičemž neexistuje žádná podobnost mezi konceptuálně zpracovanou obrazovou informací a textovým (symbolickým) vyjádřením. Jedním z řešení je postulace procesu mediace (zjednávání) perceptuálních reprezentací (např. Harnad, 1987; Höffding, 1891; Neisser, 1967). Každý amodální symbol je asociován s odpovídajícím konceptem v dlouhodobé paměti. (Barsalou, 1999). Proces je bohužel definován pouze v teoretické úrovni.

4.2.2 Rozšířený princip ukotvení

Jestliže Harnad (1990) definoval *grounding* jako proces diskriminace a identifikace předmětu, tedy jeho kategorizace a následné symbolické označení, jsou současné požadavky na proces ukotvení větší. Například Roy definuje *grounding* jako kauzálně-prediktivní cyklus, pomocí kterého agent ovládá soubor věr o okolním světě. Proces *groundingu* vyžaduje kauzální i prediktivní vztah mezi vírou a referentem (Roy, 2004). Hovoří o třech aspektech významu.

1. Jazyk referuje k entitám a situacím ve světě.
2. Jazyk je situovaný.
3. Agent používá jazyk k dosahování cílů.

Následně definuji proces ukotvení jako kauzálně-prediktivní cyklus, pomocí kterého systém ovládá soubor věr o okolním světě. *Víra* je dle něj explicitní informační struktura, která existuje v mysli agenta. Pokud má být víra (např. hrnek je na stole) ukotvena, musí být splněny 2 podmínky:

- 1) hrnek musí vytvářet víru skrze přírodní zákony (proud světla, fyzický kontakt, senzorická transdukce)*
- 2) víra musí podporovat budoucí predikce ohledně chování hrnku i vzhledem k akcím, které s ním může agent provést.*

Proces ukotvení vyžaduje kauzální i prediktivní vztah mezi vírou a referentem. Cyklický proces následně zajišťuje interpretačně kontrolní smyčku, implementovanou do systému. (Roy, 2005a). Pouhé ukotvení na základě kauzality není možné. Ukotvení také vyžaduje predikci budoucnosti s ohledem na agentovy současné akce. Požadavek prediktivní reprezentace je odklonem od čistě kauzálních teorií. Harnad (1990) navrhuje kauzální řešení postavené na kategoriální percepci senzoricky ukotvených signálů. Pokud ale nepoužijeme predikci, dostáváme se k problému zvanému *homunkulární* (Roy, 2005a), jelikož budeme muset postulovat systém uvnitř systému, který jej řídí.

V Royově pojetí jsou pro systém všechny informace o okolním světě víry s určitým procentem jistoty. Vytváří alternativu ke klasickému přístupu, jelikož systém nemůže svou znalost objektivně posoudit „vystoupením z kůže“ (Roy, 2005a). Analogová víra je brána jako distribuce nad všemi možnými pozorováními (jejich statistický průměr) vzhledem ke vstupní modalitě. Může sloužit jako element paměti, který kóduje historii pozorování v dané modalitě (kanále) a zároveň predikuje, co bude v kanále pozorováno (Roy, 2005a). Jedná se opět o přístup založený na dvou kooperujících subsystémech, které se navzájem ovlivňují. Systém věr je částečně nezávislou vrchní vrstvou, která může strukturovat agentův pohyb v prostředí a způsob akvizice nových konceptů. V modelu realizovaném v disertaci není tento subsystém zahrnut, ale v rozšířené verzi plánuji využít zmíněné principy. Využití neurčitosti (míra víry) se v této oblasti jeví jako slibný přístup. Roy rozpracovává Harnadův elementární princip do podoby, která zohledňuje míru neurčitosti umožňující systému reprezentovat neúplné informace z prostředí. Také v rovině interních procesů s reprezentacemi definuje podrobněji základní operace, pomocí kterých můžeme napodobit procesy myšlení. Důležitým aspektem je rovněž zakomponování prediktivních mechanismů do procesu ukotvení. Tento požadavek je v souladu s psychologickými teoriemi postulovanými již v počátcích kognitivní psychologie. Jedná se o percepčně akční cyklus, jak o něm hovoří Neisser (1964).

4.3 Modely ukotvení symbolů

Obečným modelem pro řešení ukotvení symbolů je hybridní systém, jak jej navrhuje Harnad ve svém článku (1990). Jedná se o propojení konekcionistického a symbolického přístupu. Proces ukotvení probíhá ve třech úrovních:

1. *Ikonizace - převod analogových signálů na ikonickou reprezentaci, která je analogovým ekvivalentem projekce vzdáleného objektu na agentovy senzory.*
2. *Diskriminace - posouzení, zda jsou vstupy stejné nebo rozdílné.*
3. *Identifikace - proces přiřazení jedinečné odpovědi.*

První dvě úrovně jsou subsymbolické, třetí slouží k ukotvení symbolů. Ikonická reprezentace se skládá z percepce jednoho typu objektů. Kategoriální reprezentace je dosaženo procesem diskriminace. Při ní systém bere v potaz pouze neměnné vlastnosti ikonické reprezentace. Taková kategoriální reprezentace je následně asociována s třídami symbolů (jmény), čímž dochází k ukotvení. Ikonizace a diskriminace jsou podprocesy, které jsou uskutečňovány pomocí neuronových sítí. Tvoří podklad pro následné pojmenování. Poté co identifikujeme symboly se subsymbolickými koncepty, je lépe použít klasickou počítačnou architekturu, pro následnou simulaci procesu myšlení. Právě díky propojení odlišných architektur, bývá nazýván Harnadův přístup hybridní.

Jelikož je jeho návrh pouze obecným řešením, které nebylo podrobněji empiricky testováno, předkládám na následujících stránkách ukázkou modelu ukotvení vycházející z obecného principu a realizované pomocí konkrétních modelů fungujících v konkrétních oblastech. Pomocí jejich analýzy můžeme ukázat obtížná místa procesu ukotvení symbolů.

4.3.1 Funkcionální model

Mayo (2003) se ve svém funkcionálním přístupu k procesu ukotvení snažil, aby agent (systém) vytvářel sémantické kategorie autonomně. Dalším požadavkem byla schopnost tvorby abstraktních jmen. Agent má nekonečně mnoho možností jakým způsobem kategorizovat kontinuální proud vstupních dat. Proto je třeba zavést nějaká omezení (zdůraznění) kategorizačních kritérií. Navrhuje funkční organizaci, jako způsob ukotvení symbolů. Kategorie je množina, která je závislá na povaze úlohy a slouží specifické funkci. Symboly jsou formovány tak, aby pomáhaly řešení úlohy. Právě povaha úlohy zdůrazní, jakým způsobem senzorická

data kategorizovat. Symboly mohou získávat různé významy, podle povahy úlohy. Množiny se mohou překrývat a měnit, a právě průniky vedou ke tvorbě abstraktních konceptů. Například slovo vítězství je spojeno průsečíkem různých druhů vítězství (tenis, šachy).

Funkcionální model zvládá abstraktní koncepty oproti Harnadovu hybridnímu modelu. Harnad tvrdí, že abstraktní reprezentace potřebují základní sémantické schopnosti, které musí být dopředu obsaženy v systému, přičemž v Mayově případě se jedná o funkcionální kritéria. Agent musí mít dopředu danou funkcionální sémantiku. V podstatě není možné, aby si agent tyto sémantické zdroje vytvořil sám.

Hodnocení: Jedná se o propojení klasického modelu ukotvení a pragmatického přístupu, vytvářející reprezentaci vzhledem k možnostem použití. Systém je podobný projektu Cospal (viz níže), jelikož klade důraz na enaktivní přístup. Posouvá problematiku ukotvení do další úrovně, ve které je třeba nalézt mechanismy identifikující funkci vhodnou pro řešení úlohy. Alternativně je možné systém postavit na základních „protomechanismech“ (reflexech či elementárních funkcích), pomocí nichž můžeme metodami učení vytvářet funkce složitější.

Mayův model je přínosný podrobnou analýzou procesu kategorizace. Jelikož člověk používá během myšlení rozdílné koncepty kategorizující totožný objekt dle rozdílných kritérií, je potřeba se zabývat otázkou, zda je informace o vhodně použitém způsobu kategorizace obsažena ve vstupech z prostředí, popřípadě zda musíme postulovat interní mechanismy sloužící v závislosti na typu úlohy ke tvorbě vhodných kategorií. Ve druhém případě tyto mechanismy nemohou být fixní, proto je nutné hledat funkční kritéria pro jejich dynamické užití. Přístup se v některých aspektech se blíží Royovu návrhu (2003) prediktivních mechanismů.

4.3.2 Intencionální model

V případě Sunova modelu (2000) se setkáváme s možností rozšířit systém o intencionální schopnosti. Autor vychází z heideggerovské filosofie a pracuje s pojmy *being-in-the-world* a *being-with-the-world*. Zdůrazňuje intencionální stavy a také schopnost agenta „být-ve-světě“, což má zajistit základní formu subjektivity. Agent je v interakci s objekty ve světě a zároveň reprezentuje sebe sama v tomto světě. Reprezentace nejsou tvořeny pouze perceptuálními vlastnostmi objektů, ale také možnostmi s těmito objekty operovat. Nazývá je teleologická pravidla. Na základě Heideggerových závěrů postuluje dvě úrovně učení. Ty jsou kvalitativně odlišné a každá kóduje uzavřenou část znalosti pro následné zpracování. Jsou mezi sebou komplementární.

První souvisí s agentovými akcemi v prostředí. V této úrovni učení nemá agent ještě vypracovanou explicitní reprezentaci. První úroveň bere v potaz pouze dva faktory: strukturu externího světa a vrozené predispozice, které ztotožňuje autor s intencionalitou prvního stupně a kvalifikuje ji jako „prereprezentační“ (implicitní). Což je počáteční základ pro tvorbu komplexnějších a složitějších forem intencionality. Opět se jedná o podobný přístup jako u funkcionálního modelu, tedy emergence konceptů z primitiv.

První úroveň probíhá na základě metody pokus-omyl. Druhá úroveň je vystavěna na základech první a nazývá se již konceptuální. První úroveň je intencionalní pro agenta ve světě objektů. Druhá úroveň vytváří z první úrovně možnosti a kombinace pro dosahování cílů. Podle Suna (2002) se jedná o dostačující metodu pro ukotvení konceptuálních reprezentací funkcionální cestou. Systém tedy obsahuje dvě vrstvy, které jsou propojeny „zdola-nahoru“. Sun poznamenává, že probíhají i procesy „shora-dolů“, přičemž agent testuje validitu reprezentací prvního stupně, jejich použitelnost v plánech, jejich možnosti, generalizace, upřesnění apod.

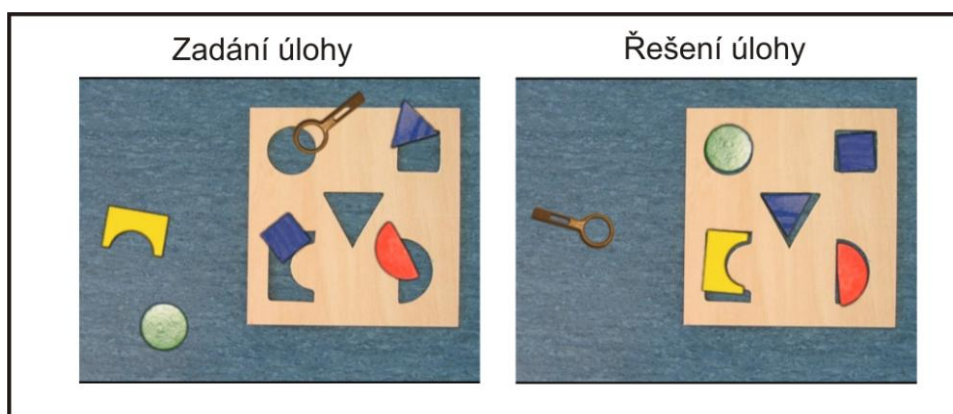
Konkrétní architektura postavená na těchto předpokladech se jmenuje Clarion. Jedná se o 4 vrstvou neuronovou síť, přičemž tři vrstvy vytvářejí podklad pro akce a čtvrtá vrstva vytváří explicitní pravidla, která převádí do symbolické úrovně. Čtvrtá úroveň neuronové sítě je totožná s druhou úrovní učení v Sunově teoretickém návrhu. Je zodpovědná za procesy „shora dolů“. Koncepty jsou kontextově závislé a jsou funkcionální pro dosahování cílů agentem. Jsou částí množiny rolí, které se agent učí při interakci se světem (Sun, 2000). Zdůrazňuje funkcionální hledisko konceptu a snaží se ukázat, že jsou tvořeny zkušeností a nejsou definovány a priori.

Hodnocení: Sun příliš nevysvětluje intencionalitu prvního řádu, která vzniká z přímých interakcí agenta s prostředím. Výhodné by bylo použití principu samoorganizace, pomocí něhož lze efektivně reprezentovat strukturu prostředí bez předchozí znalosti. Sun problém řeší dlouhodobým pobytem agenta v prostředí, přičemž pro tvorbu prvotních konceptů používá metody genetických algoritmů. Takové řešení vyžaduje designérský zásah do systému, neboť je potřeba ohodnotit funkci *fitness*, což je v rozporu autonomní tvorby konceptů agentem.

Pozitivně lze hodnotit návrh architektury založené na dvou stupních, které jsou navzájem propojené a hledají vzájemné mapování. Podobný princip můžeme nalézt i u projektu Cospal (viz níže), ve formě propojení senzorického a motorického systému. Právě součinnost dvou subsystémů pracujících odlišně, ale navzájem propojených, je důležitým inspiračním zdrojem při návrhu vlastního modelu.

4.3.3 Aplikovaný enaktivní model

Systémy založené na ukotvení symbolů se vyskytují také v oblasti aplikovaných výzkumů. Jedná se o konkrétní modely, schopné fungovat v reálném prostředí a plnit definovanou úlohu. Ukázkou může být projekt Cospal (Granlund, 2005), vyvíjený ve spolupráci s ČVUT v Praze. Je založen na enaktivním přístupu (Varela, 1991), kladoucím důraz na propojení systému s prostředím. Systém Cospal využívá motorických akcí k ukotvení perceptuálních informací. Úlohou modelu je napodobit schopnost rozlišit základní geometrické objekty (trojúhelník, kruh, obdélník, trojúhelník) a umístit je do stejně tvarovaných otvorů. Tuto úlohu dokáže vyřešit dítě během prvního roku života..



Obr. 4 Ukázka fungování systému COSPAL. Cílem je umístit geometrické tvary do správných otvorů.

Autoři vycházejí z principu konceptuálního učení. Systémy musí zachycovat informace z externího světa pomocí exploračního učení, protože komplexní propojení perceptů a konceptuálních rámců nemůže být explicitně specifikováno. Častou chybou minulých systémů vidění byl názor, že perceptuální produkce deskriptce objektu nebo scény, založené na abstrakci objektů. Problematická je mezera mezi abstraktní deskripcí a akcí. Příliš mnoho důležitých kontextuálních kvalifikátorů, prostorové a časové informace, nutné pro jednání bylo ztraceno díky procesu abstrakce vedoucí k deskripci.

Abychom překonali mezeru mezi perceptuální a akční strukturou, potřebujeme rozdělit proces na menší kroky. Centrální mechanismus je založen na percepčně-akčním cyklu. Tak tomu je ale pouze u trénovaného systému. Ve fázi učení nastává situace, ve které akce předchází percepci. Má to výhodu, jelikož stavový prostor akcí je méně komplexní než perceptuální. Systém je v prvotních fázích učení řízen částečně náhodnými akcemi. Použití pseudonáhodných mechanismů v procesu učení je biologicky plausibilní, jelikož mnoho primitivních živočichů

obsahuje generátory šumu. Bylo prokázáno, že šумы a spontánní aktivita jsou důležité pro organizaci a koordinované chování organismu. Systém je tak organizován z motorické části, nikoli senzorické.

Možnou variantou je tvorba systému, které budou mít tyto pseudonáhodné mechanismy ukotveny na základě předchozího učení, tzn. bude vytvářet „předobjektové“ a „předkategoriální“ mechanismy, založené na velmi zjednodušené tvorbě kategorií (*protokategorie* a *protokoncepty*). Obdobné principy se objevují v souvislosti s přenosem ukotvení (viz níže). Řízení systému pomocí zpětnovazebných signálů má pět důležitých funkcí:

1. *Redukuje komplexitu asociačního prostoru, který tvoří modely vzájemných vztahů perceptů a stavů či akcí.*
2. *Separuje jednotlivé modalities, takže mohou být reprezentovány odděleně.*
3. *Identifikuje pouze percepty, které se vztahují k dané akční modalitě.*
4. *Generuje výstupy na základě akcí. Bez výstupu nedokáže anonymní mód provést žádnou akci v okolním prostředí.*
5. *Kontroluje asociace ze strany akcí, což znamená větší kontinuitu, jednoduchost a uzavřenost.*

Systém kontroly ze strany akce připomíná *top-down* způsob, ve kterém abstraktní struktura kontroluje složitější percepční doménu. Na rozdíl od klasické symbolické architektury je abstraktní úroveň nonsymbolická a přímo zodpovědná za jednání. Autoři předpokládají, že symbolické reprezentace vzniknou následně pomocí principů emergence a jsou organizovány na základě akčních nebo stavových reprezentací. Symbolická reprezentace a manipulace by měla být nazírána jako doména pro efektivní zpracování relativně invariantních konceptů, bez nadbytečných prostorových, časových a kontextuálních kvalifikátorů. Invariantní formát symbolických reprezentací dělá manipulaci a komunikaci efektivnější a obecně aplikovatelnou.

Právě jednoduchý převod z perceptuální do symbolické úrovně je slabinou současných systémů ukotvování. Jestliže chceme nalézt způsob, jakým tyto systémy spolupracují (případně paralelně pracují), potřebujeme vytvořit daleko komplexnější a sofistikovanější mechanismy zajišťující spolupráci. Jelikož systém *percepce-akce* dokáže operovat pouze v úrovni tady a teď, je třeba symbolické struktury, která umožní systému pracovat v prostoru a čase výkonnějším způsobem, což je umožněno generalizací.

Hodnocení: Systém Cospal je obtížné hodnotit, jelikož byl v době vzniku disertační práce ve fázi vývoje a testování. Jeho funkčnost a efektivitu proto analyzovat nelze. V rovině návrhu se jedná o inspirativní teorii, která se snaží v praxi aplikovat principy extrémnějších variant kon-

ceptualismu, přikládajíc symbolické vrstvě pouze limitované využití, v případě nedostatečnosti konceptuálního aparátu. Systém je v konceptuální rovině velmi robustní, jelikož obsahuje dvě primární reprezentace, vizuální a motorickou, jejichž vzájemné propojení umožňuje reprezentovat okolní prostředí nejen z hlediska vnímání, ale také způsobu interakce s prostředím pomocí efektorů. Robustnost reprezentací zajišťuje větší odolnost proti chybám, jelikož primární oblasti reprezentují odlišné vlastnosti téhož objektu či jevu externího prostředí. Rozšíření o motorickou vrstvu je plánováno i v modelu realizovaném v disertační práci.

4.3.4 Model přenosu ukotvení

Rozšíření a doplnění hypotézy ukotvení symbolů následovalo v letech 2000-2002, opírajíc se o články Harnada a Cangelosiho (2001) a Cangelosiho (2000,2002). V návaznosti na Harnadovy práce navrhuje dva směry dalšího vývoje modelů ukotvení. Jedná se o *sensomotorickou dřinu* a *symbolickou krádež*. V případě *symbolické krádeže* hovoří o mechanismu transformace kategoriální percepce do ukotvených elementárních označení. Následně slouží proces *přenosu ukotvení* k akvizici nových symbolů na základě kombinace již ukotvených, přičemž může být implementován do neuronových sítí. Ty umí rozlišovat mezi množinami stimulů, extrahovat podobnosti a kategorizovat. Sítě také vykazují základy kategoriální percepce, tzn. člen třídy je více blízký (kompresi vzdálenosti uvnitř kategorie), a členové jiné kategorie jsou více odlišní (expanze vzdáleností mezi kategoriemi).

Takto budovaný systém je pokusem o hierarchickou stavbu ze základních atomů do vyšších celků, přičemž atomy jsou dostatečně elementární (a zde také ukotvené), aby se staly univerzálními prvky pro následnou tvorbu komplexních reprezentací a operací s nimi. Ve své podstatě se jedná o kompozicionalitu (schopnost skládat prvky), což je základní vlastnost symbolického systému. *Symbolická krádež* je na ní postavena. Ve svém původním článku (1990) Harnad ještě považuje tuto schopnost za vlastní pouze symbolickému systému (cf. Touretzky, Hinton, 1985).

V návaznosti na teoretické předpoklady provedl Cangelosi (2002) simulaci procesů kategoriální percepce, akvizice ukotvených jmen a učení symbolů vyššího řádu z již ukotvených. Vytvořil neuronovou síť, která měla jako vstupy umělou retinu a lingvistický vstup. Na vstup prezentoval základní geometrické útvary a jejich názvy. Proces tréninku byl následující:

- 1) *Síť je trénovaná příklady geometrických útvarů až jsou vytvořeny jejich prototypy.*
- 2) *Poté jsou asociovány tvary se jmény, čímž dochází k ukotvení symbolů v percepci.*
- 3) *Dochází k učení se kombinacím ukotvených jmen tvarů s novými arbitrárními jmény (např. symetrický x asymetrický). Je použito imitační učení kombinací jmen (např. čtverec je symetrický). Tak dochází k přenosu ukotvení z již ukotvených jmen (čtverec) na nové jména (symetrický).*

Původ a vývoj jazyka je v modelu vysvětlován hybridním symbolicko/senzomotorickým systémem. Nejdříve přichází proces kategorizace - *senzomotorická dřina* a následně je znalost použita k symbolickým manipulacím - *symbolická krádež*.

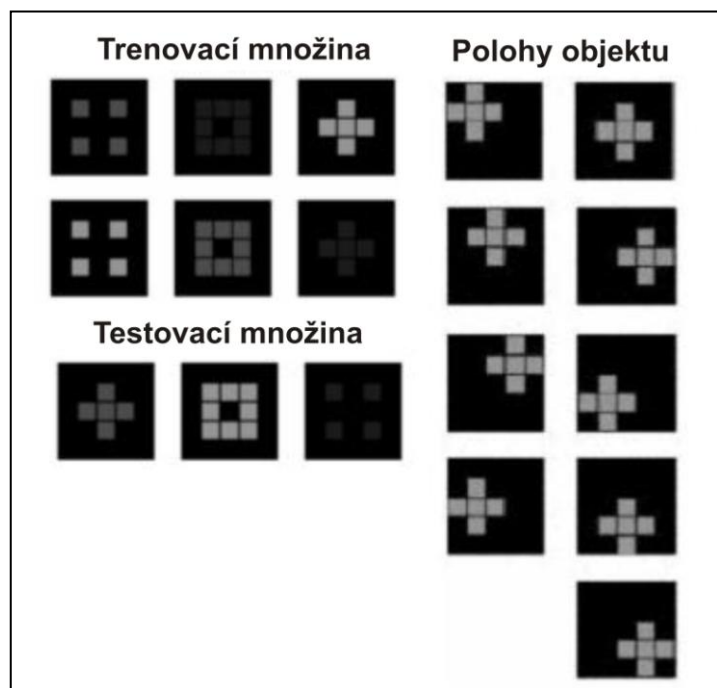
Hodnocení: V Cangelosiho přístupu nacházíme snahu o rozlišení způsobu řešení symbol groundingu. Pokouší se nalézt odpověď na otázku, který subsystém by měl být během groundingu převažující. Zda použít *senzomotorickou dřinu*, tedy velmi robustní mechanismy práce se senzorickými vstupy a předzpracování na mnoha úrovních (což je výpočetně náročné), nebo systém, ve kterém jsou ukotvené symbolické atomy následně využívány symbolickým systémem, jehož činnost je daleko méně náročná, transparentní a teoreticky zakotvena. Cangelosi se v podstatě snaží určit, ve které části převodu senzorických vstupů máme začít používat výhody symbolického systému. Ve svém modelu navrhuje větší využití symbolického systému a ukotvení v elementární úrovni.

4.3.5 Rozšířený model přenosu ukotvení

Cangelosi se věnoval problematice přenosu ukotvení i ve svých dalších pracích. V realizovaném modelu (Riga, Cangelosi a Greco 2004) je použito *učení bez učitele*, pomocí něhož dochází k formaci konceptu, a následně je kategorii pomocí *učení s učitelem* přiděleno jméno. Základní ukotvené symboly se týkají barev a tvaru prezentovaných objektů. Po ukotvení mohou být využity vznikající *kategoriální primitiva* ke tvorbě deskripcí nových kategorií. Princip *symbolická krádeže* je použit pro *grounding transfer* (přenos ukotvení). Nové symboly jsou definovány bez nutnosti přímého kontaktu s referentem (externím prostředím). Kombinací výše zmíněných způsobů umožňuje vznik perceptuálně kotvených významů, které mohou v symbolické rovině zajistit přenos ukotvení. Lze například vytvořit slovo *zebra* ze slov *kůň* a *pruhy* (Cangelosi, 2006).

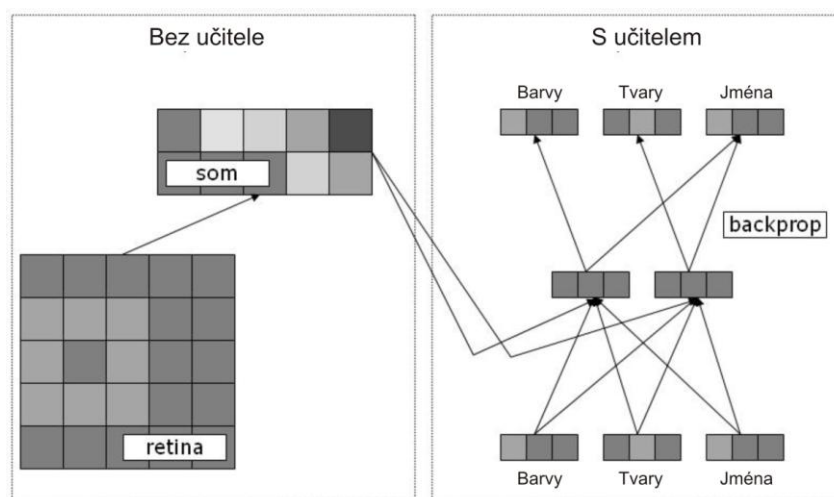
Podívejme se na model trochu blíže. V první fázi jsou prezentovány obrázky objektů různých tvarů a barev (viz Obr. 5). Systém se učí diskriminovat a vytváří mapu vlastností pomocí *učení*

ní bez učitele. V následné fázi jsou prezentovány symboly společně s obrázky. Každý obraz je asociován se symboly tvořenými arbitrární skupinou bitů reprezentující označení vlastnosti barev a tvarů. Síť musí reprodukovat stejné symboly na výstupu. Během fáze učení jsou symboly přímo asociovány se senzorickými reprezentacemi z první fáze.



Obr. 5 Vizuální vstupy modelu ukotvení. Vpravo ukázka možných poloh objektu (Riga, 2004)

Ve třetí fázi je trénovací vstup pouze symbolický. Lze například prezentovat obrazy červeného a modrého čtverce nebo zeleného a červeného trojúhelníku, ale nikoliv zeleného čtverce a modrého trojúhelníku. Následně prezentujeme symbolický popis zeleného čtverce (Čtverec & Zelený = SOD) v průběhu tréninkové fáze. Během testové fáze pak můžeme projikovat na retinu nové a neznámé obrázky, abychom potvrdili, zda došlo k přenosu ukotvení. Jestli k němu došlo, měl by se na výstupní vrstvě objevit správné označení objektů. (Riga, Cangelosi a Greco 2004). Nové symboly jsou nepřímě ukotveny, jelikož odkazují pouze na kombinace barev a tvarů, které nebyly systému nikdy prezentovány. Právě tento proces se nazývá *přenos ukotvení*, jelikož vytváříme nová označení na základě již naučených senzorických primitiv, které kombinujeme do nových sekvencí.



Obr. 6 Schéma modelu ukotvení. Vlevo vizuální část, vpravo dole auditivní, nahoře společný systém.

Architekturu systému můžeme vidět na Obr. 6. Obsahuje dva moduly a jednu umělou retinu. Prvním modulem je dvouvrstvá síť SOM, druhým pak MLP – vícevrstvý perceptron. Obraz z retiny je zpracováván SOM modulem. Druhý modul přijímá distribuované reprezentace prvního modulu a také symbolický vstup ve formě arbitrárních symbolických sekvencí (Riga, Cangelosi a Greco 2004). Během zpracování jsou použity dva algoritmy učení. Prvním je samoorganizace, druhý propojí analogové reprezentace emergující z první sítě do diskretních symbolických stimulů pomocí algoritmu *backpropagation* (zpětné šíření chyby). Vrstva SOM pracuje na základě algoritmu *vektorové kvantizace*, který vytváří mapování z multidimenzionálního prostoru stimulu do dvojrozměrné matrice která dokáže zachytit podobnosti. První modul tedy autonomně vytváří aktivační matici, ve které je zachycen vnitřní řád množiny stimulů, tzn. podobnost tvarů objektů a jejich barev při jejich různém prostorovém umístění. Druhý modul dostává spolu se symbolickým vstupem také aktivační matici prvního modulu. Vizuální, symbolický vstup a výstup jsou mapovány pomocí *supervised učení*, které je kontrolováno designérem.

Model byl odzkoušen na 30 sítích s různým nastavením (rychlost učení=0,2 v první a 0,5 v druhé a třetí fázi; momentum 0,9). Způsob prezentace stimulů během *učení bez učitele* byl náhodný, během *učení s učitelem* naopak sekvenční. V testovací fázi byla úspěšnost „přenosu ukotvení“ 90%. Předchozí model (Cangelosi, 2002) měl stejnou úspěšnost, ale neobsahoval SOM vrstvu. Oproti starším přístupům (Schyns, 1991), obsahuje model princip *přenosu ukotvení* a je vhodný pro rozšíření na složitější úlohy. V současné verzi používal model pouze logickou spojku AND (barva AND tvar), přičemž v další verzi by měl obsahovat více možností kompozicionality a produktivity (Riga, Cangelosi a Greco 2004).

Hodnocení: Zmíněný přístup proces *přenosu ukotvení* příliš zjednodušuje. Přenos a kombinace perceptuálních vlastností a jejich univerzální použití v symbolické rovině nejsou automaticky zajištěny. Například bez perceptuální znalosti *zebry* nelze v symbolické rovině určit jakým způsobem aplikovat pruhy na koně, jelikož možností propojení těchto označení do perceptuálního vjemu je mnoho (absurdně například pruhy, na kterých je namalován kůň). V systému nelze pouhou kombinací vlastností objektů definovat (ukotvit) složený objekt. Uvedené námitky, přestože jsou opodstatněné, považujeme při současném stavu modelů ukotvení předčasné. Již samotný pokus o elementární způsob kompozicionality objektů a vlastností je podstatným krokem vpřed. V rozšířené verzi vlastního modelu se touto otázkou zabývám, přičemž používám odlišný postup.

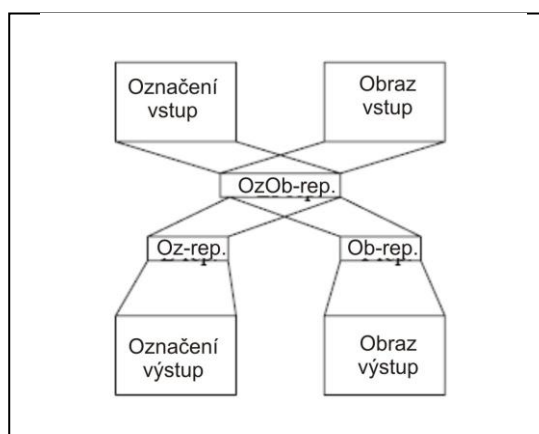
4.3.6 Konekcionistický model rozpoznávání objektů

Podobný model, obsahující retinu a auditivní vstup, nacházíme u Dorffnera (1996). Ve své práci „A Connectionist Model of Categorization and Grounded Word Learning“ provedl pokus s neuronovou sítí, přičemž použil dvě vstupní vrstvy, jednu pro perceptuální (obrazový) vstup druhou pro lexikální (auditivní) vstup, které mapoval do společné vrstvy (Dorffner, 1996). Popsaná architektura může sloužit (v modelu realizovaném v disertační práci její modifikovaná varianta slouží) jako základ pro tvorbu multimodálních reprezentací.

Autor vychází z několika teoretických předpokladů, které jsou blízké perceptuálním teoriím kognice. Slova jsou arbitrární, protože jejich tvar nijak nekoresponduje s objektem jež zastupují a nijak nesouvisí s významem. Proto učení se slovům musí probíhat na jiném principu než běžné učení se na základě podobnosti stimulů (Dorffner, 1996). Dále zdůrazňuje, že učení se slov a jejich významů probíhá v „zašuměných podmínkách“ a proto musí být mechanismus učení velmi robustní. Pro jeho pochopení využívá znalosti o lidské kategorizaci, tvorbě prototypů, taxonomii apod. (Dorffner, 1996). Při samotné realizaci modelu bylo nutné provést následující zjednodušení:

- a) *Slova a jejich porozumění nevyžaduje sofistikovanou fonologii a morfologii, jedná se o pouhé fyzické jednotky, které jsou na vstupu brány jako vzorce stimulů bez sekvenční struktury (pořadí slov či pochopení věty).*
- b) *aspekty slov, jdoucí za jejich jednoduchý význam je pomíjen (dvojznačnost, kontextuálnost).*

c) význam je získáván kategorizací stimulů a jejich přiřazení jednotlivým slovům. Což omezuje možnost použití pouze na podstatná jména. Vynechána musí však být také podstatná jména označující abstraktní entity. (Dorffner, 1996)

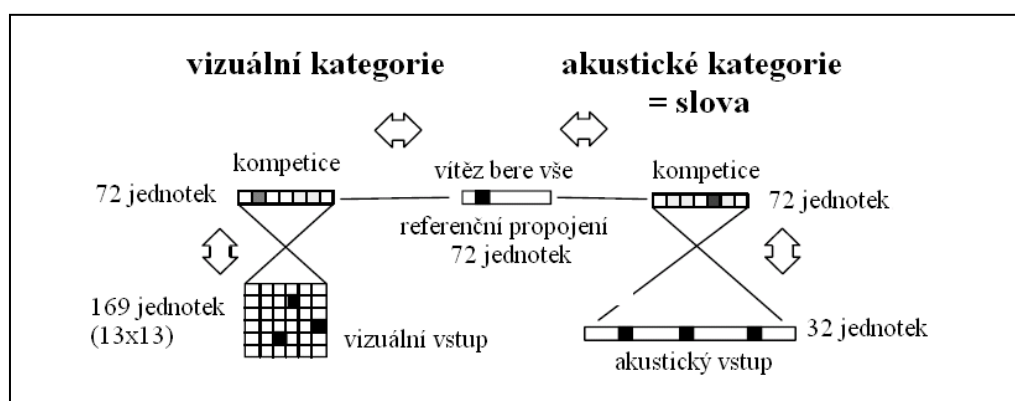


Obr. 7 Základní návrh Dorffnerova modelu. Obraz a označení jsou společně reprezentovány v jedné vrstvě (OzOb-reprezentace). Celý model je postaven na neuronových sítích (Dorffner, 1996).

Přestože nebyla provedena přesná psycholingvistická analýza, jsou považovány takové typy modelů, jako první kroky konekcionismu směrem k ukotvení významu (Dorffner, 1996). Z hlediska architektury se jedná o systém se dvěma oddělenými vstupy (viz Obr. 7). Stejně jako v podobně zaměřených pracích (Cottrell, Bartell, Haupt, 1990; Plunkett, 1993) přijímá systém vizuální a slovní vstupy, a následně je zpracovává pomocí kategorizace a mapování (přiřazování) mezi kategoriemi. Kategorizační model je založen na *kompetitivním učení*. V principu se jedná o hledání konceptu, čili jakéhosi centra, kolem kterého je daná kategorie vymezena. V psychologii tomu odpovídá pojem prototyp. Z hlediska neuronových sítí využíváme principy samoorganizace a učení bez učitele. *Kompetitivní učení* je založeno na postupném vývoji kategorizace během učení, založené na inhibici okolí, či metafoře *bohatý se stává bohatším*, což principiálně odpovídá architektuře ART3 (Carpenter, Grossberg, 1993). Podobně je tomu z hlediska vítězného neuronu. Pokud není dosažena určitá hodnota „vítěze“, je vrstva resetována a trénování pokračuje dále (Dorffner, 1996).

Architektura ART

Grossberg (Carpenter, Grossberg, 1986) používá neuronových sítí ke tvorbě konceptů a rozšiřuje architekturu o přídavné mechanismy či funkce, které jsou inkorporovány do sítě a zlepšují její schopnosti v procesu kategorizace. Jedná se o možnosti učení se konceptům při proměnlivém opakování prezentace stimulů. Klíčová je schopnost být citlivý k této změně a predikovat ji. Systém obsahuje *top down* mechanismus, který dokáže modifikovat způsob kategorizace na základě *backpropagation* učení a také na základě odhadu chyby. Je ovládán pomocí parametru *vigilance*, jehož modifikace (v pozdější fázi jí provádí sám systém pomocí predikce chyby), dokáže měnit způsob kategorizace, ve smyslu tvorby širokých abstraktních či velmi úzkých konkrétních kategorií (v extrémním případě je každý příklad kategorií), na základě změny tohoto parametru. Variací *vigilance* pak systém může rozpoznat abstraktní i konkrétní kategorie.

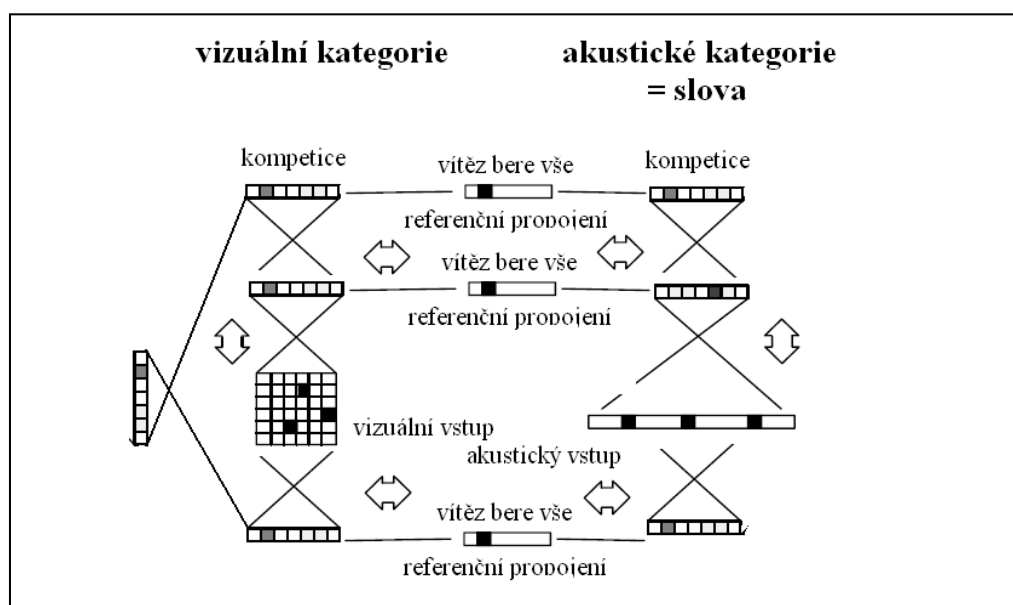


Obr. 8 Základní schéma Dorffnerova modelu. Vizuální a akustické vstupy jsou propojeny pomocí referenční vrstvy, která vytváří jednoznačné kategorie (Dorffner, 1996)

Systém obsahuje kromě dvou kategorizačních modulů také referenční vrstvu sloužící k propojení kategorizačních vrstev do jedné. Propojení v referenční vrstvě zajišťuje ukotvení arbitrárních slov pomocí obrazu. Vrstva pracuje na principu WTA (vítěz bere všechno) a posiluje pouze jeden neuron, což maximalizuje asociační vazbu.

Autor na základě stávajícího modelu navrhuje další možnosti vývoje. Rozšířený kategorizační modul může obsahovat místo jedné vrstvy celou sadu propojených kategorizačních vrstev, což umožní tvorbu více různých kategorií pro jeden stimul. Skrze takové zapojení se dostanou

již naučené kategorie opět na vstup, čímž dochází ke tvorbě abstraktních kategorií. V konečném důsledku znamená takové propojení schopnost učení synonym a homonym.



Obr. 9 Návrh rozšíření Dorffnerova modelu. Systém je schopen kategorizovat ve více úrovních, popřípadě podle rozdílných kritérií.

Hodnocení: Dorffnerův model byl v oblasti použité architektury velkým inspiračním zdrojem pro systém realizovaný v disertační práci. Obsahuje dvě primární reprezentační vrstvy, které kategorizují zvláště obrazovou a auditivní reprezentaci, což je v korespondenci s perceptuálními teoriemi kognice. Následně jsou jednotlivé primární reprezentace integrovány a vytvářejí multimodální reprezentace daných objektů a jejich označení. Oproti Cangelosiho modelu (2004) je celý systém postaven na principech samoorganizace, což zaručuje schopnost systému autonomně tvořit kategorie bez zásahu designéra, tzn. je v souladu s *učitelským přístupem*. Ve svém modelu rozvíjím Dorffnerovy návrhy, přičemž zkoumanou oblastí je vnímání prostoru. Použitím rekursivních sítí (RecSOM) navíc mohu reprezentovat dynamické děje, čímž posouvám možnosti modelu např. do oblasti reprezentace sloves či dynamických změn.

4.3.7 Model ukotvení prostorových vztahů

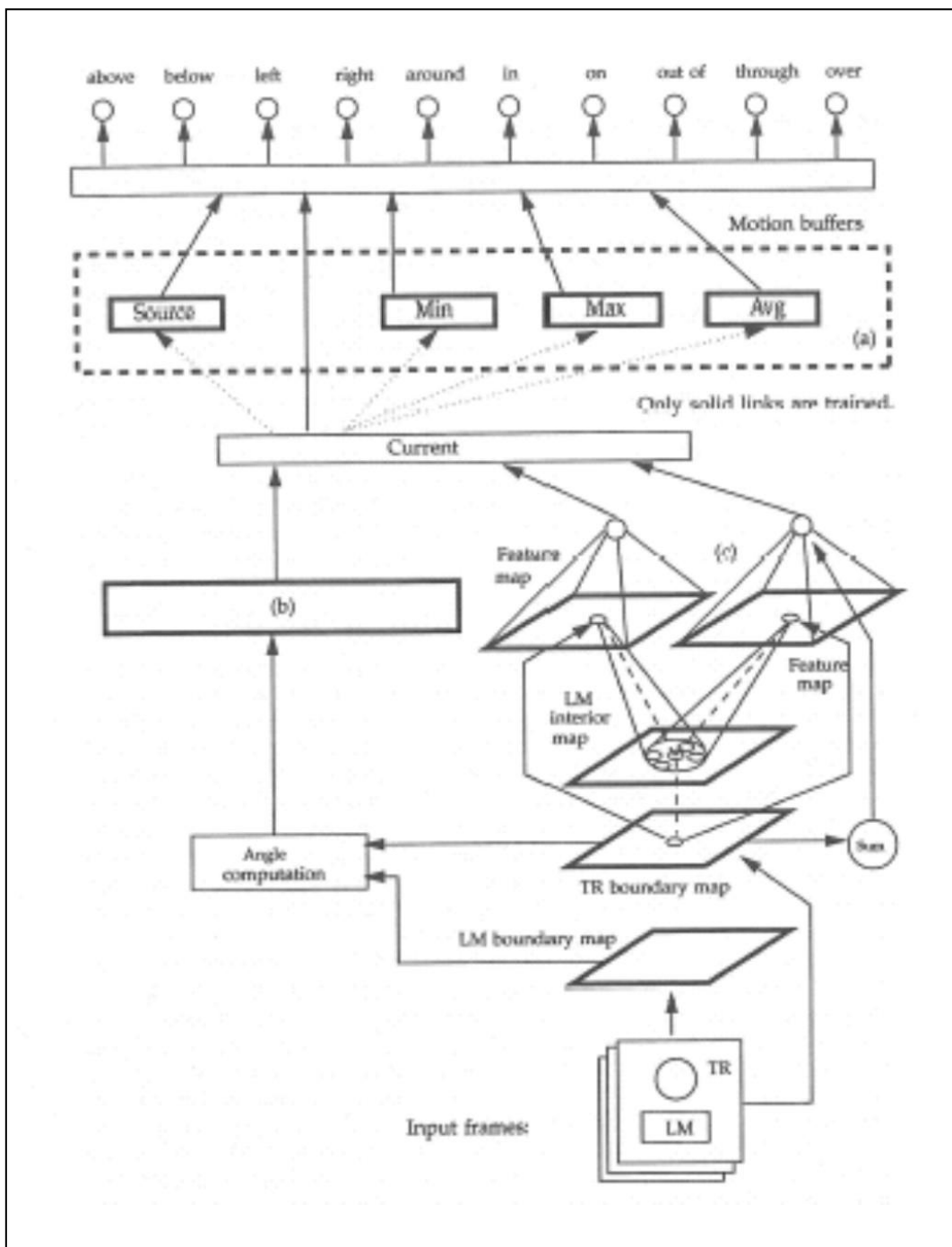
Jiným příkladem konkrétní aplikace v oblasti *symbol grounding* je Regierova *perceptuálně ukotvená sémantika*. Regier (1992) trénoval neuronové sítě v označování dvourozměrných scén obsahujících základnu (landmark) a objekt (trajektor), u něhož určujeme prostorový vztah vzhledem k základně (*na, pod, nad apod.*). Autor nazývá svůj systém *preadaptovaným strukturním nástrojem*, který plní podobné úkoly jako malé dítě v průběhu učení se lexikální sémantice prostorových vztahů. Zdůrazňuje, že jeho model je založen na poznatcích z biologie, psychofyziky a lingvistiky ale nikoliv z neurovědy, protože se nesnaží přesně kopírovat funkci neurálních drah.

Zdůrazňuje fundamentálnost prostorových vztahů, které mohou působit jako metafora v ostatních kognitivních systémech, což by znamenalo, že ovlivňují i jazykovou úroveň. Přesto v polemice o vztahu mezi nelingvistickými a lingvistickými kategoriemi a jejich vzájemné hierarchii, tedy otázce, zda je jeden systém nadřazen druhému, hovoří ve prospěch privilegovanosti lingvistických kategorií. Autor tvrdí, že děti tvoří sémantické koncepty na základě struktury jazyka (Regier, 1996). Jedná se o důraz na jazykový kód sloužící lépe k identifikaci prostorových konceptů, než je tomu u reprezentace vizuálních kategorií, které se mohou navzájem překrývat a být nejednoznačné. Přestože hovoří o privilegovanosti lingvistické (symbolické) reprezentace, odmítá formální přístupy založené na jediném symbolickém reprezentačním systému. Lingvistická reprezentace pouze lépe strukturuje koncepty, které jsou nutnou součástí systému a jsou perceptuální povahy.

Při tvorbě systémů používá architekturu *omezeného konekcionismu* (Feldman, Fanty, Goddard, 1988), což je varianta neuronových sítí, u kterých je způsob propojení vytvořen z ohledem na snadnou analyzovatelnost. V těchto sítích jsou jednotky interpretovány spíše v úrovni symbolů než subsymbolicky a jsou daleko přehlednější pro modeláře. Zjednodušení modelu bývá trnem v oku kritice, jelikož návrh obsahuje příliš mnoho parametrů nastavených designérem a splňuje kritéria pro *designérský přístup* (Ziemke, 1999).

Strukturovaný konekcionismus znamená tvorbu sítě, ve které jsou vlastnosti definovány jedním neuronem tzn. lokalisticky. Autor vidí v jednoduchosti pouze pragmatické důvody. Cílem je vytvořit systém propojených neurálních modulů, u kterých můžeme přesně určit k jaké transformaci slouží a vzájemnou součinností vytvářet konečný systém. Oproti klasickému konekcionismu, u kterého je způsob práce sítě v jednotlivých vrstvách „lidsky nečitelný“, tak získáváme designérskou výhodu. Na druhou stranu model obsahuje zjednodušení působící problémy v oblasti reprezentace významu.

Podívejme se na jeho činnost podrobněji. Na vstupní vrstvu jsou prezentovány statické obrázky či krátké sekvence, ve kterých se vyskytují dva objekty v různých prostorových konstelacích, přičemž systém je učen jejich názvům.



Obr. 10 Schéma modelu ukotvení prostorových vztahů (Regier, 1996)

Vrstvy ohraničené přerušovanou čarou jsou vybudované na základě metod *omezeného a strukturovaného konekcionismu*. Naproti tomu *vrstva current* a také vrchní vrstva obsahující názvy prostorových označení jsou nestrukturované, založené na principu *vícevrstvého perceptronu (MLP)*. Autor rozlišuje vrstvy podle toho, zda je do nich designérem implementovaná specifická funkce nebo dojde k jejich strukturaci až během tréninku. Strukturované a omezené vrstvy bere jako zabudované do systému, plnící specifickou funkci.

Výpočet polohy objektů provádí specializované *théta neurony* detekující orientaci dvou objektů. Jedná se *proximální orientaci*, což je nejbližší vzdálenost mezi objekty a *orientaci na centrum masy*, tedy vzdálenost mezi centry objektů. Celková funkce vypadá následovně.

Vstupní snímek je reprezentován pomocí mapy *okraje* trajektoru a základny. Následně specializované *théta neurony* provádějí výpočet uhlů a orientaci. Výsledek je posílán o úroveň výše do *current vrstvy*. Ta zpracovává sumarizovaný vážený výstup z množiny *théta neuronů*. Při tréninku je zpětně ovlivňováno jejich nastavení, což umožňuje modifikovat citlivost na polohu objektů.

Kromě polohy objektu je nutné identifikovat jejich vzájemnou interakci. K tomu se používá *mapa okraje* trajektoru a *mapu interiéru* základny, získanou z *mapy okrajů*. Vstupy přes *mapy vlastností* ovlivňují *hlavní neurony* detekující kontakt nebo průnik. Výstupy z *hlavní vrstvy* jsou posílány do *current vrstvy*, stejně jako výstupy z *théta neuronů*. Váhy jsou opět modifikovány pomocí *backpropagation* učení.

V *current vrstvě* dochází k integraci výsledků činnosti modulů vedoucích k označení statických prostorových vztahů (vlevo, vpravo apod.), tzn. uchovává reprezentaci vstupního snímku. Všechny vstupy musí projít přes *current vrstvu*, která je hrdlem integrujícím statické reprezentace. Neobsahuje žádnou formu paměti pro předchozí snímky, tzn. obsahuje vždy jen informaci z aktuálního snímku.

Dynamické scény jsou zpracovávány v modulech, pro které je *current vrstva* vstupem. Pomocí metody *zdroj-cesta-cíl* jsou odvozeny dynamické prostorové vztahy (skrz, kolem atd.). První snímek, který je transformován na reprezentaci v *current vrstvě* je uložen do *zásobníku zdroje* (source buffer). Další zpracování probíhá v *zásobníku cesty* vytvářejícím statickou reprezentaci pohybu. Reprezentace posledního snímku sekvence zůstává v *current vrstvě* a stává se cílem. Následně jsou vrstvy společně zpracovány v posledním zásobníku obsahujícím označení prostorových vztahů.

Mezi nevýhody modelu patří neschopnost naučit se rozpoznat kontakt či průnik objektů, pokud se nevyskytují v posledním snímku sekvence. To je způsobeno *backpropagation* učením. Dalším omezením je insensitivita k časovým údajům pohybu trajektoru na jeho cestě. Necitlivost znemožňuje zjištění aktuální polohy během pohybu po cestě, popřípadě sekvence těchto

údajů. Autor zdůvodňuje, že necitlivost k času je jeho záměrem a umožňuje jednoduchost architektury modelu.

Přestože je tento model postaven na *designérském principu*, hovoří autor o biologické inspiraci a biologické plausibilitě některých částí svého systému. Jelikož existují buňky vizuální kůry, citlivé na prostorovou orientaci stimulů, můžeme hovořit o neuroanatomických strukturách podobných *théta neuronu*. Na druhou stranu vizuální neurony reagují na konkrétní informace, na rozdíl od *théta neuronů*, které pracují s abstraktnější informací, jako je *centrum masy*. Autor (Regier, 1996) oponuje výzkumy Heyda, Peterhanse a Baumgartera (1984) prokazující citlivost neuronů v area 18 na orientaci iluzorních kontur, tedy kontur, které lidský pozorovatel vnímá, ale neodpovídají fyzickým rozdílům světelnosti obrazu (Kanisza, 1979). Model je v této úrovni částečně biologicky plausibilní.

Hodnocení: Přestože je Regierův model unikátním příkladem systému schopného kategorizovat prostorové vztahy na základě perceptuálních informací, je způsob jeho tvorby a použité mechanismy v rozporu s principy ukotvení symbolů. Například způsob učení na základě *backpropagation* algoritmu (učení s učitelem), v sobě obsahuje skrytý předpoklad řízení architektury „shora-dolů“. Metoda učení založená na zpětném šíření chyb není psychologicky plausibilní a předpokládá zásah designéra do interních mechanismů systému.

Díky tomu, že modul označení kategorií je umístěn v hierarchicky nejvyšší části architektury, je zjednodušen způsob akvizice prostorových označení. Také fixní počet prostorových označení ve výstupní vrstvě je apriorní a umožňuje tvorbu pouze lokalistických reprezentací.

Navíc autor neřeší získávání auditivní informace a její společnou integraci s vizuální senzorickou informací. Proto jsem se rozhodl (na podkladě zmíněných poznatků) pro tvorbu systému, který by k procesu ukotvení přistupoval fundamentálněji a zohledňoval auditivní i vizuální informace.

Při konstrukci modelu jsou také opomíjeny teorie vývoje neverbálních prostorových konceptů a způsob kterým ovlivňují jazykovou kategorizaci prostoru (Piaget, Inhelder, 1967, Clark, 1977, Huttenlocher, Smiley a Ratner, 1983, Mandler, 1992), což činí z modelu pouze omezené řešení akvizice prostorové sémantiky. Proto se nabízí možnost zkoumat, zda prelingvistický vývoj neobohatí prostorové vnímání o schopnosti (např. rozpoznat průnik), ještě před jazykovou akvizicí. Právě takové možnosti nabízí model prezentovaný v disertační práci.

5 Vnímání prostoru

V minulých kapitolách jsem podrobně popsal teoretická východiska, která se týkala reprezentace významů a otázek s tím spojených. Jelikož jsem si jako oblast aplikace kognitivní sémantiky vybral prostorové vztahy, je potřeba prezentovat současné přístupy, které jsou v této oblasti používány. Přestože se jedná o psychologickou práci, snažil jsem se integrovat poznatky nejen z psychologie, ale i z oblasti neurobiologie, lingvistiky a kognitivního modelování, abych vysvětlil některé principy objevující se v navrženém řešení.

5.1 Neurobiologické poznatky

Neurologické studie u opic (Ungerleider, Mishkin, 1982) prokázaly existenci dvou vizuálních systémů pro zpracování vstupů z prostředí. Podle informací, které zpracovávají, byly nazvány *what* a *where dráha*. *Where dráha* se v některých publikacích nazývá *dorzální* a je zodpovědná za reprezentaci prostoru z hlediska polohy objektů a vykonávání akcí. Na vstupu obsahuje velká receptivní pole citlivá na směr pohybu. Při následném zpracování jsou používány informace o poloze hlavy a oka. Z neuroanatomického hlediska směřuje do parietálního lobe zahrnující mediální temporální (MT), mediální superiorní temporální oblasti (MST) a posteriorní parietální ventrální (VIP) a laterální (LIP). Nejsilnější důkazy o existenci této dráhy pocházejí z výzkumů hemiprostorových neglektů - přehlížení poloviny zorného pole.

Jelikož se budeme v následujících kapitolách věnovat referenčním rámcům (souřadných osobních systémů) používaným při reprezentaci prostoru, je pro nás důležité zjištění, že parietální prostorová reprezentace používá pro zpracování více referenčních rámců (Snyder, Grieve a Andersen, 1998). Staví na poznatcích o prostorové reprezentaci a teorii referenčního systému, který je tvořen více vrstvami a prostorovými podsystémy (Andersen, Snyder, Bradley, Xing, 1997).

Další výzkumy mozkových lézí odhalily specializaci jednotlivých center. Poškození *ventrální dráhy* (*what systém*) se u pacientů projevuje neschopností identifikovat předměty na základě jejich velikosti a orientace, přestože jsou schopni objekty uchopit a správně s nimi manipulovat (Millner, Goodale, 1995). Poškození *dorzálních drah* má za následek opačný efekt. Pacienti jsou schopni rozpoznat objekt, ale nejsou schopni ho správně uchopit či manipulovat s ním (Perenin, Vighetto, 1988). Také Milner a Goodale (1995) referují o pacientovi, který je schopen manipulovat správně s objekty hapticky, ale není schopen posoudit jejich velikost

perceptuálně. Rozdíl mezi percepcí a akcí byl sledován i u normálních dospělých (Bridgeman, Gemmer, Forsman, Huemer, 2000).

V lingvistické rovině vedlo odhalení dvou specifických drah ke spekulacím o vývoji jazyka. Podle Landau a Jackendoffa (1993) souvisí *where systém* s užitím předložek, jelikož *what systém* musí rozpoznat velmi mnoho kategorií rozličných objektů, zatímco *where systém* rozliší pouze malé množství prostorových vztahů. *Where systém* produkuje pouze omezené informace o detailní geometrii vizuální scény. Ve prospěch hovoří některá neuropsychologická bádání, např. Damasio (2000) zjistil pomocí metody PET, že pojmenování dynamických akcí (chození, běhání) a prostorových vztahů (v, nad) aktivuje oblasti spojené s dorzálním *where systémem*.

Jeannerod (1994) vysvětluje odlišnost obou systémů pomocí existence sémantické a pragmatické reprezentace. První je spojena s vizuálními představami, druhá s motorickými. Motorické znamenají přípravu a plánování akce, její provádění či pouhé představy o ní. Následně uvádí tyto dva systémy do korespondence s *where* a *what dráhou*. Sémantická reprezentace kóduje detailní informace o objektu, pragmatická zachycuje vizuální vlastnosti vztahující se k affordancím, tzn. je důležitá pro organizaci motoriky směřující k manipulaci s objekty.

Důkazy pro rozdílné zpracování vizuálních vjemů nacházíme také u výzkumů, které se zabývají zpracováním odlišných slovních druhů (Cappa a Perani, 2003). Levý temporální neokortex slouží ke zpracování podstatných jmen a dorsolaterální prefrontální kortex ke zpracování sloves. Rozdíl závisí na sémantice daných slov, nikoliv na jejich gramatické kategorii. Důkazem je nepřítomnost rozdílu při zpracování podstatných jmen spojených s určitou akcí (Pulvermueller, Mohr, Schliechert, 1999).

Z výsledků výzkumů v oblasti neuroanatomie můžeme vyvodit, že pro zpracování informace o poloze objektu v prostoru je používán samostatný systém. Navržený model bude svou činností odpovídat právě zmiňovanému *where systému*. Přestože je přímo spojen s reprezentací motoriky, v první verzi modelu ještě nejsou zmíněné subsystémy zahrnuty. Je to dáno postupným vývojem a rozšiřováním modelu, který bude doplňován o přídatné systémy, pomocí kterých zajistíme zvyšování biologické a psychologické plausibility. V prvotních fázích se soustředíme na tvorbu systému obsahujícího neurony citlivé na pohyb. Právě tato schopnost se ukáže jako klíčová pro reprezentaci dynamických prostorových vztahů.

5.2 Psychologické poznatky

V klasickém modelu Baddeleye a Hitcha (1974) byla vizuální a prostorová informace zpracovávána v jednom systému nazývaném *vizuo-prostorový náčrtník*. Současné studie, které vycházejí z výše zmíněných neurobiologických poznatků, však poukazují na to, že tato část pracovní paměti je složena ze dvou specializovaných systémů - vizuálního a prostorového. Vizuální systém slouží k zapamatování tvarů, textury, barvy a metrických vzdáleností. Prostorový systém uchovává polohu a pohybovou informaci (Knauff, 2006).

Na podporu teorie slouží výzkumy Klauera a Thao (2004), kteří prokázali, že vizuální krátkodobá paměť je silněji ovlivněna vizuální interferencí a prostorová paměť naopak prostorovou interferencí. Prostorové úlohy aktivují okcipitální a parietální regiony pravé hemisféry, tedy oblasti dorzální *where dráhy* a vizuální úlohy aktivují inferotemporální oblasti levé hemisféry, oblasti ventrální *what dráhy* (Knauff, 2006).

5.2.1 Psychopatologie vnímání prostoru

Velmi mnoho cenných informací přinášejí výzkumy, které zkoumají patologie v oblasti reprezentace prostoru. Mezi nejznámější patří Williamsův syndrom. Jedná se o vzácné genetické onemocnění postihující prostorové schopnosti při zachování jazykových. Vzniká na základě mikrodelece sedmého chromozomu. U pacientů se projevuje specifickým profilem tváře, poškozením srdce a orgánů a lehkou až střední retardací. Z hlediska výzkumu prostoru je nejdůležitější deficit v prostorové orientaci při zachovaných jazykových schopnostech. Je zachováno rozpoznávání tváří, rozpoznávání objektů, percepce biologických pohybů a některé aspekty prostorového jazyka. Pokud mají pacienti zachovány jazykové schopnosti, nabízí se otázka: Jak může emergovat jazyk na těchto deficitních systémech prostorové reprezentace?

Odpověď můžeme hledat v reprezentační jednoduchosti jazykových označení. Pacient dokáže jazykem označovat, protože poškozené struktury jsou schopné takto jednoduchého zpracování, ale v případě komplexnější úlohy, kterou je potřeba provést v představě, selže, jelikož poškozené systémy takto komplexní operaci neprovedou (Landau, Hofmann, 2005).

Existuje také alternativní možnost vysvětlení. Jazyk emerguje nezávisle na prostorových strukturách, tzn. jazyk a prostorová reprezentace jsou nezávislé a pravděpodobně modulární (Bellugi et al., 1988). Podle autora zpracovává lingvistický a nelingvistický systém úlohu odlišně, s využitím jiných mechanismů. Jazyk je založen na syntaxi a morfologii, nelingvistický systém na rekonstrukci prostorových polí.

Zmíněné teorie nejsou protichůdné. Oproti teoriím uváděným na začátku této práce (viz Analytická filosofie), ve kterých je reprezentace prováděna pomocí jednoho systému (čistě symbolického nebo čistě konceptuálního) se posouváme dále, jelikož psychologické přístupy se snaží nalézt způsob propojení mezi lingvistickou (symbolickou) a nelingvistickou (konceptuální) vrstvou reprezentace. Řešíme otázku ukotvení symbolů, hledáme způsob propojení konceptů se symboly a jejich vzájemné kooperace. Oproti řešení procesu ukotvení „zdola nahoru“ metodou modelování se pohybujeme v oblasti psychologických výzkumů, při kterých mají zkoumané osoby lingvistický a nelingvistický systém vytvořen, tzn. neřešíme otázku vzniku subsystémů, ale jejich vzájemné kooperace. Jedná se o podobnou problematiku, kterou se zabýval Paivio (1985) ve své teorii duálního kódování.

Vraťme se k otázkám týkajícím se Williamsova syndromu (WS). Současné výzkumy se shodují, že deficit pacientů s WS je způsoben poruchou vývoje referenčního systému (souřadných osových systémů) vizuální pozornosti. (Brown, Johnson, Paterson, Gilmore, Longhi, Karmiloff-Smith, 2003). Deficit prostorové konstrukce objevující se později, může reflektovat selhání základních struktur, např. referenčních systémů.

Experiment s pacienty trpící WS naopak prokázal **zachování** osového referenčního systému (Landau, Hoffman, 2005) při reprezentaci směrů na osách, jak při nelingvistickém posouzení polohy (ukazování objektů), tak při pojmenování polohy (označení prostorového vztahu). Výsledky prokázaly sdílenou strukturu nelingvistické a lingvistické prostorové reprezentace.

Experiment (Landau, Hoffman, 2005) přinesl také přesnější identifikaci rozdílů mezi normální ZO a pacienty s WS. Prokázalo se, že mají deficit v reprezentaci prostoru, který neznamena úplnou absenci referenčních rámců či systémů. Děti s WS mají například podobnou organizaci osových systémů jako normální ZO. Problematické jsou objekty vzdálené od os a také objekty kolem horizontální osy. Nejzajímavějším výsledkem je zachování referenčního systému jako organizující struktury jak v lingvistických, tak v nelingvistických úlohách. Byla prokázána fundamentálnost referenčního systému v procesu prostorových reprezentací a potvrzeno propojení mezi lingvistickým a nelingvistickým prostorovým vnímáním.

Z experimentu je zřejmé, že deficit jednoho subsystému se projeví zvýšenou závislostí celého systému na subsystému druhém. Což nás opět směřuje k teoriím založeným na reprezentaci pomocí lingvistických a nelingvistických znalostí. Z výsledku nejsme schopni rozhodnout, jakým způsobem kódování jsou jednotlivé reprezentační subsystémy tvořeny (deskriptivní či depiktivní), ale můžeme konstatovat, že systém by měl obsahovat více primárních struktur (lingvistické a nelingvistické), které odlišným způsobem reprezentují empiricky získané znalosti. Jejich následným propojením vzniká kvalitativně nový systém využívající společný potenciál obou. Jedná se o mapování mezi jednotlivými subsystémy. Takový způsob reprezenta-

ce znalostí je použit i v navrženém systému. Dva primární reprezentační subsystémy jsou na základě učení mapovány do společné multimodální úrovně, ve které jsou asociovány na základě společného výskytu.

5.2.2 Osové systémy

V předchozí kapitole jsme se setkali s pojmem referenční systém, který bývá v oblasti vnímání prostoru používán v různých kontextech. Výše zmíněný termín znamená užití osových systémů, tedy systémů pravoúhlých os, které jsou používány jako referenční systém pro určení prostorových vztahů. Existují odlišné názory na jejich užití a privilegovanost v kontextu lingvistických a nelingvistických systémů.

Přestože je jazyk specializovaným systémem s vlastními formálními pravidly, v případě prostorové sémantiky je ukotven referenčními rámci, které jsou používány i pro nelingvistické účely. Empirický důkaz o vzájemné homologii lingvistických a nelingvistických systémů přinesli Hayward a Tarr (1995). Zjistili, že polohy, které byly nejkonzistentněji pojmenovány základními prostorovými termíny (nahoru, dolů, vlevo, vpravo), byly také nejlépe zapamatovány v nonlingvistické paměťové úloze. Tyto polohy odpovídají hlavním osám objektů.

S těmito závěry je možné polemizovat, jelikož v úloze intervenuje statistický průměr, s jakým jsou používána prostorová označení během života, tzn. jejich privilegovanost. Nelze rozpoznat, zda jsou systémy homologní, protože pracujeme se subjekty, kteří mají osový systém naučený a odpověď je dopředu známa. Výsledky nejsou překvapivé, pouze potvrzují nejčastější výskyt vzájemné polohy objektů. O homologii obou systémů se ale z takového experimentu nedozvíme nic průkazného. Výsledky jejich experimentu byly replikovány kroslingvisticky (Munnich, Landau, Doshier, 2001) na populaci dospělých rodilých mluvčích angličtiny, korejštiny a japonštiny. Použili metody Haywarda a Tarra a našli shodu v kódování prostorových vztahů, ale zároveň odhalili specifické používání prostorových označení v jednotlivých jazycích.

Odlišný názor na užití osových referenčních rámců nacházíme u Huttenlocher (1991). Podle ní má osový systém roli prototypu v lingvistické reprezentaci a roli hranic kategorie v nelingvistické. Jejím závěrem je, že oba jsou nasazené na třetí strukturu, skrze kterou se odlišně realizují. Základem odlišného přístupu je názor, že schopnost kategorizovat se zvyšuje se vzdáleností, přičemž největší je na okraji kategorie (Regier, 2003). Přesto lze namítat, že takový způsob kategorizace funguje u rozpoznání objektů, ale v případě vnímání prostoru neexistují jasně odlišitelné prostorové kategorie, jelikož na okrajích nevzrůstá jasnost, do které

kategorie objekt patří. Kategorie jsou *fuzzy*, což vede k závěru, že existují pouze pravděpodobné způsoby kategorizace. Huttenlocher však zastává názor, že rozlišovací schopnost (efekt kategorie) na osách je nejpřesnější, protože tam leží hranice dvou kategorií (Regier, 2003). Výzkum nabízí alternativu oproti výzkumům Haywarda a Tarra, jelikož zadání úlohy eliminuje vliv užití klasického osového referenčního systému.

5.2.3 Blížkost objektů

V oblasti vnímání prostoru se setkáme kromě absolutních prostorových označení také s relativními. Například pojem blízkosti je relativní, jelikož závisí na velikosti objektů. Vzhledem k jejich rozměrům posuzujeme, kde končí blízkost. Kromě toho se do výsledku promítá vliv prostředí, ve kterém jsou předměty umístěny (rozlehlost). Jestliže chceme posoudit vzdálenost předmětů na stole či na poli, budou naše odhady různé. Důležitým faktorem je i poměr velikostí jednotlivých objektů. Při posuzování dvou objektů mají lidé tendenci porovnávat menší objekt k většímu. Omezení plynou i z hlediska funkce objektu. Použití věty „plot je blízko kola“ je možné v úrovni představy, ale v rovině jazyka se nepoužívá.

V případě vlivu velikosti objektů na odhad prostorového vztahu se nabízí použít funkci, která vytvoří mezi velikostí a vzdáleností vztah. V úvahu je ovšem nutné vzít také další faktory. Jestliže posuzují objekty, které se pohybují, popřípadě se mohou pohybovat rozdílnou rychlostí (kolo/formule), promítne se to do odhadu a je nutné tedy stávající funkci rozšířit o daný faktor. Podobně je tomu s funkčním vztahem objektů. Pokud jsou předměty v poloze, která umožňuje funkční interakci (židle otočená směrem ke stolu), promítne se do posouzení vztahu blízko/daleko.

Oproti ostatním objektům, které lze porovnávat v absolutním měřítku, nemůžeme o vztahu blízko/daleko říci, které systémy se na něm podílejí, jelikož užití je vždy relativní a svázáno s kontextem (Coventry, 2004). Posuzovaná vzdálenost je subjektivní a nelze ji použít v absolutním měřítku. Proto je třeba vzít v potaz všechny zmíněné faktory, bez nichž je velmi obtížné zodpovědět „Jak blízko je blízko?“.

V souvislosti se zmíněnými experimenty se nabízí otázka, zda je reprezentace prostoru v mentálních modelech euklidovská? McNamara (1986) dochází k závěru, že po fázi učení jsou blízké objekty více asociovány (*primují* jeden druhý), než je tomu u vzdálenějších objektů. Zároveň objekty v jednom ohraničeném regionu jsou více *primovány* než stejně vzdálené objekty ve dvou regionech. *Prostorový priming* (který v mnohém připomíná *gestalt principy*) není závislý pouze na vzdálenosti, ale i na úrovni vztahu mezi objekty (Langston, Kramer a

Glenberg, 1998). Závěrem je, že prostorová informace není reprezentována pouze v euklidovském systému, ale také hierarchické a prostorové vztahy ovlivňují jeho fungování (Langston, Kramer a Glenberg, 1998).

Nacházíme i zcela neeuklidovská řešení problematiky. Tversky a kolegové (1990,1992) tvrdí, že prostorová označení jsou kódována pomocí *prostorové soustavy*, která je rozhodně neeuklidovská. Pomocí experimentu, ve kterém ZO memorovaly scénu a následně si měly vybavovat objekty v ní různými způsoby pohybu v rámci vizuální představy scény (shora-dolů, zleva-doprava apod.), zjistili funkční závislost reakčních časů vybavení pro různé způsoby prohledávání scény, přičemž nejrychleji si lidé vybavují objekty při pohybu shora-dolů a nejpomaleji zleva-doprava.

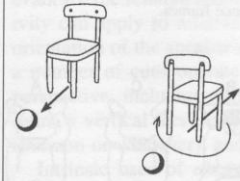
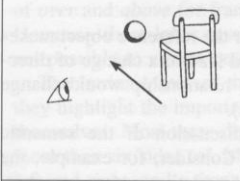
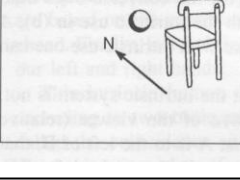
Ve výzkumech týkajících se odhadu blízkosti objektů se objevují názory, že reprezentace jsou v rozporu s euklidovskými principy geometrie, díky porušení axiomu symetrie (vzdálenost z místa A do B je stejná jako vzdálenost z B do A). Mc Namara a Diwakhar (1997) to vysvětlují pomocí modelu *kontextuálního škálování*. Za asymetrii v určení vzdálenosti mohou kognitivní procesy spojené s odhadem velikosti, posuzované v určitém kontextu (např. „Jak daleko je televize od pohovky?“). V případě zrušení vlivu kontextu v zadání (místo věty „Jak daleko je televize od pohovky?“ je použito věty „Jak daleko je televize a pohovka navzájem?“) se ve výsledcích asymetrie neobjeví, jelikož není jasné, který objekt je referenční (Ferenz, 2000).

Právě díky závislosti na kontextu, který vyplývá z velikosti, tvaru a funkčního vztahu porovnávaných objektů, jsem se reprezentací tohoto pojmu ve svém modelu nezabýval. Systém schopný rozeznávat blízkost objektu by v sobě musel obsahovat modul pro rozpoznávání objektů (*what dráhu*) a další reprezentační mechanismy, které by umožnily plausibilní model daného prostorového označení.

5.2.4 Centrum deixie

Jestliže jsme hovořili o referenčních rámcích jako o osových systémech, které používáme při posuzování prostoru z hlediska pozorovatele, existuje také dělení zohledňující použití jazyka při interakci s okolním světem. Oproti předchozímu fixnímu posuzování referenčního rámce z hlediska pozorovatele můžeme při jazykových promluvách umisťovat počátek osového systému do prostředí. V literatuře zabývající se prostorovou kognicí je často referenční systém reprezentován jako soustava ortogonálních os, jejichž centrem je retina, hlava, tělo nebo další body, objekty či pole v prostoru (Behrmann, 2000; Colby a Goldberg, 1999; McCloskey,

2001). Díky rozličným úlohám může být použito více referenčních systémů. Nejlépe tyto referenční rámce diferencuje taxonomie prostorových vztahů na Obr. 11.

	Rotace		
	Pozorovatel	Ref. Objekt	Celé pole
Vnitřní míč je před židlí 	Stejný popis? Ano	Stejný popis? Ne	Stejný popis? Ano
Relativní míč je vlevo od židle 	Ne	Ano	Ne
Absolutní míč je severně od židle 	Ano	Ano	Ne

Obr. 11 Vliv referenčních rámců na prostorové označení

V případě uvažování o prostorových vztazích navrhuje Levinson (1996) tři druhy referenčních rámců:

Intrinsický (objektový) rámeček - souřadný systém je orientován podle referenčního objektu - např. talíř je před chlapcem, talíř je za chlapcem.

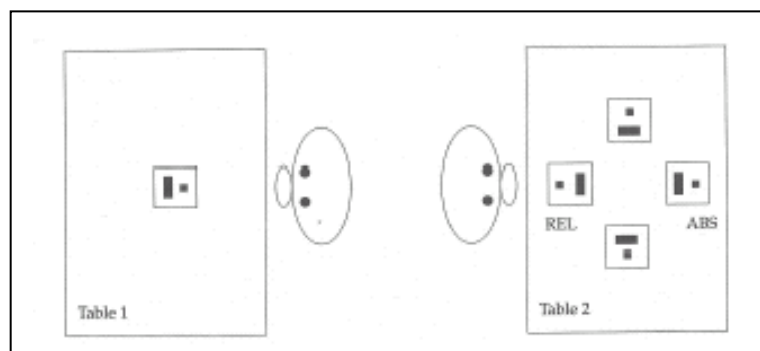
Relativní (vzhledem k pozorovateli/centru deixie) rámeček - souřadný systém je umístěn mimo objekt, který je popisován. Jedná se o centrum deixie, které může umisťovat pozorovatel na libovolný objekt sledované scény, tedy i do centra senzoru - např. chlapec je před talířem, chlapec je za talířem, ale i talíř je přede mnou, talíř je vedle mě.

Absolutní (vzhledem k prostředí) rámec - jsou používány univerzální fixní souřadné systémy, které lze objektivizovat - např. *chlapec je na sever od talíře, chlapec je nad talířem.*

rámec	vztah	centrum deixie
vnitřní	Rel (Loc,Ref)	centrum je v Ref
relativní	Rel (Loc,Ref,V)	centrem je entita V
absolutní	Rel (Loc,Ref)	centrum je dáno fixně okolím gravitace, světové strany

Tab. 1 Referenční rámce

Tyto referenční systémy se liší napříč kulturami. Některé jazyky obsahují pouze absolutní rámce (Regier, 1996).



Obr. 12 Členové kmene Guugu Ymathirr používají pouze absolutní rámec k označení prostorových vztahů. Pokud se nejprve podívají na konfiguraci předmětů na prvním stole (Table 1), a následně se otočí o 180 stupňů ke druhému stolu (Table 2), budou hodnotit jako shodnou konfiguraci tu vzdálenější (ABS). V ostatních kulturách bývá hodnocena jako shodná bližší konfigurace (REL), jelikož používají relativní rámec. (Regier, 1996)

Z hlediska referenčních rámců se mohou objevovat různé způsoby reprezentace prostoru v závislosti na objektu, který se ve scéně objevuje. Některé objekty totiž obsahují vlastní prostorová schémata (rámce), která bereme při určení polohy v potaz. U většiny živých tvorů vytváříme prostorová označení na základě orientace jejich obličeje a končetin. Jedná se o *vnitřní lidské schéma*. To se projevuje při určování prostorových vztahů a může jej zásadně ovlivnit. Prostor okolo každého živého tvora je dělitelný na šest směrů - nahoru, dolů, vlevo, vpravo, dopředu a dozadu. Tyto směry jsou implicitně dané podle orientace hlavy a končetin.

Schémata však mohou být různá pro odlišné typy živočichů. Například u hada nepřisuzujeme polohu vepředu podle umístění břicha. Kromě levé a pravé strany jsou ostatní směry přizpůsobeny specifické stavbě a způsobu pohybu v prostředí.

U schémat neživých objektů závisí použitý rámec na způsobu interakce s těmito objekty, např. použití označení „u stolu“, přistupujeme-li k němu z pozice úředníka či člověka, který přišel s žádostí (*relativní schéma*). Kromě *vnitřního* a *relativního schématu* ještě existuje třetí, které leží na pomezí mezi nimi. Jedná se o *náhodné vnitřní schéma*, nazývané také *pohybové vnitřní schéma* (Jackendoff, 1996). Například u pohybující se kulečnickové koule závisí její vnitřní schéma na směru pohybu a může se měnit. Díky jejímu tvaru nemůžeme aplikovat fixní schéma podobné lidskému, a proto jej přiřadíme až na základě pohybu. Přední část objektu může být také závislá na poloze pozorovatele. Pokud stojíme v kostele, je přední část u oltáře, pokud před kostelem, je přední část u vchodu (Fillmore, 1982).

Otázkou je, jaký referenční rámec si vybíráme při označení prostorového vztahu? Carlson-Radvansky a Irwin (1993) zjistili v sérii experimentů, ve kterých ZO posuzovaly prostorový vztah „nad“ při různých natočeních objektů (člověk, židle, míč) odchylky při používání referenčních rámců. Účastníci volili nejčastěji absolutní, následně vnitřní a nejméně preferovaným byl relativní rámec. V dalších experimentech potvrdili, že použití rámců je simultánně aktivované (Carlson-Radvansky, Irwin, 1993) a také automatické (Carlson-Radvansky, Logan, 1997). Kromě zjištění ohledně výběru či inhibice referenčních rámců (Carlson, West, Taylor, Hendon, 2002) došli až ke zjištění modulace neurální aktivity ve formě EEG při užití rozdílných referenčních rámců.

Názory na vývoj deixie (centra osového referenčního systému) se mohou lišit. Například Piaget a Inhelder (1967) vycházejí z předpokladu, že dítě během vývoje nejprve používá pouze egocentrický pohled, který je ztotožňován s relativním rámcem, s centrem deixie přímo v pozorovateli. Podle výzkumů Johnston a Slobina (1979) se děti nejdříve naučí výrazy *v*, *na*, *pod* a *vedle* a potom *mezi*, *vpředu*, *vzadu* (pro základní objekty a potom pro ostatní). Autoři přisuzují toto pořadí mnoha faktorům, např. konceptuální komplexitě prostorových vlastností.

5.3 Lingvistické poznatky

I v oblasti kognitivní lingvistiky se zaměřením na prostorová označení narážíme na rozpor mezi klasickými symbolickými teoriemi a reprezentací významu na základě jeho senzorického ukotvení. Zastánci první z nich (Fodor, Pylyshyn, 1988) se přiklánějí ke způsobu tvorby konceptů na základě atomických jazykových označení, které můžeme podle syntaktických pravidel kombinovat do nekonečného množství promluv. Lakoff (1987) oponuje, že jsme při tvorbě konceptů omezeni naší tělesnou zkušeností, která determinuje strukturu konceptů a zdůrazňuje proces ukotvení. Jako klíčový prvek při reprezentaci prostoru považuje vtělenost (embodiment). Podobně jako ostatní kognitivní lingvisté (Langacker, 1989) zastává názor, že konceptuální systém je tvořen na základě tělesné zkušenosti a jádrové koncepty jsou ukotveny v percepci, tělesných pohybech a zkušenosti fyzického či sociálního charakteru. Svou teorii rozšiřuje o mentální akty. Pokud není koncept ukotven ve zkušenosti, je možno jej vytvářet na základě metafor, metonymie či mentálních představ. Což jsou teoretické požadavky, které nás utvrzují o správnosti přístupu k reprezentaci významu z hlediska tvorby perceptuálně získaných konceptuálních reprezentací, ale neodhalí nám podrobnosti o zkoumané doméně. Proto se podívejme na konkrétnější výzkumy.

V souvislosti s osovými systémy byl zmíněn výzkum Haywarda a Tarra (1995). Autoři interpretují své výsledky také v úrovni lingvistiky. Sérií experimentů potvrdili, že lingvistická organizace staví na nelingvistické. Prokázali privilegovanost osového systému, na který jsou napojeny prototypy prostorových vztahů a následně prostorová označení.

Huttenlocher naopak poukazuje (1991), že některé prostorové označení se člověk učí bez nutnosti ukotvení v nelingvistické prostorové reprezentaci, např. *nahore* jako antonymum *dole* nebo rozdíl mezi *between* a *among*. Taková argumentace však není z hlediska kognitivní sémantiky možná. Už samotná možnost antonym vyžaduje prostor (v mentální představě), ve kterém se může opositum nacházet. V případě *between* a *among* potřebujeme znát počet objektů v prostoru (opět nelingvistickou perceptuální informaci), abychom mohli vytvořit prostorové označení. Ve výsledku opět narážíme na spory týkající se privilegovanosti jednoho systému nad druhým.

Problematika privilegovanosti lingvistických a nelingvistických reprezentací je dobře demonstrovatelná také na sporu Chomsky vs. Kay a McDaniel, ohledně lingvistických univerzálií (tvrzení, která jsou platná pro veškeré jazyky). Chomsky (1965) tvrdí, že učení se jazyku není možné bez vrozených lingvistických pravidel. V oblasti transformační gramatiky však existují kromě Chomského teorie také alternativní přístupy. Comrie (1981) porovnává Chomského

přístup k lingvistickým univerzáliím s výsledky Greenberga (1963). Ten zkoumal více jazyků a místo striktního systému nabízí několik alternativních výkladů univerzálií. Podle Comrieho je v Chomského přístupu největším problémem kladení důrazu na vrozenost, která je podle něj vědecky prázdná. Je to pouhé jméno, kterým označujeme množinu univerzálií, tzn. cokoliv pozorujeme jako univerzální, můžeme nazývat jako vrozené. Jak bylo zmíněno, pojem vrozenosti nelze falzifikovat.

Kay a McDaniel (1978) tvrdí, že sémantické univerzálie (podkategorie lingvistických univerzálií) se dají odvodit z perceptuálních univerzálií. Dokazují to na svém experimentu s barvami a jejich univerzálním počtem ověřeným intersubjektivně. Barvy však nehrají v oblasti kognitivních procesů významnou úlohu, proto se podívejme na sémantické univerzálie z hlediska vnímání prostoru.

Zastánci teorie neverbální akvizice prostorových vztahů (Clark, Johnston, Slobin, Mandler, Ratner) tvrdí, že většina konceptuálních struktur je prelingvisticky kódována. Podle nich žádný jazyk nemůže kódovat něco, co je konceptuálně zcela neznámé pro posluchače v jiných jazycích. Naopak Regier (1996) tvrdí, že jazyk Guugu Ymithirr a jeho absolutní orientace prostorových vztahů (viz Obr. 12), je důkazem proti takovým tvrzením. Jeho argument není zcela přesný. Jestliže koncepce prostorového vztahu neobsahuje deiktické centrum v pozorovateli, neznamená to, že neobsahuje prelingvistickou koncepci prostorového vztahu. To, že se koncept nemění s rotací hlavy, je pouze jeden aspekt, který nemodifikuje celý koncept a nečiní jej nejasný pro libovolného uživatele libovolného jazyka.

Jiným příkladem pro odlišnost prostorových konceptů napříč jazyky (soustředěná na jeden aspekt konceptu, nikoliv na jeho fundamentální vlastnosti) je longitudinální studie Choi a Bowermana (1992), která ukázala, že preverbální prostorové koncepty ovlivňují rozdílným způsobem význam prvních slov a asociace na ně u korejských a anglických dětí. Podobných výsledků pro australské děti dosáhl Bavin (1990).

V konečném důsledku je obtížné odlišit, které jsou esenciální (fundamentální) vlastnosti prostorového konceptu (předpokládáme, že jsou shodné napříč jazyky v preverbální úrovni (ortogonálnost, bipolární dimenze na osách apod.)), a které vlastnosti intervenují během akvizice jazyka, ale nemají na prostorový koncept zásadní vliv, pouze jej zpřesňují či zrelativňují. Navíc se setkáváme i s tvrzením, které existenci sémantických univerzálií v souvislosti s prostorovými koncepty zcela popírá.

Podle Gopnik (1984) neexistují sémantické univerzálie, protože dětské koncepty se mění radikálním způsobem. Vypožorovala, že termíny *nahoru*, *dolů* a *na* znamenají něco zcela odlišného u dětí a u dospělých. U dětí znamenají touhu přemístit patřičným směrem objekt, zatímco u dospělých statické označení. Tvrdí, že koncepty nejsou vybudovány z bloků, o kterých se

hovoří v jazyce dospělých. Koncepce prostoru není pevnou bází, ale spíše terénem pohyblivých písků, jehož proměnlivost činí pojem sémantických univerzálií prázdný. Podle Regiera (1996) je přístup Gopnikové argumentem proti fixní „lingua universalis“ (Fodor, 1975) než proti univerzáliím. Vývoj prostorových univerzálií lze vysvětlit i dle závěru Gopnikové, jelikož můžeme předpokládat, že univerzálie se vyvíjí.

Problematické v tomto sporu je identifikace fundamentálních a proměnlivých vlastností. Sémantické univerzálie jsou platné, pokud si prostorový koncept během vývoje udrží neměnné prototypické jádro. Přikláním se spíše k názoru potvrzujícímu existenci univerzálií, protože orientace v prostoru je natolik základní pro intersubjektivní komunikaci, že lze předpokládat nejen univerzální prototypy, ale na druhou stranu také robustní mechanismy, které prostorovou informaci doplňují a podléhají ontogenetickému vývoji. Podle Clarka je akvizice prostorových neverbálních konceptů ukončena v okamžiku, kdy dítě začne mluvit.

V našem modelu je vývoj konceptu omezen pouze na dobu, během které je mu prezentována trénovací množina. Změny na úrovni způsobu vnímání a operace s koncepty nejsou pro jednoduchost modelu možné.

5.4 Jednotný model reprezentace prostoru

V této části se soustředíme na konkrétní oblast vnímání prostoru, která souvisí s činností realizovaného modelu. Pokud posuzujeme reprezentaci prostorových vztahů z hlediska polohy objektů v zorném poli pozorovatele (v jediném relativním referenčním rámci, kdy je centrem deixie střed oka), narazíme na rozdílné teorie, které se snaží o zachycení vztahů z hlediska jediného systému, popřípadě užitím více systému či úrovní zpracování. Pro účely této práce jsem zvolil Coventryho dělení teorií reprezentace prostoru, které je vhodné nejen z didaktického hlediska, ale také pro možnost přímé aplikace v oblasti modelování vnímání prostoru. Coventry (2004) navrhuje integrovat modely prostorových vztahů do jednotného rámce, tvořeného třemi vrstvami, který slouží ke zpracování informace z prostředí a její reprezentaci. Činnost jednotlivých vrstev je následující.

- 1. Vrstva geometrických vztahů - prostorový vztah je reprezentován na základě polohy dvou objektů v prostoru. V potaz je brána pouze tato konfigurace.*
- 2. Vrstva extra-geometrických vztahů - prostorový vztah je reprezentován pomocí sil, které na objekty působí, například gravitace, směry pohybu, vzájemné interakce.*

3. *Vrstva znalosti objektů – prostorový vztah je reprezentován pomocí předchozí znalosti objektu, která ovlivňuje prostorové označení na základě funkčního vztahu mezi objekty.*

Navrhuje zahrnout tyto přístupy do funkčního geometrického rámce (FGR), který bude zohledňovat společné vlastnosti. Kromě toho je nutné vzít v potaz také vliv přirozeného jazyka, který používá prostorové předložky velmi specificky, přičemž některé výjimky neberou v potaz ani prostorovou konfiguraci ani působení sil. Tyto poznatky je do výsledného modelu také nutné zakomponovat, v případě že od něj požadujeme schopnost porozumění a produkce v přirozeném jazyce.

5.4.1 Geometrické vztahy

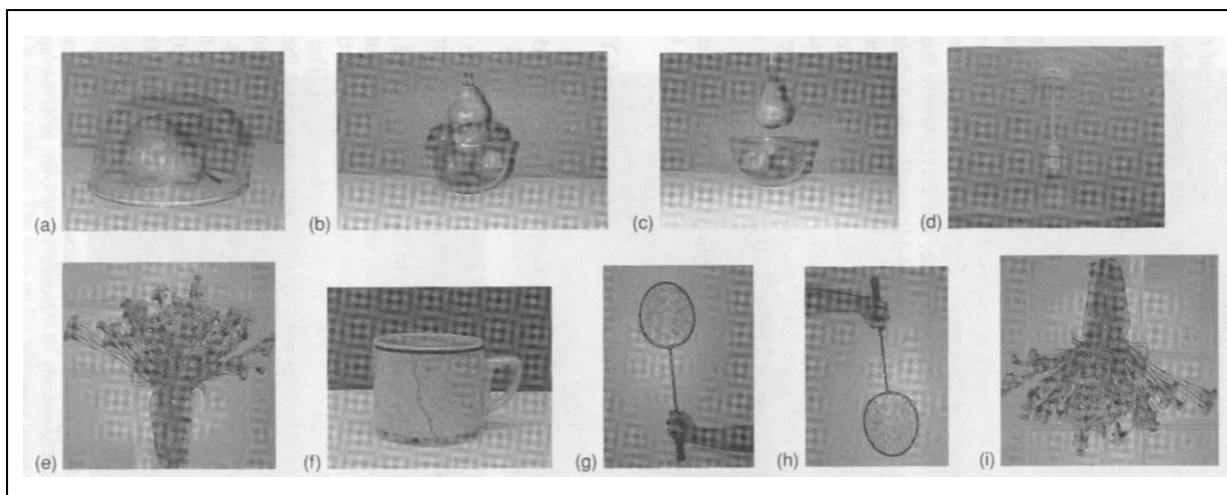
Způsob reprezentace geometrických vlastností se objevuje v teoriích mnoha badatelů, protože se jedná o základní vlastnost, která nás v této oblasti zajímá. V následujícím textu sumarizuji základní přístupy.

5.4.1.1 Symbolická reprezentace prostoru

I v oblasti vnímání prostoru nalézáme řešení, které vycházejí ze symbolických teorií kognice založených na logickém aparátu. Pokud však ze zmíněných teorií vycházíme, narazíme hned v počátcích na obtíže. Příkladem může být pokus o definování významu pomocí *jádrového významu*, což je snaha vyjádřit co nejúspornějším symbolickým popisem daný prostorový vztah (Cooper, 1968). Například definice označení „*v*“ jako:

$$V(X,Y) \text{ iff } \text{Umístění}(X, \text{Interior}(Y))$$

Jedná se o vyjádření pomocí logiky prvního řádu. Přístup je efektivní, protože nabízí mapování 1:1 mezi perceptuální a symbolickou rovinou. Tímto způsobem je možné popsat prostorový vztah velmi úspornou formou. Jedná se o ideální způsob formalizace prostoru pomocí metod klasické komputace, které můžeme využít v oblasti umělé inteligence. Již v začátku však narazíme na množství problémů. Varianty prostorových vztahů, jak jsou používány v přirozeném jazyce, leží mimo rámec úzkého logického vymezení. Označení *v* totiž může znamenat mnoho odlišných možností.



Obr. 13 Prostorové označení „v“ může nabývat mnoho podob, přičemž některé jsou chybné, přestože splňují podmínky jádrového významu, např. obr. a), c) či i). (Coventry, 2004)

Jak je zřejmé, mnohé perceptuální prostorové konfigurace nejsou zachytitelné daným pravidlem. Proto přichází Miller a Johnson-Laird v roce 1976 s rozšířením, které má definovat význam „v“ i pro objekty větší než interiér druhého objektu, nebo pokud nejsou v druhém objektu zcela uzavřené. Nová definice je delší:

$V(X,Y)$ if ($\text{Část}(X,Z)$ and $\text{Obsahuje}(Z,Y)$)

Jedná se o popis vztahů mezi objekty X a Y, pomocí vlastností *obsahovat* a *část*. Jelikož je výraz zapsán v predikátové logice, neexistuje možnost *stupňovité odpovědi*, ve smyslu míry adekvátnosti daného označení a neexistují tedy dobré či špatné příklady prostorových výrazů. I z empirického hlediska definice selhává. Do takové definice významu spadají i označení „láhev je v uzávěru“. Symbolická definice významu prostorového vztahu nám nedostačuje k označení veškerých perceptuálních možností, jak se s nimi běžně setkáváme. Podobně je tomu u ostatních prostorových označení (na, přes apod.). Snaha o symbolické uchopení selhává, čímž se vzdalujeme možnosti reprezentovat význam pomocí pravdivostní hodnoty na základě prezentovaného formalismu. Zároveň narážíme na problémy, které byly zmíněny v souvislosti s formální sémantikou. Podle Searla (1974) je pravdivostní hodnota daného značení závislá na znalosti v pozadí, význam je tedy do značné míry flexibilní. Jelikož je citlivost prostorových označení na kontextu zřetelná na mnoha výjimkách při jejich užití (Wood, 1967), je reprezentace pomocí *jádrového významu* neudržitelná. Použití logických operátorů a symbolických označení je vhodné pro manipulaci v abstraktní rovině, ale při tvorbě modelu prostorové sémantiky nedokáže zachytit všechny důležité vlastnosti.

5.4.1.2 Víceúrovňová reprezentace prostoru

Nacházíme také přístupy, které symbolickou reprezentaci rozšiřují o další úrovně, čímž eliminují výše zmíněné nedostatky. S víceúrovňovou reprezentací prostorových vztahů přichází Herskovits (1988). Ve svém přístupu zohledňuje pragmatické hledisko při reprezentaci prostoru pomocí samostatné úrovně. První úroveň je obdobná s *jádrovým významem*, nazývá se *ideální význam*, a zachycuje abstraktní geometrickou definici daného prostorového vztahu. V další úrovni obohacuje reprezentační bázi o *typy použití*, které jsou založené na *ideálním významu*, tzn. specifikuje použití pro různé (pragmatické) kontexty. V další úrovni jsou *typy* doplněny o principy, které je umožňují přizpůsobit ještě širšímu užití v přirozeném jazyce. Jedná se o *salienci, relevanci, typičnost a toleranci*. Nejlépe nám ke způsobu reprezentace poslouží ukázka:

Voda je ve váze

Ideální význam: v/ve - obsažení geometrického pojmu v 1 až 3 rozměrném geometrickém pojmu

Typy použití: z 8 možných typů pro „v/ve“ nejlépe odpovídá typ - prostorová entita v kontejneru

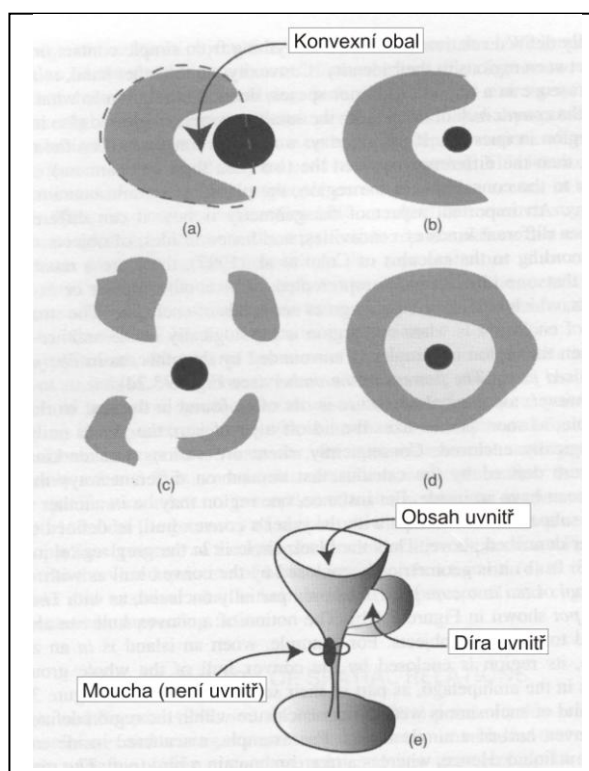
Herskovitsian ideal meanings and use types for <i>in</i> , <i>on</i> , and <i>at</i>
<p>Ideal meaning: <i>In</i>: inclusion of a geometric construct in a one-, two-, or three-dimensional geometric construct</p> <p>Use types: <i>In</i>:</p> <ul style="list-style-type: none"> Spatial entity in container Gap/object "embedded" in physical object Physical object "in the air" Physical object in outline of another or a group of objects Spatial entity in part of space or environment Accident/object part of physical or geometric object Person in clothing Spatial entity in area Physical object in a roadway Person in institution Participant in institution <p>Ideal meaning: <i>On</i>: for a geometrical construct X to be contiguous with a line or surface Y; if Y is the surface of an object O_y, and X is the space occupied by another object O_x, for O_x to support O_y</p> <p>Use types: <i>On</i>:</p> <ul style="list-style-type: none"> Spatial entity supported by physical object Accident/object as part of physical object Physical object attached to another Physical object transported by a large vehicle Physical object contiguous with another Physical object contiguous with a wall Physical object on part of itself Physical object over another Spatial entity located on geographical location Physical or geometrical object contiguous with a line Physical object contiguous with edge of geographical area <p>Ideal meaning: <i>At</i>: for a point to coincide with another</p> <p>Use types: <i>At</i>:</p> <ul style="list-style-type: none"> Spatial entity at location Spatial entity "at sea" Spatial entity at generic place Person at institution Person using artefact Spatial entity at landmark in highlighted medium Physical object on line and indexically defined crosspath Physical object at a distance from point, line, or plane

Tab. 2 Ukázka ideálních významů a typu užití pro různá označení, jak je navrhuje Herskovits (Coventry, 2004)

Následují pragmaticky blízké principy, které mohou doladit význam pokud je kontext nejasný. *Salience* umožňuje podstatnému jménu, které referuje k celému objektu, aby zastupovalo pouze část (kočka je pod stolem). *Relevance* souvisí s cíly komunikace, tedy s intencí řečníka a upřesnění kontextu (olej je na pánvi, olej je v pánvi). *Tolerance* vyjadřuje odchylku, kterou může předmět od typické polohy zaujímat (například jablko je v misce, v případě že se jedná o jablko, které je na vrcholu hromady, přičemž prostorově leží mimo (nad) miskou). *Typičnost* bere v potaz znalost v pozadí, která není přímo v prostorovém vztahu obsažena (auto je za domem - předpokládáme že auto je blízko domu, přestože to předložka nevyjadřuje). Obecně lze v tomto přístupu nalézt důraz na pragmatickou komponentu prostorových vztahů, která je pomocí klasického logického aparátu nezachytitelná. Přesto autorka uvádí, že ani její kategorizační systém nedokáže zachytit veškeré výjimečnosti objevující se v přirozeném jazyce v oblasti prostorových vztahů.

5.4.1.3 Region connection calculus

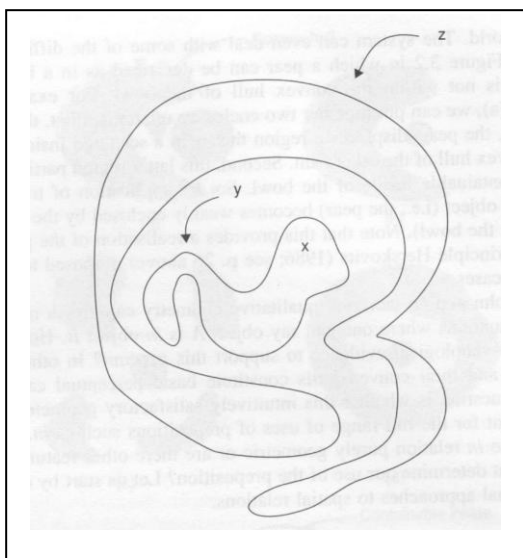
Další metodou reprezentace prostorových vztahů je RCC (Region connection calculus). Jedná se o kvalitativní vyjádření prostorových vztahů, podobně jako zachycují Allenovy kvalitativní vztahy časové události (Allen, 1983). Základními parametry definice objektů je jejich propojení a konvexnost (Cohn, 1996). První parametr zachycuje způsob propojení objektů, druhý definuje konvexní obal, jako nejmenší vypouklý útvar okolo daného objektu. Pomocí rozdílu ploch mezi obalem a skutečným objektem můžeme zjistit, zda je daný objekt konvexní či konkávní, tzn. umožní nám rozpoznat konkávní „díry“ v objektu. Konvexní obal slouží jako idealizovaná hranice, která rozliší zda došlo k propojení objektů. Obal lze použít i pro určení vztahu pro skupinu předmětů (například jeden ostrůvek leží uvnitř skupiny ostatních, kterou definujeme pomocí konvexního obalu). Pokud aplikujeme tato pravidla na 3D objekty, je třeba stanovit dodatečná opatření pro některé prostorové vztahy, např. díra v předmětu apod. Otázkou zůstává, zda je použití konvexních obalů psychologicky plausibilní.



Obr. 14 Ukázka konvexního obalu pro 2D objekty, který lze použít i pro skupinu objektů (c). V případě 3D předmětů se setkáváme s obtížemi, jelikož předmět může obsahovat část (ucho poháru), které nelze konvexním obalem popsat (díra je uvnitř ucha, ale přitom vně poháru).

5.4.1.4 Vizualní rutiny

Ullman (1996) postulují vizuální rutiny jako abstraktnější úroveň zpracování vizuálního počítka. Jedná se o následné zpracování primárního vizuálního vstupu. Na příkladu (viz Obr. 15), ze kterého nelze na první pohled posoudit prostorové vztahy dokazuje, že je třeba abstraktnějšího zpracování informace, abychom dokázali určit vzájemnou souvztažnost. V podstatě upozorňuje na nutnost zpracování vjemu na více úrovních, při kterém dochází k syntéze informací z úrovně nižších.



Obr. 15 Některé vzájemné prostorové vztahy nelze odhadnout pomocí jednoduchých mechanismů a je nutné použít zpracování na abstraktní úrovni (např. posouzení, která z částí x,y, a z je uvnitř) (Ullmann, 1996).

5.4.1.5 Teorie prostorových vzorů

Tento přístup je založen na požadavku společného rámce pro prostorové vztahy a jejich lingvistické označení (Logan, Sadler, 1996). Jedná se o aplikaci jazykových mřížek na vizuální scénu, přičemž mřížka vymezuje, která část scény je pro dané prostorové označení adekvátní. Použití mřížek znamená digitalizaci a rozdělení spojitého prostoru na jednotlivé části, čímž dochází ke sloučení symbolické a subsymbolické roviny. Jednotlivé části mřížky jsou zároveň prostorovým i lingvistickým vymezením daného označení, což může vést k obtížím o oblasti reprezentace významu.

A	A	A	G	A	A	A
A	A	A	G	A	A	A
A	A	A	G	A	A	A
B	B	B	■	B	B	B
B	B	B	B	B	B	B
B	B	B	B	B	B	B
B	B	B	B	B	B	B

Obr. 16 Ukázka prostorového vzoru „nahoru“. Správná poloha je označena (G), akceptovatelná (A) a špatná (B). Použitím vzoru na konkrétní situaci můžeme odvodit polohu objektů.

Přínosem přístupu je aplikace schémat pro specifické události, tedy zdůraznění, že nelze univerzálně stanovit prostorový vztah jediným schématem, ale je nutné aplikovat více schémat pro speciální případy určitého prostorového vztahu. Jejich počet je omezen počtem jazykových označení, které nekorespondují s prototypickým vyjádřením daného vztahu (například jablko je v misce i pokud je nahoře na hromadě jablek a miska jej neobklopuje). Výhodou systému je schopnost postihnout polysémický vztah objektů. Pokud existuje více významů prostorového vztahu, aplikujeme na něj odlišné vzory.

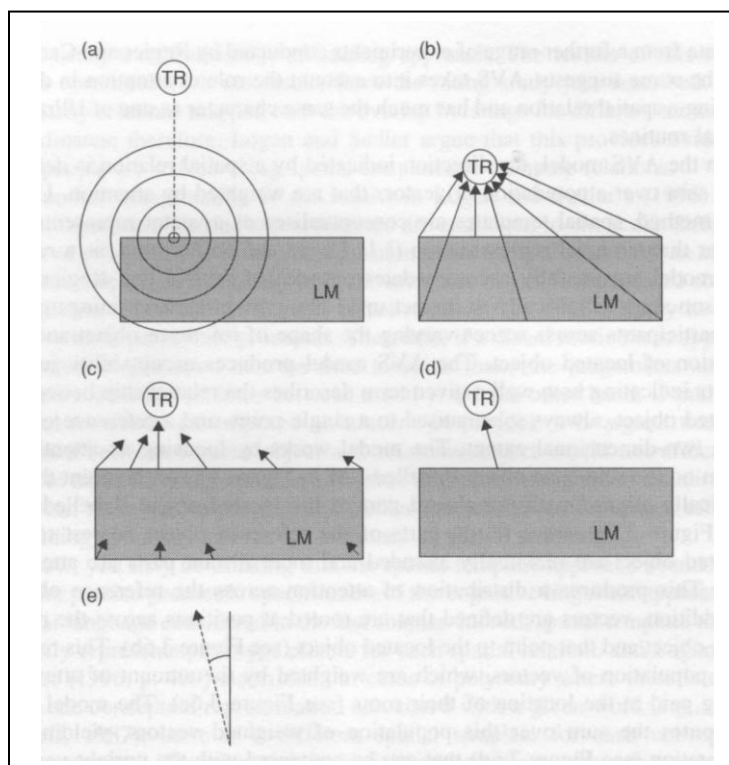
5.4.1.6 Model sum vektoru pozornosti

Regier ve svém rozšířeném modelu (Regier, Carlson, 2001) využívá pro odhad prostorového vztahu většího počtu posuzovacích dimenzí. Kromě *orientací center mas a blízkých bodů* je to *vzdálenost objektů* a také užití *vertikály* vycházející z nejvyšší části zkoumaného objektu.

Prostorové vztahy jsou odvozeny na základě *sum vektorů*, které jsou váženy pomocí *centra pozornosti*. Nejprve je vytvořeno *centrum pozornosti*, které je zacíleno na *základnu* (landmark), v místě které je nejbližší trajektoru, tedy objektu, ke kterému určujeme prostorový vztah. Ten je určován na základě populace vektorů směřujících od základny k trajektoru, přičemž největší váhu mají vektory nacházející se v blízkosti centra pozornosti. Z takto vážených sum vektorů je následně vypočítána orientace.

Model vykazuje velmi dobré výsledky na empirických datech. Jedná se o rozšířenou verzi původního modelu (Regier, 1996), který byl založen na *vektorech nejbližších bodů* a *center mas* objektů. Rozšíření spočívá právě v přidání *centra pozornosti* a většího počtu vektorů.

Celková činnost modelu závisí pouze na geometrických vztazích a nebere v potaz síly působící na objekty ani lingvistické specifika. Ty chce Regier zakomponovat do rozšířené verze modelu.



Obr. 17 Model vektoru sum pozornosti. Nejprve dochází a) zacílení pozornosti na základnu (LM), b) stanovení vektorů mezi základnou (LM) a objektem (TR), c) zvážení vektorů dle pozornosti a d) nalezení vektoru sum pozornosti (Regier, Carlson, 2001).

Regierův model byl kritizován autory RCC, protože dokázal pracovat pouze s 2D objekty. Jak již bylo řečeno, převod úlohy do třetího rozměru znamená nárůst složitosti během řešení, kterou Regierův systém není schopen zvládnout.

5.4.2 Extra-geometrické vztahy

Jestliže byla většina výše zmíněných přístupů založena na geometrických vztazích, podíváme se nyní podrobněji na druhou vrstvu Coventryho modelu a budeme zkoumat vlastnostmi sil působících na objekt. Základní teoretická východiska nacházíme již v pracích Gibsona (1979) či Michotta (1963), u kterých se setkáváme s pojmy jako affordance či funkční vztahy. Důraz je kladen na prostředí a způsob, jakým ovlivňuje naše reprezentace. Tyto názory byly dále

rozvíjeny v pracích Talmyho (1988) či Vandeloise (1994). Kromě prostorové pozice mají na způsob reprezentace vliv také síly, které na objekty působí. Jelikož nejsou pozorovatelné na úrovni geometrické, nazýváme je funkční či extra-geometrické. Například Talmy (1988) postulují agonisty a antagonisty a vztahy jejich vzájemné interakce. Přestože aplikuje své poznatky pouze na slovesa, je možné vidět přesah i v rovině prostorových předložek.

V souvislosti s extra-geometrickými vztahy se setkáváme s pojmem *reprezentační momentum*. Jedná se aplikaci dynamických faktorů a sil působících na statické objekty). Odkazy nalezneme i u Regiera (1996), který však tyto faktory do svého modelu nezahrnul. Podle některých badatelů (Coventry, 2004) je u Regiera princip zahrnut do faktoru *centra mas* a *center blízkosti*, ale pomocí těchto vektorů jsou síly nevyjádřitelné.

Důkazy pro existenci extra-geometrických vztahů nacházíme také v oblasti experimentální psychologie. Výzkumy prokázaly, že lidé posuzují vztah *v/uvnitř* jako jistější pro neživé objekty, než pro živé (Garodd et al., 1999). Jestliže je objektem uvnitř nádoby vážka, lidé to posuzují jako slabý příklad obsazení jednoho objektu druhým. V případě neživého objektu (mince) je míra jistoty větší. Jedná se o vyjádření míry *kontroly umístění*. V některých případech je však tento faktor méně relevantní, např. prostorový vztah *ostrov uvnitř souostroví*, ve kterém není objekt uzavřen ani ohraničen. Proto je potřeba pro některé případy použít již zmíněvaný Region connection calculus (RCC), který dokáže tyto výjimky postihnout (Cohn, 1996).

Kromě výše zmíněných požadavků při reprezentaci prostoru, musíme brát v potaz také informaci o funkci objektu, jehož polohu vyvozujeme, jelikož i tyto informace se podílejí na správném prostorovém označení. Obecně se tedy ptáme nejen na otázku *kde* objekt leží, ale také *jak* interaguje s okolím.

Byly provedeny experimenty v kosmu, kde je působení fyzikálních sil (např. gravitační) odlišné od pozemských podmínek. Podle výsledků experimentů Friederici a Levelta (1990) se ZO přizpůsobily podmínkám a vytvářely si referenční rámce pro vnímání prostoru podle polohy hlavy a centra retiny, což hovoří ve prospěch naučených schémat a rigidního způsobu vnímání gravitační síly. Vznikají také zajímavé posuny ve významu předložek, jelikož při neexistující gravitaci není možné označovat více objektů v misce vztahem *v*, protože neplatí vztah podpory jednoho objektu druhým.

Existují i další ukázky vlivu působení sil na prostorové označení prostoru. Příkladem může být vložení překážky mezi dva posuzované objekty znamenající změnu v prostorovém označení. Richard a Coventry (2003) prokázali, že pokud vložíme překážku mezi objekty, které mají funkční vztah (pošták a schránka), změní se referenční rámec a místo „pošták je před schránkou“ (vnitřní rámec), používají ZO relativní rámec, tzn. „pošták je nalevo od schrán-

ky“. V případě experimentu s letadlem, které přelétá nad domem, se opět prokázala závislost působení sil či funkčních závislostí. Pokud měly ZO posoudit polohu civilního letadla a bombardéru, určovali polohu v závislosti na typu letadla (Coventry, Mather, 2002).

Podobně jako geometrické vlastnosti jsou zachytitelné pomocí geometrických rutin, je možné zachytit extra-geometrické pomocí rutin dynamicko-kinetických. Jestliže pro první zmíněné používáme RCC či AVS model, je potřeba nalézt mechanismy pro reprezentaci těch druhých. Přestože nalézáme základy v Regierově přístupu, neexistuje dosud ucelený přístup k jejich reprezentaci v oblasti modelování. Ani model navržený v této disertační práci zmíněné mechanismy neobsahuje, přestože jsou pro reprezentaci prostorových vztahů nutné. Uvádím je zde pouze jako součást Coventryho teorie reprezentace prostoru. V dalších fázích vývoje modelu bych je rád do stávajícího řešení zakomponoval. Působení sil však do značné míry souvisí s rozpoznáním objektů (*what dráha*), což je z hlediska kognitivního modelování velmi obtížná úloha.

5.4.3 Znalost objektu

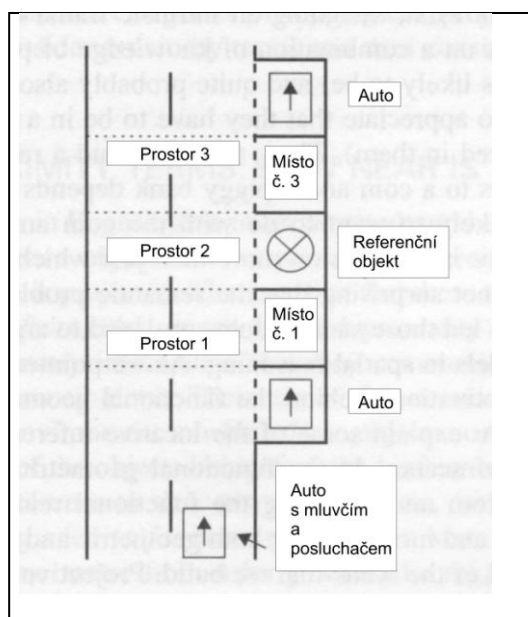
Na určení prostorového vztahu se podílí také znalost objektu, tedy jeho vlastnosti, které mohou mít vliv na správném označení polohy v prostoru. Jedná se vlastně o přechod od systému *where* do systému *what*. Při odhadování prostorového vztahu totiž vezmeme v potaz předchozí znalost o objektu, která nám pomůže určit správné označení. Pokud je šálek pod konvicí, je jeho prostorová poloha jiná, než pokud je muž pod deštníkem. Právě vlastnosti objektu určují správné prostorové označení. V některých případech může způsobit rozdíl pouhé užití jiného výrazu pro stejný objekt, např. talíř a nádobí. Předložku *na* použijeme ve vztahu s talířem, v případě nádobí však volíme jiný druh vyjádření. Znalost objektu souvisí také s pragmatickým použitím. O prostorovém vztahu objektů hovoříme z hlediska způsobu operace s ním. Lidé mohou sedět u piána nebo za stolem, přičemž pokaždé je ze zkušenosti jasné, jakým způsobem jsou s objektem v interakci a jaké je prostorové uspořádání (Coventry, 2004).

V potaz je nutné brát také specifika daného jazyka, např. používání předložek. Příkladem je použití spojky *on* v anglických větách *man is in/on the bus in/on the train*, ale nelze říci *man is on the car*. Výsledkem je použití spojky s ohledem na velikost dopravního prostředku, přičemž nelze přesně určit hranici, kdy je předložka adekvátní. V češtině podobný kontrast neexistuje, což poukazuje na lingvistické rozdíly napříč jazyky. Odlišnosti a výjimky nalézáme u ostatních jazyků a příkladů. Vliv znalosti objektu se objevuje také v souvislosti s tvarem objektu. Nacházíme vliv efektu tvaru objektu na přisouzení prostorového vztahu. Pokud má kon-

tejný jiný poměr stran, nazývají jej odlišnými jménem (šálek, pohár, sklenice) (Labov, 1973). Také míra uzavřenosti objektu hraje při posuzování roli. Navíc je možné zjistit změnu v užití předločky pomocí označení objektu z hlediska funkce či materiálu, ze kterého je vyroben (Ferenz, 2000).

Rozdíl mezi vlivem extra-geometrických vztahů a znalostí objektů může být z hlediska experimentálního zadání velmi nepatrný, přesto na výstupu obvykle získáme diametrálně rozdílné výsledky. V případě experimentu, kde byla mezi dva objekty umístěna překážka (pošťák a schránka), provedli autoři (Carlson-Radvansky, 1996) modifikaci, užitím referenčního objektu bez funkčního vztahu k posuzovanému (pošťák, ptačí budka). ZO označovaly obrázek „pošťák-schránka“ jako „pošťák před schránkou“, přičemž užívaly vnitřního referenčního rámce podobně jako při experimentu bez překážky, ale obrázek „pošťák-ptací budka“ označovaly „pošťák vlevo od ptačí budky“ pomocí relativního rámce, podobně jako by mezi předměty byla bariéra. Způsob interakce s objektem tedy hraje při posouzení také významnou roli. Zároveň bylo prokázáno, že funkcionalita objektů nemusí být reprezentována pouze pro stereotypní objekty (deštník nad hlavou), ale bývá posuzován jako adekvátní směrové označení pro objekty, které s danou funkcí běžně nespojujeme, např. *kufř nad hlavou* jako ochrana před deštěm (Coventry, 2004). Kontext je tedy důležitý, ale stereotypnost nikoliv.

Zajímavé výsledky pro efekt kontextu zadání úlohy a užití rozdílných referenčních rámců pro rozdílné objekty dosáhl Grabowski a Miller (2000). Jádrem experimentu je instrukce pro ZO, která je řidičem, aby zaparkovala dle situace vyobrazené na Obr. 18.



Obr. 18 Ukázka experimentálního zadání (Grabowski, Miller, 2000)

V případě, že je referenčním objektem auto a instrukcí zaparkovat před ním, zaparkovala většina ZO na místě 3. V případě stromu jako referenčního objektu a stejné instrukce, provedli parkování na místě 1. Jedná se o rozdíly mezi užitím *vnitřního schématu*, které je v případě auta podobný lidskému (lépe řečeno hadímu, kde označení „před“ znamená před čelem). U stromu podobné schéma není, a proto vytvářejí ZO referenční rámec na základě deiktického centra stromu, jehož přední část je orientována směrem k řidiči.

Rozdíly jsou způsobeny také vlivem sociálního kontextu. Jestliže byly ZO v pozici frekventanta autoškoly, zastavovali v místě 3 pro referenční objekt auto, v případě že byli instruováni jako řidiči vezoucí domů přítele, zastavovali v pozici 1. Pro referenční objekt strom nenastala žádná změna. Zmíněné výzkumy nám ukazují, že vlivy kontextu mohou vést k situaci, kdy ani jazyková instrukce nezaručí přenos významu prostorového označení a dostáváme se do situace, kdy neexistuje jednotná intersubjektivní shoda pro posouzení dané situace. Jelikož výše zmíněné příklady souvisí s rozpoznáním objektu, jsou pro současné možnosti simulace nedostupné. Přesto je bude nutné při vývoji budoucích modelů zohlednit. Pomocí experimentálních zjištění můžeme přesněji vymezit požadavky na systém a při jeho konstrukci využít empirické poznatky.

6 Shrnutí

Propracovali jsme se až k samotnému závěru teoretické části. Byly prezentovány všechny důležité aspekty, které se týkají realizovaného modelu. Pro větší přehlednost se je pokusím shrnout. Vycházíme z principů kognitivní sémantiky, přičemž základ tvoří Peircova teorie znaku. Jedná se o tvorbu znaku, který referuje k externímu objektu pomocí konceptu (*interpretant*) a arbitrární symbolické vrstvy (*representamen*). Konceptuální úroveň je vytvářena na základě perceptuálních vstupů, což je v souladu s perceptuální teorií kognice. Tím se přístup liší od Saussurovy teorie znaku a také od principů analytické filosofie. Oproti formálnímu symbolickému systému a formální sémantice je tento systém odlišný užitím perceptuálně získané konceptuální úrovně.

Realizovaný model je v teoretické úrovni řešením problematiky ukotvení symbolů zabývající se způsobem převodu perceptů do konceptuální úrovně a následným propojením s arbitrárním symbolickým označením. Jeden z možných způsobů převodu, který je inspirován teorií perceptuálního symbolu, spočívá v tvorbě multimodálních reprezentací, tzn. integrací informací z různých senzorických modalit do společného rámce. Podobnou inspiraci nacházíme také v teorii duálního kódování. Systém nejprve primárně reprezentuje auditivní a vizuální informace a následně je integruje do společné multimodální vrstvy.

Konkrétní oblasti aplikace je reprezentace prostoru. Přestože se objevují teorie o privilegovanosti osových systémů (Hayward, Tarr, 1995), popřípadě privilegovanosti kategorií mimo osy (Huttenlocher, 1991), vycházím při konstrukci z předpokladu, že prostorové kategorie vznikají na základě časté frekvence výskytu, tzn. ortogonální osy vznikají způsobem interakce člověka s prostředím a není nutné je fixovat pomocí apriorních předpokladů. Použití rozdílných referenčních systémů (absolutní, relativní, vlastní) není do modelu zakomponováno. V prvotní fázi je vytvořen pouze jeden systém, jehož centrem je střed senzoru, který slouží k odvození prostorového vztahu. V rovině teorie vychází model z Coventryho FGR teorie (2004), která odvozuje prostorové vztahy na základě geometrických, extra-geometrických vztahů a také znalosti objektu. V počáteční fázi systém řeší pouze absolutní geometrické prostorové vztahy dvou objektů, které nevyžadují rozpoznávání tvaru. V další fázi bude model vyvíjen v souladu s Coventryho teorií, pomocí přidávání dalších modulů, schopných reprezentovat relativní vztahy a také extra-geometrické informace nutné k řešení složitějších úloh.

Z hlediska privilegovanosti lingvistické a nelingvistické reprezentace se přikláním ke kombinaci zmíněných teoretických poznatků. Ani jeden systém není privilegován, multimodální reprezentace je tvořena pomocí obou. Přestože je pro strukturaci důležitá auditivní (lingvistic-

ká) informace, která je jednoznačná na rozdíl od vizuální, jejíž kategorie se překrývají a jsou zatíženy mírou neurčitosti, pro rozpoznávání objektů tomu může být opačně, jelikož pomocí jednoznačného tvaru lze odlišit slova, která jsou v symbolické úrovni identická, přestože nabývají odlišných významů (polysémie).

Architekturou pro realizaci modelu byly zvoleny samoorganizující se neuronové sítě a metodou *učení bez učitele*. Důraz je kladen na *učitelský přístup*, jehož cílem je co nejmenší zásah designéra. Právě proto jsou neuronové sítě vhodnou volbou. Také užití *učení bez učitele* zaručuje, že systém je vytvářen, a jeho reprezentace strukturovány, na základě informací z prostředí, což omezuje fixnost architektury a reprezentačního systému. Poté, co bude model rozšířen o motorickou část a systém efektorů, které se na základě učení naučí mapovat své reprezentace do multimodální vrstvy, stane se systém plnohodnotným zástupcem enaktivního přístupu ke kognici.

Protože v oblasti kognitivních věd neexistuje jednota přístupů ke konstrukci umělých inteligentních systémů, a jednotlivé proudy používají rozličná teoretická východiska, shrneme si před praktickou částí na jakých přístupech je model založen.

<i>Teoretický rámec</i>	Symbolický přístup (analytický)	Konceptuální přístup (empirický)
<i>Forma reprezentací</i>	Symbolická	Subsymbolická
<i>Způsob realizace</i>	Kognitivismus	Konekcionismus
<i>Existence reprezentací</i>	Nonreprezentacionismus	Reprezentacionismus
<i>Způsob tvorby systému</i>	Designérský	Učitel'ský

Tab. 3 *Tabulky použitých přístupů (jsou vtištěny tučně)*

II. Praktická část

7 Východiska pro návrh modelu

V oblasti aplikované umělé inteligence narážíme při tvorbě systému na problém, který se týká jejich schopnosti interagovat s okolním prostředím. V posledních desetiletích došlo k rozvoji vtělené robotiky, která se snaží umělé systémy propojit s prostředím pomocí dostatečné senzorické výbavy. Potíže nastávají při zpracování dat ze senzorů. Stále se nedaří odhalit mechanismy vedoucí k tvorbě interní reprezentace dostačující pro následnou rekonstrukci vjemu či jeho části bez přímé přítomnosti objektu. Současné výzkumy jsou zaměřené na hledání kompromisu mezi mohutností a efektivností reprezentace, jelikož systémy vytvářející reprezentace ze senzorických vstupů dokážou překonat obtíže související s čistě symbolickým způsobem reprezentace, tedy vyřešit problém ukotvení symbolů (Harnad, 1990).

V současné době je pozornost výzkumníků zaměřena na první fázi tvorby senzorických reprezentací, což je spojeno s výzkumem procesů kategorizace. Cílem je tvorba tříd a hierarchií tříd sloužících umělému systému k optimální reprezentaci okolního prostředí. V průběhu zpracování senzorických informací je nejobtížnější fází převod získaných dat do reprezentační úrovně zachycující konstantní vlastnosti a znaky a vytvářející konceptuální úroveň schopnou reprezentovat význam a následně ji propojit se symbolickou vrstvou.

7.1 Inspirační zdroje

Badatelé přistupují k problematice ukotvení odlišnými technikami, takže výčet modelů *groundingu* je v současné době rozsáhlý. Autoři se zabývají reprezentací jednotlivých objektů, rozlehlých prostor (Kuipers, 2007), sloves (Siskind, 2001) apod. Při návrhu vlastního modelu mne nejvíce inspirovaly níže zmíněné práce.

Důležitým aspektem při tvorbě systému je volba vhodné architektury. Např. Worgan (2006) se snaží pomocí systému tvořeného umělým uchem a hlasivkami a vybaveného dvěma samoorganizujícími se mapami (SOM) pro recepci a produkci řeči implementovat systém, který je na základě náhodného mapování a následné reorganizace schopný přijímat i produkovat řeč. Použití samoorganizace je v procesu ukotvení velmi efektivní, protože nám umožňuje zachování strukturální topologie vstupu pomocí *učení bez učitele*. Pokud použijeme SOM mapy pro rozdílné vstupy, můžeme zachovat a reprezentovat jejich specifické vlastnosti. Daný přístup

překonává námitky, o kterých jsme se zmínili v souvislosti konceptuálním empirismem, kritikou reprezentacionismu a enaktivním přístupem. Jedná se o využití informační bohatosti a stálosti prostředí při vytváření jeho interních (mentálních) reprezentací. Je to právě struktura prostředí, na jejímž podkladě jsme schopni vytvořit reprezentace. Systém se tedy neorganizuje sám, pouze se zformuje na základě svých vstupů z prostředí. Překonáváme námitky související s potřebou apriorních konceptů (koncepty lze odvodit z prostředí).

Pro učení se na základě podnětů z externího prostředí je nutné, aby architektura dokázala dostatečně reprezentovat externí objekty a byla schopná se v případě potřeby rozšiřovat. Ilustrativním přístupem je agent, napojený na server umožňující celoživotní učení, sloužící jako paměťový sklad pro reprezentaci objektů (Seabra Lopes, 2006). Systém je schopen pomocí *učení s učitelem* vytvořit libovolně velký slovník, přičemž každý objekt obsahuje konceptuální reprezentaci získanou z perceptuálních dat. Možnosti zvětšujícího se slovníku jsou pro lingvistickou či auditivní reprezentaci zásadní. Robustnější verzi předchozího modelu je architektura CELL (Roy, 2002), která zajišťuje mapování mezi obrazovým a zvukovým vstupem, přičemž používají odlišné mechanismy ukotvení. V současné době chtějí autoři využít architekturu ke zpracování obrazových a zvukových informací během prvních třech let vývoje dítěte. Projekt s názvem Speechome (Roy, nepublikováno) se stane nejrozsáhlejším pokusem o vytvoření systému, schopného reprezentovat okolní prostředí na základě kategorizace perceptů.

Nyní můžeme přistoupit k samotné definici navrhovaného systému. Pro jeho základ sloužily výše zmíněné modely ukotvení, které byly doplněny a rozšířeny. Architektura využívá jako základní prvek pro vizuálního i auditivního reprezentační vrstvu vlastnosti samoorganizujících se map (Kohonen, 1989). Pro vizuální systém použil zmíněný typ sítí již Riga (2004). V jeho případě však mapy uchovávaly invariantní reprezentace objektu a barev, nikoliv prostorových vztahů. V navrhovaném modelu jsou navíc samoorganizující se mapy (SOM) použity jak pro vizuální informace, tak pro auditivní vstupy. Právě reprezentace zvukového vstupu, který má strukturu jazykové promluvy a lze jej nazírat jako symbolický, je slabinou předchozích modelů. Jak Riga (2004), tak Regier (1996) tuto část systému vytvářejí redukcionisticky. Regierův model auditivní vrstvu vůbec neobsahuje. Prostorová (jazyková) označení jsou výstupní vrstvou neuronové sítě, což můžeme interpretovat různě. Model buď symbolickou vrstvu apriorně obsahuje, nebo dochází k auditivnímu předzpracování a slova můžeme reprezentovat jako prostorové označení kategorií, které nemají žádné auditivní vlastnosti. Tím ale pomíjíme jejich potenciál a výhody. Rigův model auditivní vstup obsahuje, ale je reprezentován jako arbitrární řetězec, asociovaný na perceptuální kategorie. Takový způsob ukotvení opět plně nevyužívá potenciál auditivního vstupu.

7.2 Návrh systému

Oblast vnímání prostoru jsem zvolil z několika důvodů. Z neuroanatomického hlediska je informace o poloze a vlastnostech objektů zpracovávána odděleně (Ungerleider, Mishkin, 1982), přičemž samotné rozpoznávání polohy objektu je pro následné zpracování fundamentální. Vytvořením modelu reprezentace prostorových vztahů můžeme danou oblast zkoumat samostatně a následně ji integrovat s rovinou rozpoznávání objektu. Model je zároveň možné analyzovat v kontextu výzkumu ukotvení symbolů, tedy převodu perceptuální roviny do roviny symbolické. Jelikož existuje mnoho výzkumů či modelů prostorových vztahů, máme zajištěné podmínky pro komparaci navrhovaného řešení.

Mezi nejznámější modely ukotvení prostorových vztahů patří Regierův z roku 1996. Použil síť s více moduly, ve které dochází k ukotvení prostorových vztahů na základě učení s učitelem. Model realizovaný v disertační práci je odlišný v návrhu architektury a způsobech zpracování vstupů. Oproti jiným řešením (Regier, 1996; Kuipers, 2007) přijímá systém jazyková označení jako auditivní vstup, který je primárně reprezentován a následně v multimodální vrstvě integrován s vizuální informací. Jak bylo zmíněno výše, nepředpokládáme, že symbolická úroveň reprezentace je systému daná apriorně a fixně, ale stejně jako perceptuální informaci se jí systém učí. Sémantika takového systému je pak tvořena jako kombinace senzorických vstupů, přičemž využíváme jejich společného potenciálu. Zpracování auditivního vstupu umožňuje tvorbu jednoznačných kategorií (jazyková označení), které jsou následně kombinovány s „nejednoznačnou“ vizuální informací (prostorové kategorie). Abstraktní rovina ve formě společné multimodální reprezentace je tvořena procesy učení ze dvou odlišných vstupů. Oproti klasickému přístupu, postulujícím fixní abstraktní symbolickou rovinu (apriorně definovanou designérem), je v tomto modelu možné multimodální reprezentaci měnit na základě učení. V modelu je „významem“ reprezentace auditivních, vizuálních a multimodálních kategorií. Multimodální kategorie tvoří společnou úroveň integrující vlastnosti obou předchozích.

Při návrhu systému vycházím také následujících úvah. Jazyk a obraz (či ostatní perceptuální informace) jsou zpracovávány odlišným způsobem. Forma jejich prvotní reprezentace musí zohledňovat odlišnosti ve způsobu jejich akvizice. Cílem je systém, ve kterém se odlišné kódy (reprezentace) prolínají, fungují synergicky a dokážou doplnit slabosti předchozího systému. Nezbytností se ovšem stává potřeba nalezení mechanismu vzájemného překladu mezi jednotlivými systémy. V psychologii se zabýval společným způsobem reprezentace už A. Paivio (1986), ale nedokázal problém definovat až do úrovně vhodné pro modelování (Vavrečka,

2006b). Abychom mohli přistoupit k integraci reprezentací do společného rámce, je třeba nalézt odpovědi na otázky, které přestože zní triviálně, jsou podstatné pro efektivní způsob ukotvení.

Jaké jsou základní rozdíly reprezentace obrazu a jazyka?

Ve kterých oblastech je jeden efektivnější než druhý?

Ve kterých oblastech jsou zásadním způsobem omezené (omezující)?

Jakým způsobem můžeme integrovat reprezentace z odlišných subsystémů?

Odpovědi se pokusím nalézt v architektuře realizovaného systému. V metaforické rovině jej můžeme považovat za narozené dítě, které nezná významy okolního světa, ale je vybaveno dostatečnými senzorickými a reprezentačními mechanismy, aby dokázalo okolní svět pochopit, a našlo způsob, jak jeho auditivní a vizuální vjemy navzájem korespondují.

7.3 Způsob reprezentace

Z toho vyplývá, že prezentovaný model je založen na zpracování auditivní a vizuální modality. Jelikož auditivní informace má vlastnosti symbolické úrovně (je arbitrární a diskrétní, tzn. jednotlivé slova jsou jednoznačně identifikovatelná) a vizuální informace je naopak perceptuální (obsahuje informaci o externím prostředí a je velmi neurčitá a *fuzzy*), vznikne jejich integrací multimodální reprezentace, ve které jsou spojeny vlastnosti obou.

Model tedy dokáže vytvořit primární reprezentace jednotlivých senzorických vstupů a následně společnou reprezentaci. Z hlediska teorie zpracování informace se jedná o výpočetně náročné procesy, přičemž je obtížné určit algoritmickou posloupnost zpracování. Proto je použito metody modelování pomocí učících se neuronových sítí, které umožní umělému systému nalézt vhodné algoritmy na základě adaptačního mechanismu. Při adekvátním nastavení počtu neuronů a metod učení můžeme docílit reprezentační úrovně blízké lidským schopnostem.

Dalším požadavkem při návrhu je tvorba dostatečně robustních reprezentací sloužících při řešení úlohy. Vlastnosti obsažené v reprezentacích by neměly „mít mezery“, které systém během řešení následných úloh nedokáže přeskočit nebo obejít. Reprezentace musí obsahovat konceptuální úroveň abstrahující podstatné vlastnosti, které jsou v případě vnímání prostoru subsymbolické. Následné myšlenkové operace jsou „sémanticky“ omezovány reprezentovanými vlastnostmi okolního světa (získanými perceptuálně).

System je navržen tak, aby si mohl sám vytvořit dostatečný počet překrývajících se kategorií, schopných pokrýt prostor v rozsahu senzorických vstupů. Pro reprezentaci kategorií používám architekturu, která se v průběhu učení adaptuje na topologickou orientaci prostoru, popřípadě fonetické vlastnosti auditivního vstupu.

7.3.1 Auditivní vstup

Podívejme se nejdříve podrobněji na možnosti, jak zareprezentovat auditivní informaci. U předchozích badatelů (Riga, 2004; Cangelosi, 2006) se jednalo o sekvence arbitrárních symbolů, které nahrazovaly slova. V realizovaném modelu jsem se od takového přístupu odklonil a uvažoval o reprezentaci auditivní informace jako n -dimenzionálního vektoru fonetických vlastností. Pro stanovení optimálního počtu dimenzí jsem vycházel z následujících úvah.

V případě jednodimenzionálního vstupu zdůrazňujeme arbitrárnost symbolů a blížíme se výše zmíněným přístupům. Abstrahujeme od schopností slov být asociovány (vytvářet strukturu na základě podobnosti). Tím mám na mysli schopnost vybavit si slova s podobnou strukturou (pes, les, ves) bez velké počítačové a časové náročnosti. Pouze jednodimenzionální reprezentace neumožňuje vznik asociace na základě zvukové (tvarové) podobnosti ani na základě kauzální blízkosti (výskyt slov vedle sebe ve větě). Taková reprezentace není psychologicky plausibilní a v dalších úrovních zpracování zvyšuje výpočetní náročnost.

V případě dvojdimenzionálních vstupů, které byly během realizace modelu použity, už můžeme reprezentovat rozdílné vlastnosti a vytvářet asociční vazby. U vícedimenzionálních vektorů lze zachytit další akustické, ale i lingvistické vlastnosti, např. společný výskyt slov v promluvě. Přesto nejsou lingvistické vlastnosti v základní verzi využity. Důvodem je větší přehlednost a snadnější interpretovatelnost modelu z hlediska procesu ukotvení symbolů.

V navrhovaném modelu se snažím pracovat se zvukovým vstupem podle následujících pravidel. Podobně jako je vizuální vstup kategorizován podle topologických vlastností, je audio vstup také nejdříve reprezentován topologickou mapou. Vycházím z předpokladu, že ze zvukového vstupu, který je strukturován jako jazyková promluva, lze bez apriorní sémantické či syntaktické znalosti získat mnoho informací, které nám mohou pomoci při tvorbě primární reprezentace. Pokud vycházíme z předpokladu, že malé dítě dokáže v prvních fázích analyzovat promluvu na úrovni slov, nabízejí se následující možnosti. Bez předchozí znalosti lze členit slova podle zvukové podobnosti, podle délky, či podle frekvenční křivky, neboť informace je v promluvě bezprostředně obsažena. Přestože se může jevit, že vlastnosti neusnadňují proces integrace zvukových reprezentací s vizuálními, není tomu tak. Jestliže přistupujeme

k otázce ukotvení symbolů s požadavkem minimálního zásahu designéra (tzn. žádné externí interpretace), je nutné využít vlastností vstupů při tvorbě primárních reprezentací v co největší míře. Reprezentujeme-li auditivní vstup samostatně v primární auditivní reprezentaci, bylo by krátkozraké redukovat veškeré vlastnosti, které jej charakterizují, jelikož bychom při použití principu samoorganizace ztratili informace sloužící k tvorbě jednoznačných kategorií. I když se nabízí možnost reprezentovat jazykovou promluvu jako arbitrární shluk symbolů, pokusíme se uvažovat odlišně. Přestože samotná promluva není sémantická (k ničemu neodkazuje), ve způsobu svého kódování obsahuje informace (jak bylo zmíněno výše), které nám pomohou při její reprezentaci, ale i v následných transformacích. V našem zjednodušeném modelu, vyjádříme jednotlivé slova 2D vektorem vyjadřujícím frekvenční křivku a časovou délku daného označení.

7.3.2 Vizuální vstup

Vizuální informace jsou reprezentovány také pomocí SOM sítí. Takto vytvořená vrstva je schopna pomocí laterálně propojených neuronů vytvořit prostorově uspořádaná prototypická centra, citlivá na vstupy určitých hodnot. Vstupem je dvourozměrný vektor reprezentující prostorovou polohu předmětu. V modelu abstrahuji od prvotního zpracování informace, které lze snadno vytvořit pomocí umělé retiny. Jedná se pouze o architektonické zjednodušení. Systém by mohl mít na vstupu umělou sítnici, na kterou by přicházel obraz z prostředí, ale pro simulaci je postačující vektorový vstup, jelikož nedochází k redukci informace o signálu z prostředí.

V realizované verzi je také oproti Regierovu modelu (1996) odlišně řešeno zpracování sekvencí. Na rozdíl od identifikace dynamické scény ve „vrchních“ vrstvách sítě, je v modelu vytvořen duální systém zpracování již v prvotní úrovni. V případě statické scény je informace zpracována samoorganizující se mapou citlivou na polohu objektu. V případě dynamického scény jí zpracovává RecSOM mapa (Voegtlin, 2002) obsahující neurony citlivé na pohyb. Ty pomocí rekurzivity dokážou klasifikovat dynamické prostorové vztahy (nad, pod, ven, skrz, okolo).

7.3.3 Multimodální vrstva

Jádrem systému je multimodální vrstva, která zajišťuje proces identifikace, tedy tvorbu jednoznačné kategorie a zároveň její reprezentace. V případě modelu se jedná o nalezení vhodného propojení mezi primárním auditivním a vizuálním výstupem, který je multimodální vrstvě prezentován v jeden okamžik. Cílem učení na základě těchto dvou informací je nalézt kategorie, které reprezentují společné vlastnosti vstupů. Rozdíl oproti klasickému Harnadovu přístupu je podrobně rozebrán v závěrečné kapitole. V teoretické rovině je výsledkem analýzy modelu odhalení principu *privilegovanosti jednoznačného vstupu* (Vavrečka, 2008), což je nový přístup k řešení otázky privilegovanosti lingvistické a nelingvistické reprezentace, podle kterého není privilegovaná reprezentace apriorně daná, ale závisí na aktuálních vstupech z prostředí. V obecné rovině se jedná o příspěvek do diskuse ohledně platnosti Sapir-Whorfovy hypotézy (Whorf, 1956).

7.4 Učitelský přístup

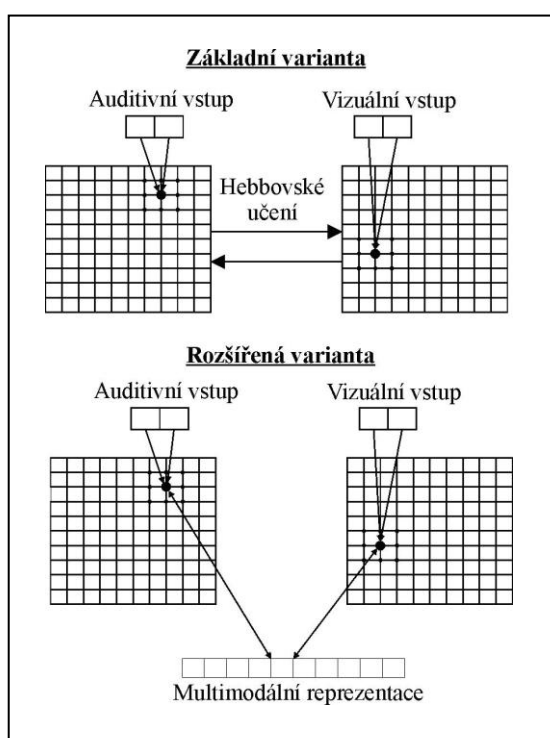
Jak bylo zmíněno teoretické části, je cílem realizovaného modelu zajistit takový způsob ukotvení symbolů, který není v rozporu s *učitelským přístupem* (Ziemke, 1999). Rozlišení způsobů tvorby komponent, které jsou pro systém předprogramované a těch, které systém může sám vytvořit a používat, je otázkou častých sporů v oblasti kognitivních věd. Obecně lze říci, že se jedná o nalezení takového poměru mezi *designérským* a *učitelským přístupem*, abychom ve výsledku dospěli k systému, který je schopen smysluplně manipulovat se svými reprezentacemi, ale i modifikovat některé své mechanismy na základě a v závislosti na učení.

Stručně lze charakterizovat *designérský přístup* jako tvorbu agentovy funkce a *učitelský přístup* je pak založen na agentovi, který se svou funkcí učí. O tom, že *učitelský přístup* se dá nahlížet různě, bylo již zmíněno v souvislosti se *Z kritériem*. Tadeo (2005) hovoří o nemožnosti procesu učení se symbolům za pomoci externího pozorovatele (učitele). Externí pozorovatel ale nesmí zasahovat pouze do tvorby konceptů. V (čistě) symbolické úrovni již může fungovat jako učitel, protože bez jeho účasti není symbolická rovina přítomna (ve smyslu designérový produkce symbolů). Pokud bychom přesto vyžadovali striktní nepřítomnost sémanticky vybaveného učitele, lze takovému požadavku vyhovět. Alternativou je získávání symbolů z knih či jeho média. Během takového učení však docházíme k paradoxu učení se symbolického systému, bez znalosti k čemu přísluší. Kompromisem je návrh systému, který se naučí manipulovat se svým symbolickým systémem (např. přirozeným jazykem), na zákla-

dě asociačního učení podobného učení se ze slabikáře (asociace obrazu se slovem). V takovém případě je ukotvování symbolů podobné učení se hluchoněmých dětí psanému a čtenému projevu, aniž mají předchozí zkušenost se symbolickým systémem. Většina dětí se ale naučí znakovou řeč jako první, takže se nám nabízí pouze úzké pole působnosti při získávání podrobnějších a přesnějších informací ohledně této problematiky pomocí psychologických experimentů. Pro naše účely je postačující, že model vytváří strukturu svých vnitřních reprezentací na základě *učení bez učitele* a auditivní vstupy zastupující symbolickou vrstvu, jsou učený pomocí arbitrárních vektorů, které reprezentují pouze fonetické vlastnosti prostorových označení. Jediným designérským zásahem je volba typu sítí a způsobu učení, bez kterých by systém nebylo možno realizovat.

8 Model 1 – statické prostorové vztahy

První model slouží jako základ celého systému a měl by být schopen klasifikovat statické scény. Při jeho realizaci lze použít metodu Hebbovského učení, které posiluje propojení auditivní a vizuální SOM na základě společného výskytu. S takovou architekturou se setkáme u modelu pro akvizici jazyka (Li, Farkaš, Mac Whinney, 2004). Druhým řešením je doplnění modelu o další vrstvu, která bude schopná vytvářet z výstupů primárních SOM map společnou multimodální reprezentaci. Základní schéma možných variant je na Obr. 19, přičemž pro realizaci byla zvolena druhá (rozšířená) varianta.



Obr. 19 Návrh architektury pro ukotvení prostorových vztahů

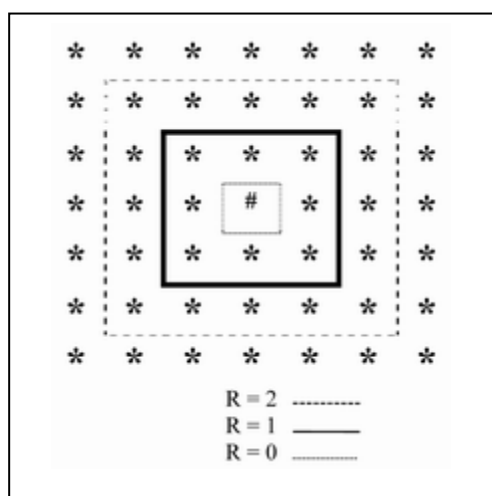
Ze samotného modelu je jasné, že přistupuji k problematice ukotvení symbolů z pozice, která explicitně nepředpokládá subsystemy pracující na čistě symbolické úrovni. Spíše se snažím o emergenci perceptuálně ukotvených elementů přirozeného jazyka, které pomocí vhodných mechanismů reprezentace získávají vlastnosti přisuzované symbolickému systému. Systém musí fungovat na základě ukotvených významů a nikoliv na základě externě dodávaných principů či mechanismů. Pro reprezentaci kategorií by měla být použita taková architektura, která by v průběhu učení zachovávala topologickou orientaci prostoru, kdy kategorie jednotlivých částí prostoru korespondují s externím prostředím. Přistupme nyní k zadání úlohy.

8.1 Zadání úlohy

Systém se musí naučit reprezentovat pět prostorových označení na základě auditivních a vizuálních vstupů. Jedná se o označení *nahore dole, vlevo, vpravo, střed*. Cílem je nalézt adekvátní počet neuronů pro jednotlivé části sítě umožňující modelu identifikovat jednotlivé prostorové označení. Systému jsou prezentovány pouze auditivní a vizuální informace, na jejichž základě je reprezentace vytvářena. Model je po natrénování testován na auditivních nebo vizuálních vstupech, na jejichž základě přiřazuje správnou multimodální kategorii. Zajímá nás, od jakého počtu neuronů dokáže systém správně kategorizovat.

8.2 Metoda

Pro řešení úlohy byla použita klasická SOM architektura (Kohonen, 1989), která je svojí schopností samoorganizace a shlukování vstupů s podobnými vlastnostmi vhodná pro aplikace rozhodování, rozlišování a třídění objektů. Jedná se o síť laterálně propojených neuronů, které jsou trénovány na základě *učení bez učitele*. Nejčastěji má struktura formu dvourozměrné čtvercové nebo obdélníkové mřížky, hexagonálního útvaru, někdy také jednorozměrného vektoru. Tvar uspořádání neuronů má vliv na volbu tzv. okolí neuronu R , které vymezuje jeho sousedy (sousední, nejbližší postavené neurony). V mřížkovém uspořádání neuronů (čtvercové nebo hexagonální) je velikost okolí rovná počtu „řad“ neuronů od centrálního neuronu – viz Obr. 20. Váhy každého neuronu definují jeho polohu v prostoru vah.



Obr. 20 Možné struktury uspořádání neuronů (*) s definicí okolí R vítězného neuronu (#)

Mřížce neuronů se postupně předkládají vektory vstupního signálu (x) tak, že se zvlášť porovnává rozdíl příslušných hodnot vektoru vah (koeficientů w) každého neuronu s hodnotami vektoru vstupního signálu. K vyjádření rozdílu se může využít různých algoritmů, ale nejčastěji se dává přednost výpočtu euklidovské vzdálenosti D , tj. součet rozdílů příslušných hodnot:

$$D = (x_1 - w_1)^2 + (x_2 - w_2)^2 + \dots + (x_n - w_n)^2.$$

Výsledkem je počet hodnot D , rovný počtu neuronů ve struktuře (např. 100 hodnot v matici 10 x 10 neuronů). Následně se vybere jediný neuron s nejmenším D a označí se jako tzv. vítěz (winner). Váhy tohoto neuronu totiž nejvíce ze všech odpovídají hodnotám právě předloženého signálu. Při předkládání prvního vstupního vektoru se jeho hodnoty porovnávají s náhodně vygenerovanými hodnotami vah (koeficientů) jednotlivých neuronů (počáteční náhodné nastavení). Váhy W vítězného neuronu se pak upravují (updatují), aby se co nejvíce přiblížily hodnotám právě předloženého vstupního vektoru (x). Využívá se vzorce:

$$W_{i \text{ nové}} = W_{i \text{ staré}} + \alpha (x - W_{i \text{ staré}}),$$

kde α je učicí koeficient vyjadřující rychlost učení (může nabývat hodnot 0 až 1, např. $\alpha = 0.6$), W_i je vektor vah (koeficientů) i -tého neuronu $W_i = [W_{i1}, W_{i2}, \dots, W_{in}]$ a x je vstupní učicí vektor $x = [x_1, x_2, \dots, x_n]$.

Při opakování dávky učicích vektorů (vstupů) nebo postupným předkládáním dalších nových dávek se učicí koeficient (rychlost učení) obvykle snižuje. Spolu s vítězným neuronem se mění i ty sousední v definovaném okolí R (viz Obr. 23). Jejich váhy se upravují stejným způsobem jako u vítěze, pouze s tím rozdílem, že koeficient α je nahrazen koeficientem β , přičemž platí $\alpha < \beta$. Při opětovném opakování dávky učicích vektorů se může provádět i snižování hodnoty okolí R až na $R = 0$, tzn. adaptuje se pouze vítěz.

Ve výsledku by se mělo dosáhnout stavu, kdy v maticové struktuře neuronů vznikne několik významných center, tzv. shluky, mezi nimiž se výrazně liší hodnoty vah neuronů. Počet shluků by měl být shodný s počtem odlišných vlastností nebo parametrů, které Kohonenova mapa našla v předložených dávkách učicích vstupních vektorů. To také znamená, že funkčnost mapy a neuronových sítí obecně, výrazně závisí na složení signálů a informací v učicích dávkách (vstupů z prostředí).

Pro jednodušší kontrolu a přehlednější dohled nad učením mapy se využívá grafického zobrazení shluků, které vyjadřuje vztahy mezi neurony v prostoru vah. Váhové vektory neuronů bývají zobrazeny jako body v dvojdimenzionálním prostoru, které zároveň tvoří centra shluků. Čáry představují přímky spojující váhové vektory sousedních neuronů.

Schopnost „neustálého doučování“ během činnosti umožňuje přizpůsobení funkce sítě při dlouhodobě trvajících změnách. Tím se blíží funkci našeho mozku, který též dokáže přizpůsobovat naše chování změně prostředí. Samoorganizace a samoučení je vhodná metoda z hlediska *učitelského přístupu*, jelikož veškeré procesy probíhají na základě vstupu z prostředí, kterým je přizpůsobována vnitřní reprezentační struktura sítě. Na SOM sítích je postaven celý systém.

8.3 Realizace

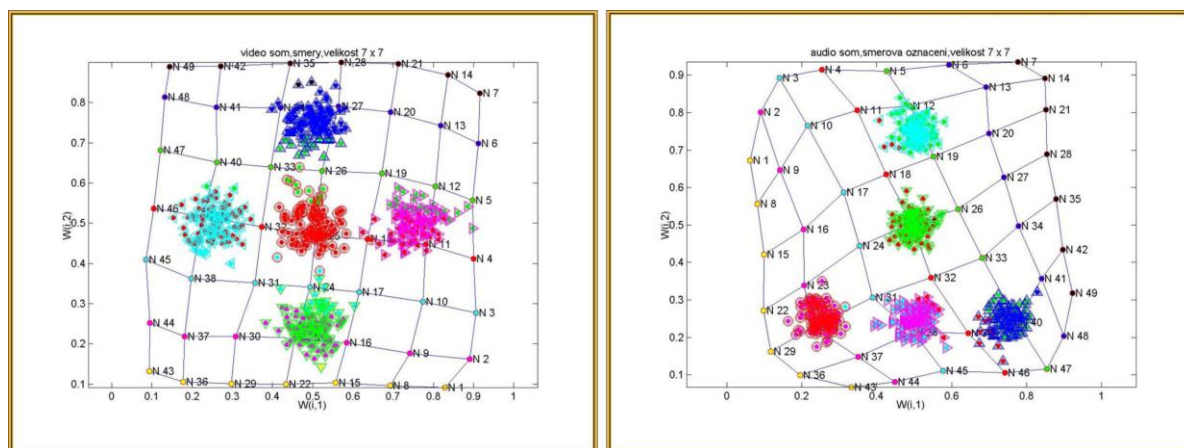
Pro realizaci modelu byl zvolen program Matlab umožňující modifikovat a doplňovat základní funkce systému a stupňovitě z nich vytvářet komplexní systém. Pomocí něj byla vytvořena celá struktura modelu. Využil jsem funkce z knihoven Neural Network Toolboxu a následně SOM Toolboxu, protože umožňuje lepší možnost nastavení parametru sítě.

auditivní vstup	2D audio vektor
vizuální vstup	2D vektor polohy
auditivní síť	SOM 2D mřížka
vizuální síť	SOM 2D mřížka
výstupní funkce	WTA, WTM
multimodální vstup	výstup auditivní a vizuální
multimodální síť	1D SOM nebo Neural Gas
výstup	WTA (pro ověření správné kategorizace)

Tab. 4 Přehled jednotlivých částí modelů a způsobů zpracování

Postup při realizaci byl následující. Nejprve jsem vytvořil vstupní informace pro trénování sítě. Jednalo se o sadu 100 vzorků pro jednotlivé prostorové označení, celkem tedy 1000 vzorků pro auditivní a vizuální vstupy. V případě vizuálního vstupu obsahuje jeden vzorek souřadnice x a y bodu v prostoru v rozsahu $(0;1)$. Centrum senzoru, ke kterému jsou odhadovány prostorové vztahy, leží uprostřed tohoto pole a má hodnotu $\langle 0.5, 0.5 \rangle$. Auditivní vstup je také dvojdimenzionální, přičemž každý vzorek obsahuje informaci o délce slova (osa x) a

frekvenční křivce slova (osa y). Při jejich grafické reprezentaci vidíme, že slova jsou podél osy x rozdělena na 1 slabičné (střed), dvouslabičné (dolů, vlevo, vpravo) a tříslabičné (nahoru). Podél osy y se slova dělí podle frekvenční křivky hlásek na „tvrdá“ (střed, vpravo, nahoru) a „střední“ (vlevo) a „měkká“ (dolů). Jednotlivá označení pak vytvářejí shluky, které demonstřují způsoby výslovnosti daného slova. Nejlépe je rozložení shluků vidět na Obr. 21.



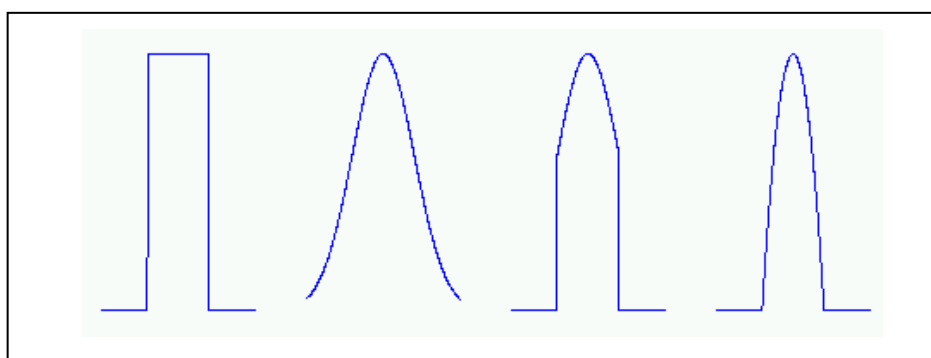
Obr. 21 Ukázka trénovacích vstupů pro vizuální (vlevo) a auditivní (vpravo) primární vrstvu. Vizuální vstupy se zpočátku nepřekrývají, s postupným vývojem modelu jsou shluky zvětšovány, až do podoby blízké reálným vstupům.

Nyní můžeme přistoupit k trénování sítě. Nejprve je inicializována, tzn. je nastaven počet neuronů, typ propojení, způsob a rychlost učení, stanoven počet okolních neuronů, které jsou modifikovány zároveň s vítězným neuronem a zvolen způsob jejich učení. Je vytvořena odpovídající síť s náhodně nastavenými váhami, ale fixním laterálním propojením neuronů. Poté je možné začít síť trénovat. Aby byla schopná detekovat vstupy v celém rozsahu (0;1), trénoval jsem auditivní i vizuální vrstvu pomocí 5000 náhodných vstupů, po dobu 100 epoch. Jednotlivé neurony vytvořily mřížku a jejich váhy se upravily do podoby, že systematicky pokrývaly celý prostor vstupu a vytvořily detekční a reprezentační systém, který dokáže adekvátně diferencovat vstupy z prostředí. Každý neuron se stává detektorem signálů z prostředí, citlivý na určité hodnoty. Pokud bychom auditivní síť trénovali pouze vstupy, které souvisí s prostorovým označením, nastavily by se váhy neuronů jen pro hodnoty, odpovídající pěti kategoriím (nahoru, dolů, vlevo, vpravo, střed). Následně by síť nedokázala dobře diferencovat nová označení. Proto je potřeba v počátku trénovat pro celý rozsah možných hodnot. Tento požadavek je psychologicky plausibilní, jelikož se během vývoje neučíme pouze prostoro-

vé označení, ale je nám prezentováno velkém množství odlišných slov, které jsme schopni diskriminovat a také reprezentovat. V případě vizuální vrstvy bychom mohli použít k tréninku pouze vstupy jednotlivých prostorových kategorií, jelikož pokrývají celý vstupní prostor, ale zvolil jsem také náhodné vstupy, aby byla zajištěná rovnoměrná distribuce hodnot.

Během úvodního tréninku je potřeba nastavit parametry sítě tak, aby rozložení vah neuronů pokrývalo celý vstupní senzorický prostor. Toho docílíme nastavením parametrů pro rychlost učení vítězného neuronu a jeho okolí. Pokud jsou tyto hodnoty příliš velké či malé, budou váhy neuronů chybně rozděleny a způsob jejich činnosti při detekci vstupních signálů není efektivní. To souvisí se špatným nastavením rychlosti učení a velikosti okolí vítězného neuronu. Pokud jsou hodnoty příliš velké, jsou váhy neuronů „stahovány“ jedna druhou a vytvářejí shluky, což neumožní nastavení vah kopírující vstupní hodnoty.

První fáze tvorby modelu proto souvisí s hledáním optimálních hodnot rychlosti učení a velikosti okolí. Okolo hodnoty rychlosti učení 0,3 dokázala síť po 100 epochách pravidelně pokrýt prostor možných vstupů. Důležité je také nastavení způsobu učení okolních neuronů. Jelikož lze tuto funkci v originálním Neural Networks Toolboxu špatně nastavit, byl pro další vývoj modelu použit SOM Toolbox, který je speciálně vyvinut pro tvorbu SOM sítí. Zde je možné zvolit parametr učení okolí jako křivku, která, se vzdáleností od vítězného neuronu klesá a snižuje se tedy rychlost učení. Můžeme definovat *funkci okolí*, zajišťující učení okolních neuronů, oproti klasickému SOM modelu, ve kterém je fixně daná rychlost učení α pro vítězný neuron a β pro okolní neurony.

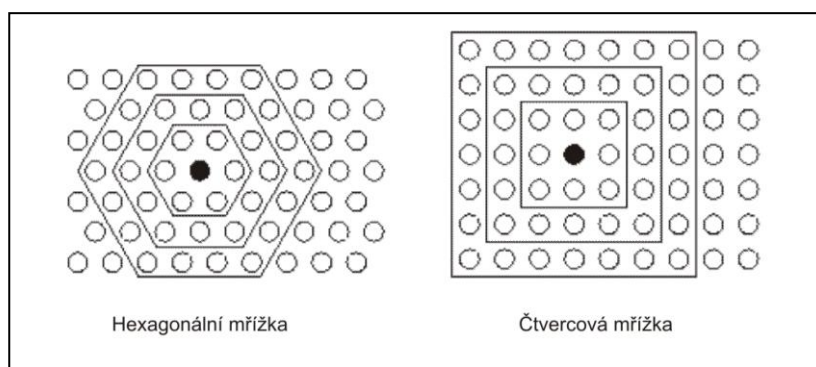


Obr. 22 Funkce učení v SOM Toolboxu. Jedná se o pokles hodnoty rychlosti učení v závislosti na vzdálenosti od vítězného neuron (střed symetrie křivky). Toolbox nabízí a) bubble, b) gaussian, c) cut gaussian, d) epanechnikov funkci.

Při tvorbě navrženého modelu se osvědčily funkce *gaussian* a *cut gauss*, při jejichž užití vykazovala distribuce vah neuronu optimální rozmístění a pokrývala celé rozmezí vstupů. V souvislosti s okolními neurony také volíme typ mřížky, která určuje vzájemné způsob pro-

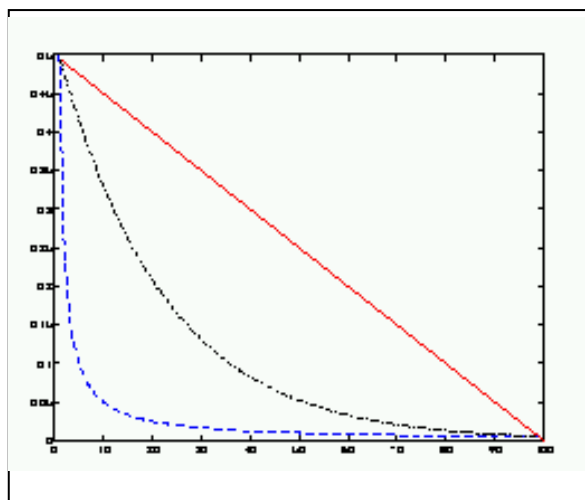
pojení sousedních neuronů. Pro dvourozměrné sítě se nejčastěji využívá čtvercová a hexagonální mřížka, které apriorně a fixně určují způsob propojení sousedních neuronů již během inicializace sítě. Z hlediska vzdáleností sousedních neuronů existuje mezi typem mřížek rozdíl. Jestliže použijeme čtvercovou mřížku, mají nejbližší sousední neurony rozdílnou euklidovskou vzdálenost. V případě použití hexagonální mřížky jsou neurony stejně vzdálené, což zajišťuje přesnější aplikaci funkce okolí, která určí, jaká bude rychlost učení okolních neuronů. Rozdíl ve vzdálenostech je patrný z Obr. 23.

Během tvorby modelu jsem vyzkoušel oba typy mřížek, ale rozdíly ve funkčnosti sítě jsem nenalezl. Důvodem může být náhodná distribuce tréninkových vektorů, která eliminuje vliv vzdálenosti sousedních neuronů.



Obr. 23 Vzdálenost sousedních neuronů dle typu mřížky.

Posledním faktorem, se kterým se v souvislosti s trénováním sítě setkáváme, je nastavení poklesu *rychlosti učení* v závislosti na počtu epoch. Jedná se o funkci, která se zvyšujícím se počtem epoch snižuje rychlost učení α , až do hodnoty 0 v poslední epoše, kdy již nedochází k modifikaci vah neuronu. Způsob poklesu je možné nastavit před začátkem učení. Kromě jediné křivky se také používá metoda, která rozdělí proces učení na dvě etapy. V první dochází k poklesu α až po určitou hodnotu a následuje *jemné ladění*, při kterém je hodnota α minimální a slouží k jemným pohybům nastavení vah neuronu, aby dokázaly co nejlépe detekovat signály v daném segmentu. Pro naše účely je postačující první metoda. Z predefinovaných křivek se nejlépe osvědčila metoda *power*, která je vyznačena na Obr. 24 černou přerušovanou čarou.

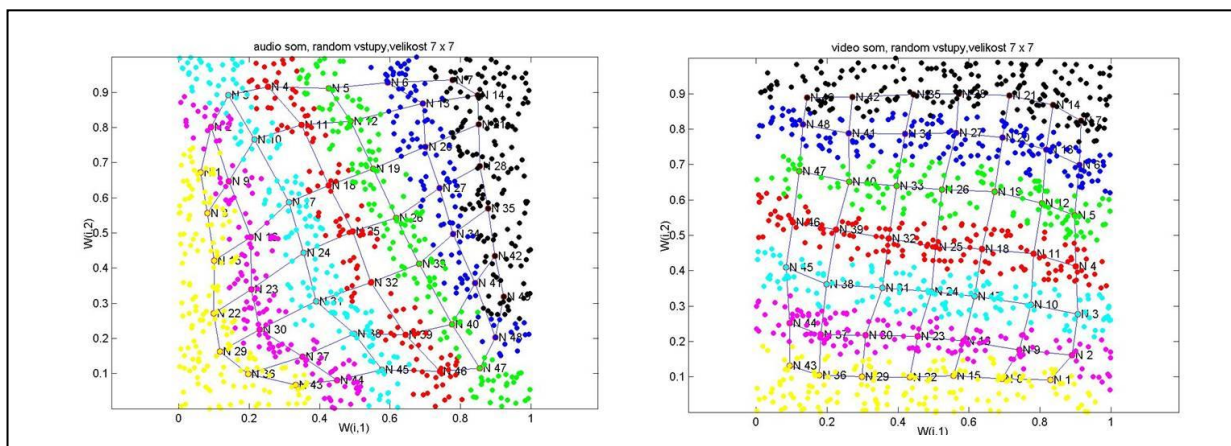


Obr. 24 Závislost rychlosti učení (osa y) na počtu epoch (osa x).

Tato křivka zajistí, že během prvních epoch dochází k velkým změnám vah neuronů, tzn. váhy rychle obsáhnou stavové pole vstupů a následně dochází ke zmenšení učebního kroku a váhy neuronu již vykazují pouze malé posuny a začínají v daných oblastech „usazovat“ a snaží se dosáhnout hodnot, kdy dokážou co nejlépe obsáhnout vstupy daného rozsahu. Pokud bychom rychlost učení α nesnižovali, síť nebude stabilní, jelikož by docházelo ke stálým úpravám vah a z toho plynoucí nestabilitě výstupů sítě. Pokud bychom nastavili hodnotu od začátku příliš malou, nelze během učení pokrýt pole vstupních hodnot, díky malé „pohyblivosti“ neuronů.

Toto jsou základní parametry, pomocí kterých ovládáme, jakým způsobem probíhá učení sítě a jak efektivně bude pracovat. Pokud tedy vytvoříme primární auditivní a vizuální vrstvu podle výše zmíněných parametrů, dostaneme v závěru síť neuronů, uspořádaných do dvojrozměrné mřížky, která pokrývá celý auditivní a vizuální vstupní prostor a dokáže reprezentovat patřičné vstupy. Správně natrénovaná síť vypadá podobně jako na

Obr. 25.



Obr. 25 Pokud je síť správně natrénovaná, pokrývají váhy neuronů (N) celý vstupní prostor. Výsledkem je čtvercová mřížka ve vizuální (vlevo) a auditivní (vpravo) vrstvě. Díky trénování náhodnými vstupy jsou téměř identické.

Nyní můžeme přistoupit k dalšímu kroku a prezentovat síti vizuální a auditivní vstupy související s prostorovým označením. To je jádro principu ukotvení spočívající v prezentaci dvou vstupů, které budou následně integrovány do společného rámce. Vizuální vstup je tvořen vektorem obsahujícím souřadnice bodu v prostoru. Auditivní informace obsahuje vektor, ve kterém jsou obsaženy údaje o fonetických vlastnostech slov. Oba údaje jsou prezentovány síti v jeden okamžik. Jedná se o způsob učení, kdy je systému prezentován objekt nebo vlastnost a přitom proneseno jeho/její jazykové označení. S podobným přístupem se setkáváme během dětského vývoje a jedná se o klasický přístup při akvizici slovní zásoby. Po prezentaci vstupů je (auditivní a vizuální) informace zpracovávána odděleně, přestože se jedná o architektonicky identické subsystémy.

Vizuální informace je nejprve detekována pomocí neuronů, které předtím byly vytrénovány na náhodných vzorcích. Neuron, který pokrývá prostor vizuálního vstupu, tzn. je vstupu svým nastavením vah nejbližší, se stává vítězem a jeho výstupní funkce posílá do multimodální vrstvy signál o hodnotě 1 (neuron detekoval signál a šíří dále informaci o jeho přítomnosti). Takový způsob zpracování a výstupní funkce se nazývá WTA (Winner takes all). Vítězný neuron jako jediný reaguje na signál výstupem, ostatní neurony mlčí.

Během vývoje modelu bylo použito také způsobu WTM (Winner takes most). V tomto případě reaguje vítězný neuron a na výstupní funkci posílá 1. Poté je zjištěn druhý nejúspěšnější neuron a ten posílá z výstupní funkce hodnotu menší než 1. Podle toho, jak velké okolí stanovíme, je vytvořena skupina vítězných neuronů, které mají podle pořadí klesající výstupní funkci. Oproti WTA, ve které se výstupní funkce skládá pouze z reakce jednoho neuronu, je

v případě WTM výstupní funkce složena z několika odstupňovaných výstupů. Z vizuální vrstvy je výstup posílán na vstup multimodální vrstvy.

Na identickém principu funguje auditivní vrstva. Vstupní vektor je podobný vizuálnímu, pouze obsahuje odlišné vlastnosti. Je zpracováván auditivní vrstvou, která byla trénována pomocí náhodných vstupů. Také zde je signál zpracován na základě WTA, popřípadě WTM a výstupní funkce posílá informaci do multimodální vrstvy.

V té dochází k integraci vstupů z auditivní a vizuální sítě. Samotná multimodální vrstva je koncipována jako jednorozměrná SOM, která obsahuje pět neuronů propojených do řady, což odpovídá pěti prostorovým kategoriím. Tento fixní počet neuronů je v rozporu s *učitelským přístupem*, proto v kapitole věnované Modelu 3 prezentuji architekturu, která potíže překonává. Nyní se podívejme na způsob práce multimodální vrstvy v Modelu 1.

Z auditivní a vizuální vrstvy do ní vstupuje propojení, jejichž počet odpovídá počtu neuronů vizuální a auditivní vrstvy. Při pokusech hledajících optimální počet neuronů to znamenalo rozmezí 4-800 vstupních propojení na každý z pěti neuronů. Pomocí nich je přenášena informace o vítězných neuronech z primárních vrstev. V multimodální vrstvě jsou tyto informace společně zpracovány, aby se po tréninku nastavily její váhy způsobem, umožňujícím rozlišovat jednotlivé prostorové vztahy. Také multimodální vrstva je založena na principu samoorganizace, tzn. způsob realizace je podobný. Nejprve inicializují 1D síť pěti neuronů, pomocí parametrů α , *funkce okolí a pokles rychlosti učení*. Hodnota těchto parametrů je obdobná jako u primárních vrstev. Následně je síť učena pomocí výstupů z primárních sítí po dobu 100 epoch. Poté jsou síti prezentovány testovací vzorky, které mají určit, zda dokáže model kategorizovat jednotlivé prostorové vztahy. Pomocí nich můžeme poznat, zda bylo učení efektivní.

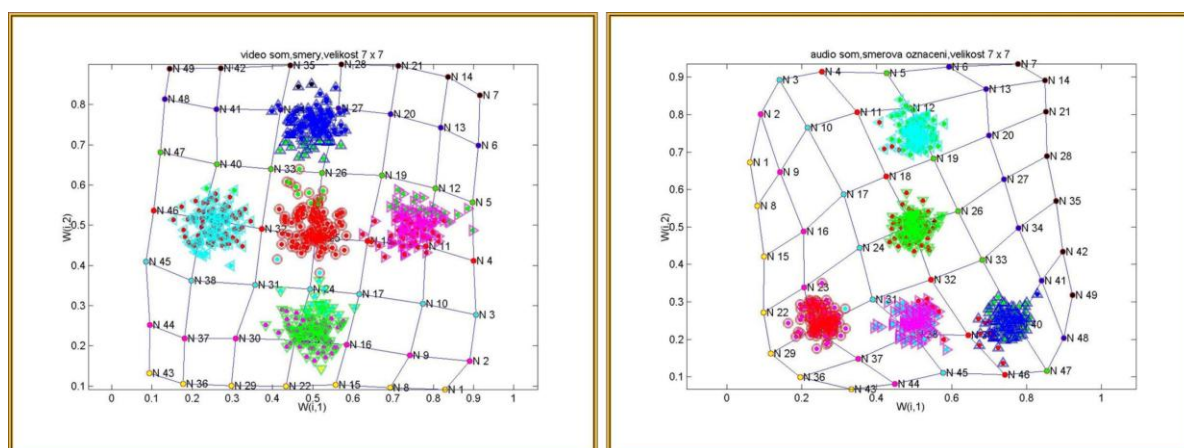
Kromě metody SOM se dá použít také modifikovaná varianta, založená na algoritmu *neurálního plynu* (*Neural Gas*). Jediným rozdílem je způsob propojení sousedních neuronů, který není na rozdíl od SOM fixní a apriorně daný dopředu (Martinetz, Schulten, 1991). Okolí neuronu je tvořeno na základě podobnosti vah, nikoliv podle počáteční konfigurace. Tento typ sítě dokáže lépe reagovat na úlohy s velkým počtem vstupů, tzn. s velkým počtem dimenzí. Při trénování se tedy nastavují shodné parametry jako u SOM, liší se pouze výsledky a efektivitou po natrénování.

Výstupní funkce multimodální vrstvy (SOM i Neural Gas) je přizpůsobena analýze a interpretaci výsledků modelů a proto obsahuje WTA výstup, který označí pouze jeden vítězný neuron, podle něhož určíme správnost kategorizace. Takový výstup se dá napojit na symbolickou část modelu, ve kterém můžeme s reprezentacemi operovat a řetězit je dle pravidel formálního systému. Pokud bychom chtěli použít v další části systému opět neuronové sítě, stačí vhod-

ným způsobem změnit výstupní funkci, která bude vytvářet distribuované reprezentace, tzn. správná kategorie je kódována ve výstupní aktivitě všech pěti neuronů.

8.4 Výsledky

Jestliže známe architekturu modelu, můžeme přistoupit k prezentaci výsledků. První úlohou bylo nalezení minimálního počtu neuronů pro auditivní a vizuální vrstvu, tak aby síť byla schopná správně kategorizovat. Test začal se 4 neurony (mřížka 2x2 neurony) a pokračoval až do počtu 400 neuronů (20x20). V první variantě se jednalo o užití Neural Networks Toolboxu a výstupní funkce WTA. V počáteční fázi je model testován na auditivních a vizuálních shlučích, které byly oddělené a jednotlivé kategorie se nepřekrývaly. Na Obr. 26 můžeme vidět vytrénovanou síť a vstupy prostorových vztahů.



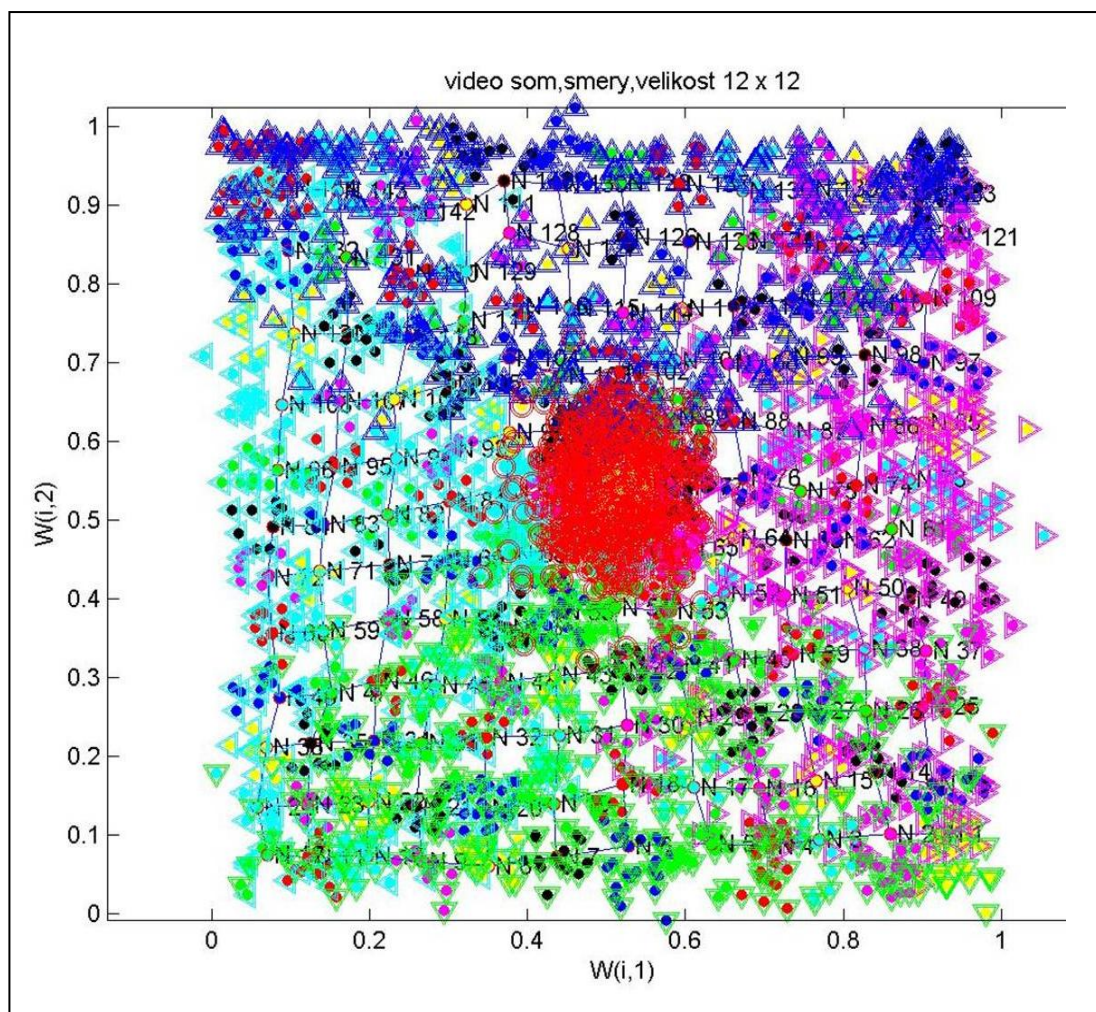
Obr. 26 SOM mřížka po natrénování obsáhne celé senzorné pole. Na obrázku je vidět mřížka 7x7 neuronů a shluky vstupů jednotlivých modalit (vlevo vizuální a vpravo auditivní).

Síť dokázala správně kategorizovat, tzn. vytvářet adekvátní multimodální reprezentace, od velikosti 7x7 neuronů v auditivní a vizuální vrstvě. Od tohoto počtu neuronů dokáže multimodální vrstva identifikovat 100 procent testovacích vstupů. Jedná se však pouze o shluky vstupů pro jednotlivé kategorie, které se ve vizuální vrstvě nepřekrývají, což je v rozporu s psychologickými východisky. V reálném prostředí jsou prostorové kategorie *fuzzy* a není přesně jasné, kde leží jejich hranice.

Proto bylo třeba síť trénovat pomocí překrývajících se shluků ve vizuálním poli, abychom úlohu přiblížili reálným podmínkám. V této fázi začíná model selhávat. Přestože je počet neu-

ronů zvětšen na 20x20 pro obě vstupní vrstvy, nedaří se systému správně kategorizovat. Jako chybné identifikují použití WTA, které na vstup multimodální vrstvy dodává pouze informaci o vítězném neuronu, což je nedostačující pro správnou kategorizaci překrývajících se shluků.

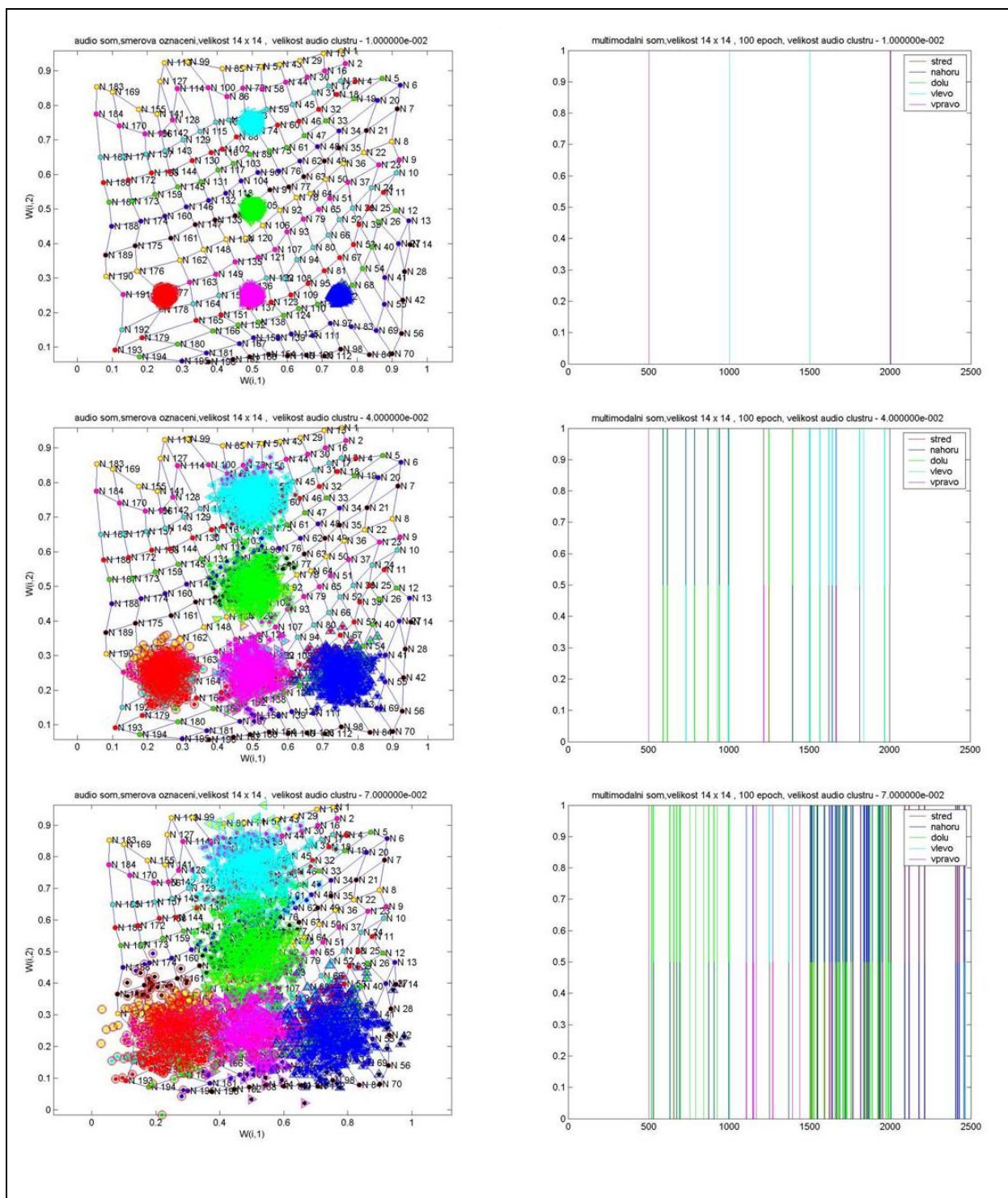
Proto je pro následné trénování použito WTM, které dokáže na multimodální vrstvu dodávat informace o prvních 5 vítězných neuronech, jejichž výstupní aktivita klesá se stoupající vzdálenosti ke vstupnímu vektoru. Pro testování je opět použito sítí se stoupajícím počtem neuronů (od 2x2 po 20x20), přičemž nás zajímá, od jaké velikosti začne síť správně kategorizovat. Během testování se podařilo prokázat, že model dokáže od velikosti 10x10 neuronů pro auditivní i vizuální vrstvu správně kategorizovat jednotlivé prostorové vztahy. Model byl nejprve trénován na umělých vizuálních shlucích, tvořených kolem ortogonálních os, které se postupně zvětšovaly, až do fáze, kdy se překrývaly. Model dokázal správně kategorizovat jak nepřekrývající, tak překrývající se vizuální shluky od velikosti 10x10 neuronů, přičemž se vzrůstajícím překrytem nebylo nutné zvětšovat velikost sítě. V poslední fázi byly modelu prezentovány vizuální vstupy velmi podobné reálnému zadání. Například kategorie vlevo zahrnuje oblast vlevo nahoře, vlevo dole, či vlevo blízko středu, tzn. zasahuje do ostatních kategorií. Značný stupeň vzájemného překrytu je vidět na Obr. 27.



Obr. 27 Ukázka reálných vizuálních vstupů. Jednotlivé prostorové vztahy se překrývají a není jasné, kde leží jejich hranice.

Jelikož síť dokázala úlohu zvládnout, bylo přistoupeno k fázi testování míry rozptylu auditivních označení. V jejich případě znamená rozptyl nepřesnost ve výslovnosti, popřípadě interindividuální rozdíly v řeči. Čím větší rozptyl, tím více různých výslovností slova dokáže síť rozpoznat a správně přiřadit. Pro testování byly opět použity zvětšující se shluky. Vizuální vstupy zůstaly nastaveny na přirozené shluky s velkým překrytem (Obr. 27). Při postupném zvětšování se auditivních shluků dokázal model správně kategorizovat až do okamžiku, kdy se začaly překrývat. Efektivitu můžeme vidět na

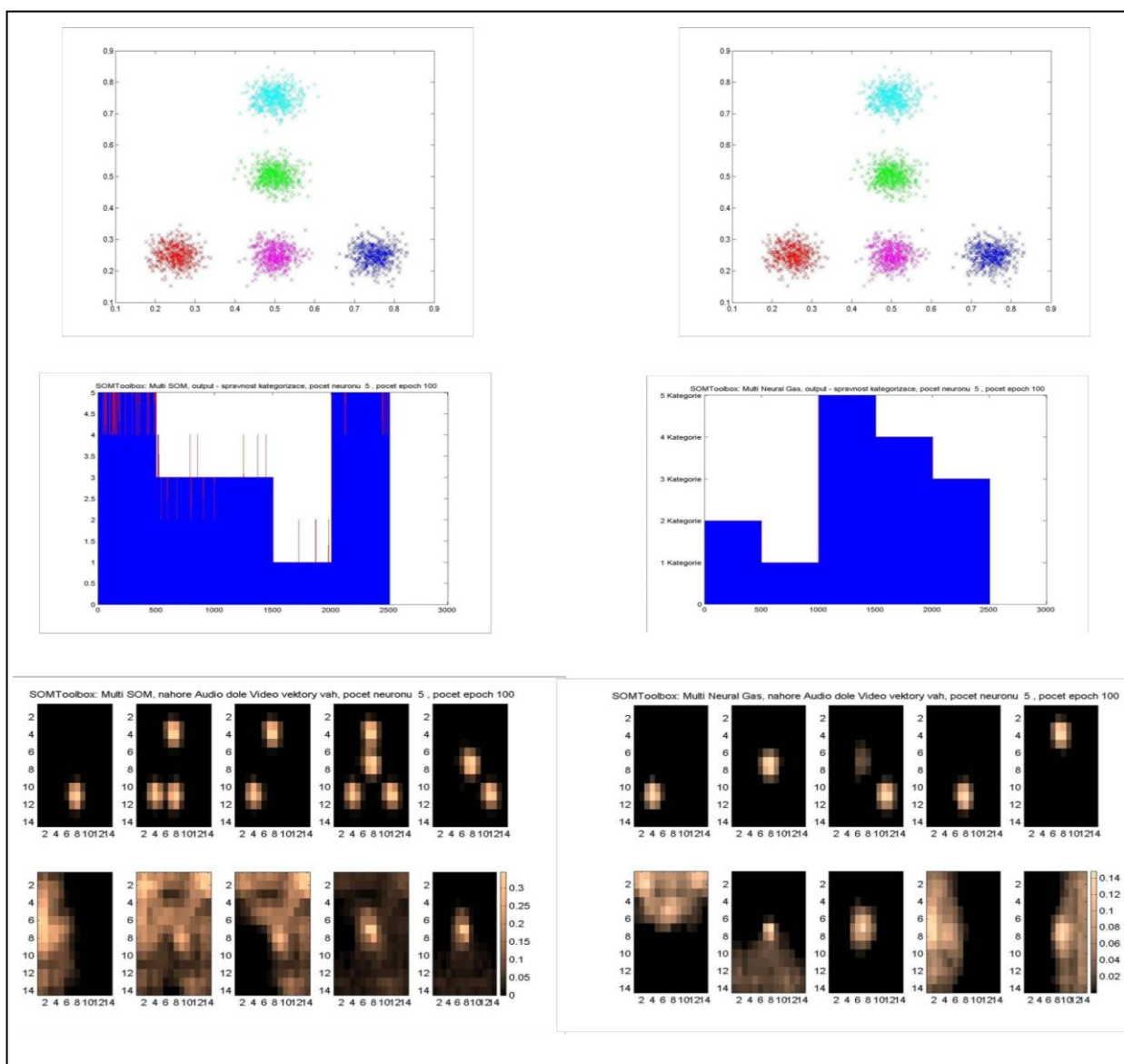
Obr. 28.



Obr. 28 Se zvětšujícím se rozptylem auditivních vstupů (vlevo), při překrývajících se vizuálních vstupech, klesá schopnost kategorizace v multimodální vrstvě (vpravo). Testovací množina obsahovala 500 vzorků každého označení, prezentovaných sekvenčně (osa x). Pokud systém kategorizoval správně, byla během prezentace jednoho vztahu (500 vzorků) vybírána pouze jedna kategorie (osa y). Při chybné kategorizaci byly kategorie odvozo- vány špatně (vpravo uprostřed a dole).

Pokud se shluky překrývaly, nepomohlo ani zvětšování sítě na velikost 20x20. Úloha totiž přestává být řešitelná, jelikož naráží na omezení, která lze odvodit teoreticky. Více je k této problematice uvedeno v diskuzi.

V poslední fázi vývoje modelu bylo užítí algoritmu Neural Gas pro multimodální vrstvu. Při zvětšující se velikosti auditivních shluků ztrácela postupně 1D SOM na efektivitě a vykazovala chyby kategorizace až do úrovně, kdy odvodila špatný počet kategorií. V případě použití Neural Gas byla odolnost vůči chybám kategorizace větší. Efektivitu jednotlivých algoritmů učení je možné vidět na Obr. 29.



Obr. 29 Porovnání efektivity sítí Neural Gas a 1D SOM při zvětšujících se auditivních shlucích. Nahoře je velikost shluků auditivního vstupu, uprostřed kvalita kategorizace multimodální vrstvy a dole analýza vah pěti neuronů multimodální vrstvy. Vlevo síť SOM, vpravo Neural Gas.

Na obrázku vidíme kromě velikosti audio shluků a efektivity multimodální vrstvy také podrobnou analýzu vstupních vah do multimodální vrstvy. Při takovém způsobu vizualizace, kdy jsou jednotlivé vstupy seřazeny do mřížky lze vidět citlivost jednotlivých neuronů na určité typy prostorových kategorií. Podrobnější analýza následuje v diskuzi.

Tím byl ukončen vývoj a testování prvního modelu, protože bylo splněno zadání úlohy. V konečné fázi dokázal model vytvořit multimodální reprezentace pěti prostorových vztahů na základě společné reprezentace vizuální a auditivní vrstvy. V následující tabulce je shrnuto, jaké konkrétní hodnoty a typy učení jsou nutné pro efektivní činnost modelu. Počty neuronů jsou větší než minimální potřebný počet. Důvodem je zajištění správné činnosti i v případě působení náhodných vlivů. Proto je model v této oblasti naddimenzován.

auditivní vstup	shluky 2D vstupů bez překrytu
vizuální vstup	překrývající se shluky 2D vstupů
auditivní síť	SOM 2D o velikosti 14x14 neuronů
vizuální síť	SOM 2D o velikosti 14x14 neuronů
výstupní funkce	WTM
multimodální vstup	plně propojen s auditivním a vizuálním výstupem
multimodální síť	neural Gas
výstup	WTA (pro ověření správné kategorizace)

Tab. 5 Přehled jednotlivých částí modelů a způsobů zpracování

8.5 Diskuze

Jelikož se v souvislosti s realizací modelů mohou vyskytnout nejasnosti, pokusím se upřesnit některé části. Netradičním postupem je trénování sítě pomocí náhodných hodnot distribuovaných v celém vstupním prostoru. Ostatní modely bývají většinou trénovány pouze trénovacími vstupy, které jsou blízké testovacím vzorkům. Síť poté dokáže adekvátně reagovat pouze na určitou část vstupů. V našem případě ovšem vyžadujeme univerzálně použitelné primární reprezentace, které jsou schopné reagovat na veškeré podněty z okolí. Jelikož se jedná o reprezentaci základních auditivních a vizuálních vlastností, musí být síť schopná reagovat na veškeré podněty. Proto je možné síť trénovat pomocí náhodných vstupů a po natrénování váhy zmrazit.

Během testování nevznikal rozdíl mezi užitím čtvercové a hexagonální SOM mřížky mezi sousedními neurony. Kromě povahy náhodných trénovacích vstupů, můžeme tento fakt při-

soudit také typu řešené úlohy. Jestliže bylo potřeba vytvořit síť, která měla reprezentovat pouze několik (5) základních kategorií, nejedná se o náročnou úlohu, ovlivnitelnou typem mřížky. Vliv vzdálenosti sousedních neuronů související s typem mřížky by se projevil v úloze s mnoha kategoriemi.

Parametry sítě související se způsobem konfigurace neuronů a rychlostmi učení byly odvozeny na základě empirických zkušeností a využitím zkušeností odborníků zabývajících se danou architekturou v úrovni teoretického výzkumu (Kohonen, 1989). Parametry je třeba nastavit takovým způsobem, aby síť dokázala během krátké doby efektivně pokrývat celý vstupní prostor a adekvátně reprezentovat kategorie tohoto prostoru.

Jádrem systému je multimodální vrstva zodpovídající za proces identifikace, tedy vytvoření a reprezentaci jednoznačných kategorií. V případě modelu se jedná o nalezení vhodného propojení mezi primárním auditivním a vizuálním výstupem, který je multimodální vrstvě prezentován v jeden okamžik. Cílem učení na základě těchto dvou informací je nalézt kategorie, které reprezentují společné vlastnosti vstupů. Rozdíl oproti klasickému Harnadovu přístupu je podrobně rozebrán v závěrečné kapitole.

Použití funkce WTM (namísto WTA) má jasné opodstatnění. Jelikož primární vrstvy nejsou koncové, ale přenášejí informace dále do multimodální vrstvy, potřebujeme výstupní funkci, která nebude obsahovat pouze jeden vítězný neuron, ale váhově odstupňované pořadí vítězných neuronů. Jedná se o informačně bohatší zdroj, který lépe poslouží následné identifikaci v multimodální vrstvě. Použití WTA a výstupní funkce jednoho neuronu má za následek neschopnost modelu přesně kategorizovat, jelikož informace posílaná do multimodální vrstvy je nedostačující.

Za zmínku rovněž stojí analýza vstupních vah multimodální vrstvy, která je graficky vynesena na Obr. 29. Způsob vizualizace nám umožňuje pochopit, jakým způsobem jsou trénovány samoorganizující se sítě pro vícedimenzionální vstupy. Každý z pěti neuronů multimodální vrstvy je propojen se 100 výstupy z auditivní a 100 výstupy vizuální vrstvy, pokud je velikost mřížky 10x10 neuronů. Každý neuron tedy dostává 200 rozměrný vstup. My však víme, že jednotlivé výstupy primárních vrstev znamenají aktivitu neuronů citlivých na určité hodnoty vstupních dat, tedy citlivost na vstup v určité části 2D pole. Jestliže 100 výstupů reprezentujeme jako mřížku neuronů, které jsou vzájemně propojeny a zcela pokrývají 2D plochu vstupního prostoru, získáváme tak pole 10x10 neuronů. Každý neuron multimodální vrstvy obsahující 200 vstupů, se dá reprezentovat jako 2 mřížky o straně 10, které reprezentují vstupy z primárních vrstev. Grafická analýza vstupů následně ukazuje, jak se vstupní váhy multimodální vrstvy přizpůsobily primárním výstupům. Každý jednotlivý neuron znamená jednu prostorovou kategorii a jeho vstupní váhy jsou citlivé k specifickým primárním vstupům.

V konečném důsledku vidíme, že každý multimodální neuron je citlivý (světlé body) na prostorovou výšeč reprezentující jednotlivé polohy (vlevo, vpravo...) vizuálního vstupu a zároveň na kruhové shluky, reprezentující fonetické vlastnosti jazykových označení těchto výšečí (vlevo, vpravo...) auditivního vstupu. Na obrázku můžeme vidět, že multimodální vrstva dokázala správně namapovat vizuální a auditivní informaci, jelikož prostorové výšeči odpovídá auditivní shluk se správnými fonetickými vlastnostmi. Tato grafická analýza umožní kontrolovat, zda během procesu mapování nedošlo k chybě. Na Obr. 29 vidíme, že v případě použití sítí typu Neural Gas, je vzájemné mapování adekvátní, ale v případě použití sítí SOM, dochází k mapování vizuálních vstupů na dvě odlišné auditivní kategorie, což vede k chybám modelu během testování. Největší slabinou modelu je predefinovaný počet prostorových kategorií v multimodální vrstvě. Samotné použití sítě s fixním počtem neuronů by nebylo problematické. Chybné je vytvořit jejich počet, který je předpokládán v souvislosti se zadanou úlohou, jelikož tím předjímáme správné řešení úlohy. Pokud bychom chtěli zachovat *učitelský přístup*, musíme nastavit počet neuronů v multimodální vrstvě zcela náhodně a vytvořit metody učení, které dokážou reprezentovat počet kategorií odpovídající vstupním informacím. Tato varianta nebyla realizována kvůli komplikacím s tvorbou specifické metody učení, využívající pouze adekvátní počet neuronů. Problematika souvisí s lokalistickou a distribuovanou reprezentací. Jestliže chceme výstup modelu propojit se symbolickou částí, je nutná lokalistická reprezentace kategorií související s jejím jednoznačným odlišením. V případě distribuované reprezentace (redundantní počet neuronů) bychom potřebovali další reprezentační vrstvu, která by převedla distribuované reprezentace na lokalistické. Proto se jako přijatelnější alternativa jeví použití rostoucích sítí (GWR) (Marsland, 2005), vytvářejících adekvátní počet neuronů na základě vstupní informace. Podrobnosti jsou uvedeny v kapitole věnované Modelu 3.

Přestože má multimodální vrstva fixní počet neuronů, je postavena na principu *učení bez učitele*, kategorizujícím vstupní data bez předchozí, či externě vložené znalosti. Užití podobných parametrů učení pro multimodální i primární vrstvy ukazuje, že tvorba architektury je jednoduchá, ve smyslu malého počtu nastavovaných parametrů. Jelikož ani jeden z parametrů nepřekračoval hodnoty, které se používají pro trénování odlišných systémů, je použití principu samoorganizace v oblasti modelování velmi efektivní.

Výsledky fungování sítě nebyly jako v ostatních případech vyjádřeny mírou efektivity sítě na testovacích vzorcích. Díky povaze úlohy a jednoduché analýze fungování jednotlivých vrstev jsem přistoupil k variantě, ve které je efektivní pouze taková síť, která dokáže správně rozlišovat testované vzorky na 100%. Proměnnou, se kterou je v případě modelu manipulováno, je počet neuronů ve vstupních vrstvách, následované měřením efektivity modelu. Jestliže pracuje bezchybně, můžeme stanovit minimální počet neuronů, dostačující ke zvládnutí úlohy. Síť

je následně testována pomocí vyššího počtu neuronů, abychom potvrdili, že se nejedná o náhodný jev. Není potřeba vyjadřovat procentuální úspěšnost modelu, jelikož nás zajímá pouze bezchybná varianta. Závěrem procesu modelování je spolehlivá a vytrénovaná síť.

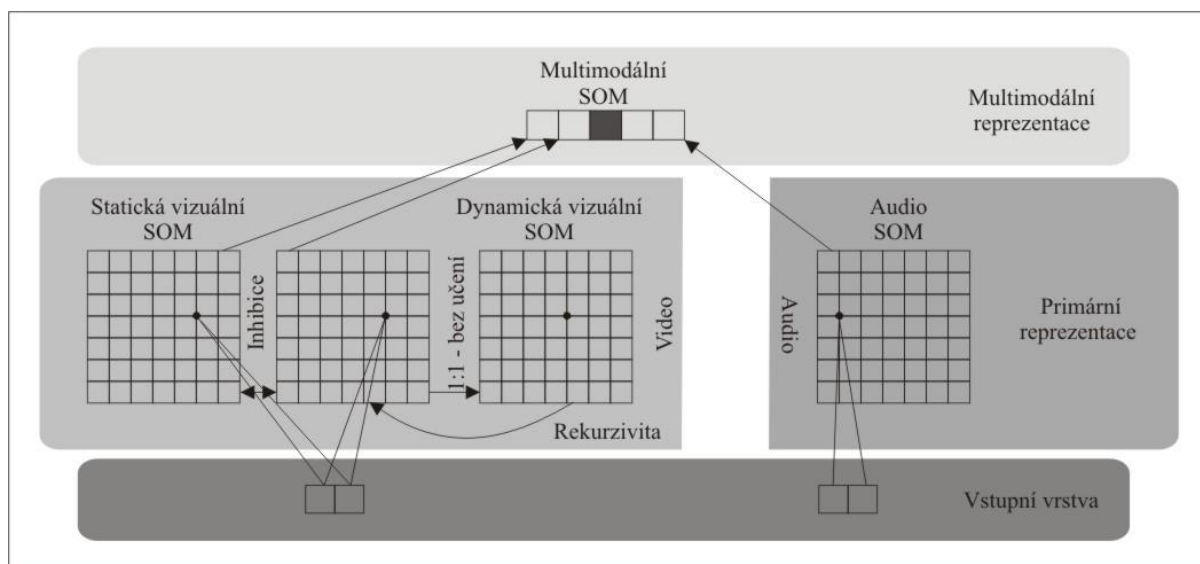
Pokud bychom se podívali na danou oblast z hlediska matematického, není funkce modelu ničím složitým či překvapujícím. Jedná o velmi jednoduché zadání, jelikož používáme pouze dvourozměrné vstupní vektory, které nejsou pro následné zpracování příliš výpočetně náročné. Také počet trénovacích vzorků neznamena enormní zatížení systémů. Způsob jejich rozložení do pěti ohraničených (auditivní vstup) a pěti překrývajících se (vizuální vstup) kategorií, znamená pro neuronovou síť s daným počtem neuronů nenáročnou úlohou. Samotný princip učení se v multimodální vrstvě není z hlediska teoretické matematiky opět ničím složitým. Hledáme mapování mezi dvěma maticemi výstupních funkcí primárních vrstev, ve kterých se díky počtu neuronů dá dostatečně reprezentovat jednotlivé informace z prostředí a díky tomu, že jeden vstup je vždy tvořen ohraničenými kategoriemi (v tomto případě vždy auditivní), lze správné mapování nalézt. Pro oblast matematiky také není překvapením, že nemůžeme odvodit správné mapování mezi dvěma zdroji překrývajících se kategorií, jelikož takové mapování není možné teoreticky ani prakticky. Přínos modelu tedy spočívá v oblasti návrhu vzájemného propojení neuronových sítí, nastavení parametrů a způsobu učení, zaručující jeho efektivní fungování. Z hlediska praxe je přínosné využití zmíněných teoretických poznatků v konkrétní oblasti reprezentace významu, které se v tomto případě soustřeďuje na vnímání prostoru.

9 Model 2 - dynamické prostorové vztahy

Jelikož prostorová označení obsahují nejen statické, ale i dynamické vztahy mezi objekty, bylo potřeba přikročit ke tvorbě modelu, který by je dokázal vhodně reprezentovat. Výraz dynamické vztahy znamená, že jeden objekt vůči druhému mění polohu v čase, přičemž jeho výsledné označení popisuje typ trajektorie. Oproti statickému modelu je v případě reprezentace dynamických scén situace komplikovanější, protože musíme brát v potaz temporální informace, tzn. nalézt vhodný způsob, jakým reprezentovat průběh pohybu v čase. Jelikož při reprezentaci času roste výpočetní mohutnost, přistupují k této problematice badatelé odlišně. Regierův model obsahoval velmi elementární mechanismus. Apriorně označil objekt, u kterého chtěl zjistit prostorový vztah vzhledem k druhému, a poté zjednodušil průběh pohybu na sekvenci *počátek-cesta-cíl*. Z nich odvozoval typ pohybu, přičemž v případě *cesty* systém detekoval, zda došlo k doteku či průniku sledovaného objektu, pro rozlišení prostorových vztahů *skrz*, *po* a *pod*. Použité schéma vychází z teoretických předpokladů Millera a Johnson-Lairda (1976). Podle nich je pro určení prostorového vztahu nutná znalost těchto tří elementů, které jsou klíčové při stanovení správného označení.

V navrženém modelu pracuji s reprezentací pohybu odlišně, jelikož Regier ve svém řešení potřebuje znát okraje objektů a dalších vlastností, které souvisí s rozpoznáváním objektů (*what dráha*). V našem případě je pozornost soustředěna pouze na informace o poloze objektu (*where dráha*), nikoliv na jeho tvarové vlastnosti. Řešení je daleko univerzálnější, protože Regierův návrh je založen na apriorních předpokladech či omezeních týkajících se tvaru a vlastností objektů, což může být zavádějící pro jejich následnou interpretaci.

Pro reprezentaci dynamických scén slouží speciální vizuální subsystém, který je založen na rekurzivních sítích typu RecSOM (Voegtlin, 2002), které dokážou reprezentovat pohyb objektů jako sekvenci elementárních pohybů, na jejichž různé směry jsou citlivé jednotlivé neurony této sítě. Jeho specifická funkce je v souladu s biologickými a psychologickými poznatky. Lidské oko obsahuje neurony citlivé jak na statickou polohu objektů, tak na pohyb. Sítě RecSOM jsou výhodné pro svou schopnost reprezentace pohybu. Oproti klasické SOM architektuře (Kohonen, 1989) obsahují kontextovou vrstvu, která dokáže zachytit časové proměny a vhodným způsobem je reprezentovat. Schéma rozšířené verze modelu je vidět na Obr. 30.



Obr. 30 Základní schéma modelu. Auditivní a vizuální vstup je reprezentován odděleně pomocí SOM sítí. Výstupy jsou následně integrovány v multimodální vrstvě. Inhibice vizuálních vrstev dosud není realizována.

Z hlediska architektury se jedná o rozšířenou verzi Modelu 1. Primární auditivní i multimodální vrstva je totožná co se týká použitého typu sítí (SOM popřípadě Neural Gas), způsobu učení a vzájemného propojení. Odlišný je pouze vizuální subsystém. Místo klasické SOM sítě, která se osvědčila pro reprezentaci statických objektů, obsahuje RecSOM modul, sloužící pro reprezentaci pohybu objektu v čase. Následné propojení s multimodální vrstvou je totožné jako v předcházejícím případě, tzn. každý s každým. Zachován je také způsob prezentace vstupů v auditivní i vizuální vrstvě.

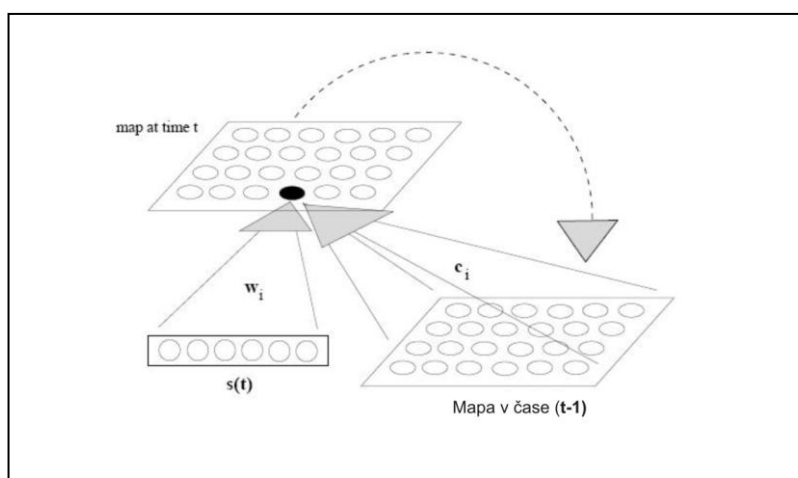
9.1 Zadání úlohy

Oproti Modelu 1 se naše pozornost přesouvá k prostorovým označením dynamických vztahů. Cílem je vytvořit systém, schopný ukotvit pět základních dynamických prostorových označení. Jedná se o významy *přes*, *pod*, *okolo*, *skrz* a *ven*. Opět se snažíme nalézt odpovídající počet neuronů v jednotlivých vrstvách modelu tak, aby po natrénování dokázal bezchybně rozpoznat testovací vzory. Pohybující se objekty budou systému prezentovány jako sekvence souřadnic, určujících polohu pohybujícího se objektu.

9.2 Metoda

Jelikož architektura SOM byla popsána v souvislosti s Modelem 1, bude pozornost soustředěna na RecSOM modul. Sítě RecSOM byly poprvé navrženy Voegtlinem v roce 2002. Jedná se o modifikaci původního návrhu SOM. Autor vychází z předpokladu, že pro reprezentaci temporálních dat je potřeba rozšířit architekturu o pomocnou vrstvu, která bude uchovávat data z předchozího časového kroku. Během učení má síť k dispozici nejen informaci o současném vstupu, ale také o vstupu předchozím a dokáže reagovat na sekvence časových dat, pomocí nichž vytváří interní reprezentace.

Základem architektury je klasická síť SOM se vstupní vrstvou a výstupní vrstvou, které jsou vzájemně propojeny metodou *každý s každým*. Novinkou je přidání další vrstvy, napojené na výstupní vrstvu, nazývaní se *kontextová* a sloužící právě k lepší reprezentaci časových dat. V případě že vložíme na vstupu hodnoty, jsou poslány do výstupní vrstvy, kde pomocí principu samoorganizace dojde k úpravě vah neuronů. Následně jsou hodnoty vah neuronů výstupní vrstvy překopírovány bez modifikace do *kontextové vrstvy*. Informace zde uložené jsou použity pro další krok učení. Výstupní vrstva totiž obsahuje dvě váhy, jedna slouží pro vstupní data, druhá pro informace z kontextové vrstvy, ve které je uložena informace o nastavení vah z předchozího kroku. Učení probíhá následovně. V prvním kroku je síť inicializována a váhy výstupní a kontextové vrstvy jsou nastaveny náhodně. Po informaci ze vstupu síť pomocí principu samoorganizace změní nastavení svých vah a informace pošle do *kontextové vrstvy*. V dalším kroku probíhá učení jak na základě vstupu, tak informace z *kontextové vrstvy*, která obsahuje informaci o nastavení vah v předchozím kroku. Tento způsob je opakován až do posledního vstupu časové řady. Schéma architektury je vidět na Obr. 31.



Obr. 31 Schéma RecSOM (Voegtlin, 2002)

Způsob nastavování vah pro vstupní data a informace z kontextové vrstvy je totožný.

$$\begin{aligned}\Delta w_i &= \gamma \cdot h_{ik} \cdot (s(t) - w_i), \\ \Delta c_i &= \gamma \cdot h_{ik} \cdot (y(t-1) - c_i),\end{aligned}$$

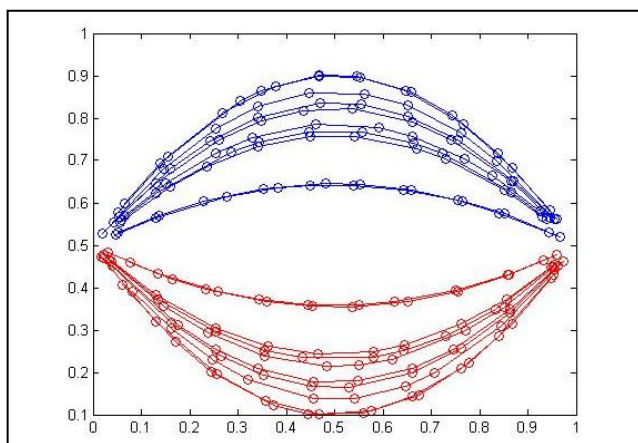
První vzorec se týká vstupní vrstvy a druhý kontextové. Gamma je koeficient učení, h určuje funkci okolí a w a c jsou váhy. Princip je obdobný jako v případě SOM sítě.

Existuje mnoho variant sítí, které jsou vhodné pro zpracování časových dat, přičemž výchozí princip samoorganizace je podobný. Mezi nejznámější patří Temporal Kohonen Map (TKM), Self-Organizing Map For Structured Data (SOMSD), Merge Self-Organizing Map (MSOM) či Sardnet. Síť RecSOM jsem zvolil pro její jednoduchou interpretovatelnost, snadnou dostupnost zdrojových kódů a výpočetní nenáročnost.

9.3 Realizace

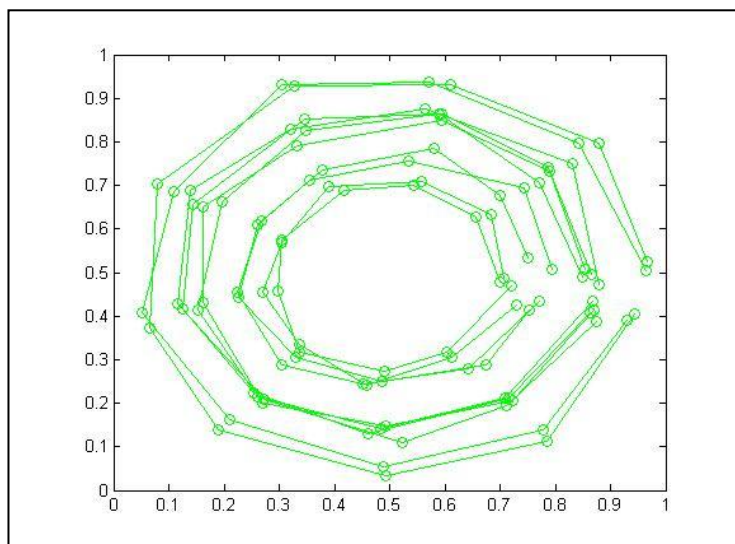
Jelikož je činnost modelu velmi obdobná s Modelem 1, popíšu způsob provedení pouze zjednodušeně. Samotná činnost vypadá následovně. Síti jsou prezentovány trénovací vzorky jednotlivých typů dynamických prostorových vztahů (*přes, pod, skrz, ven a okolo*) ve formě sekvence, která se skládá z deseti dvourozměrných souřadnic popisujících trajektorii pohybu.

Ty byly vytvořeny pomocí matematických funkcí, které popisují daný pohyb. V případě pohybů *nad* a *pod* se jednalo o půlkruhy, opisující oblouk nad nebo pod středem zorného pole. Aby se síť naučila veškeré druhy pohybů, obsahují trénovací vzorky různé způsoby trajektorie, tzn. mění se poloměr půlkruhu. Souřadnice trajektorií jsou zároveň změněny o malou hodnotu náhodného čísla, simulující šum a nepřesnosti, jako je tomu v reálných podmínkách. Vstupní sekvence pohybů jsou zřetelné z Obr. 32.



Obr. 32 Ukázka vstupů *nad* a *pod*

V případě pohybu *okolo* se jednalo o kruhy, které měly střed v oblasti středu vizuálního senzoreckého pole. Vstupy byly opět rozdílného poloměru, zaručující různé varianty pohybu. I v tomto případě používám náhodné číslo, které změnilo polohy bodů definujících trajektorii tak, aby se nejednalo o přesnou matematickou funkci. Vstupy nejlépe demonstruje Obr. 33.



Obr. 33 Ukázka vstupu *okolo*

Jestliže předchozí pohyby měli kruhovou trajektorii v případě směrů *skrz* a *ven* se jedná o přímky. U směru *ven* se jedná o paprskovitě vybíhající přímky ze středu senzoru, vycházející do všech stran. Směr *skrz* je reprezentován pomocí trajektorií probíhajících přes střed senzoru. Opět je zde použito náhodné číslo, které upraví ideální souřadnice trajektorie tak, aby obsahovaly šumy a chyby.

Přestože souřadnic obsahují náhodné chyby, aby se podobaly reálným datům, byly při jejich tvorbě provedeny redukce a zjednodušení. V případě směrů *nad* a *pod* je počátek pohybu umístěn vždy v levé části zorného pole a končí v pravé části. V reálném životě se můžeme setkat také s příklady *pod* a *nad*, které probíhají opačným směrem. Přestože model není trénován na všechny možné příklady označení, můžeme to pro naše účely považovat za dostačující. Zahrnutí opačných směrů znamená naučit systém kategorizovat správně i trajektorie, které vedou z opačné strany. Zvýšila by se výpočetní a tréninková náročnost a bylo by potřeba pravděpodobně zvětšit počet neuronů. Podobný důsledek by znamenalo trénování sítě na příklady, jejichž trajektorie není přibližnou matematickou křivkou, například označení *nad* probíhající po náhodné trajektorii, kterou nelze přesně matematicky vyjádřit, ale přesto patří do této kategorie. V modelu jsou tréninkové vzorky pouze základní trajektorie, jelikož způsobů jakým se dá vyjádřit např. pohyb *nad* je tolik, že by trénovací vzorek byl

příliš velký. Přesto si uvědomuji nutnost vytvořit architekturu modelu tak, aby dokázala všechny tyto trajektorie správně kategorizovat. Testování podobných variant je plánováno v budoucí verzi modelu, přesto považuji za nutné redukci současného řešení zmínit.

Jestliže máme připravenou sadu auditivních a vizuálních tréninkových vzorů a sadu testovacích vzorů (vzniknou obdobným způsobem) můžeme přistoupit k trénování. Nejprve je třeba síť SOM i RecSOM inicializovat. Auditivní vrstva je opět trénována na náhodných vstupech, jelikož potřebujeme získat síť, která je schopná reagovat na vstupy v rámci celého stavového pole. V případě RecSOM sítě nemůžeme postupovat obdobně, jelikož by velikost trénovacího vzorku, který obsahuje náhodné varianty časových průběhů trajektorií, byl natolik obsáhlý, že by trénování probíhalo neúměrně dlouho. Proto jsou u RecSOM modulu trénovací vzorky omezeny pouze na příklady konkrétních trajektorií.

Síť jsem trénoval pomocí 500 trajektorií pro každé prostorové označení po dobu 100 epoch. Během tréninku jsou jí prezentovány střídavě vždy po jedné trajektorii z každého směru, tzn. je prezentováno 10 vstupních vektorů pro jeden směr a následně je resetována kontextová vrstva, aby mohl být trénován další vstup. Princip resetu kontextové vrstvy je nutný k tomu, aby síť oddělila jednotlivé trajektorie od sebe. K tomu slouží fixní počet vektorů, které reprezentují trajektorii. Přestože je tento požadavek v rozporu s *učitelským přístupem*, nelze model realizovat jinak. V současné době není znám typ sítě, který by byl schopen samostatně rozpoznat konec jedné sekvence pouze na základě vstupních dat. Model s podobnými vlastnostmi by musel obsahovat interní mechanismus sloužící k oddělení jednotlivých vstupů, jehož realizace je nad současné možnosti. Proto je použit fixní počet deseti kroků trajektorie.

Po natrénování obsahuje výstupní vizuální vrstva neurony citlivé na specifické změny polohy. Je toho docíleno architekturou RecSOM, která dokáže pomocí trénování vytvořit neurony, přičemž každý z nich je citlivý na jiný směr pohybu. Celkový pohyb je následně reprezentován jako sekvence změn a výstupní funkce posílá do multimodální vrstvy sekvenci vítězných neuronů. Informace je zpracována společně s auditivním vstupem (označení prostorových vztahů) a dochází ke tvorbě multimodálních kategorií. Čas je reprezentován jako sekvence výstupní funkce neuronů citlivých na jednotlivé změny pohybů, pomocí které můžeme odlišit jednotlivé typy dynamických prostorových vztahů již v rovině zpracování vizuálního podnětu a vytvořit jejich elementární kategorie (v úrovni elementárních částí celkového pohybu). Auditivní informace je pro shodný způsob prezentace zpracovávána obdobně jako u statických prostorových vztahů. Model je rozšířen pouze o rekurzivní síť reprezentace pohybu, jinak je identický s modelem předchozím, což zajišťuje dobrou možnost jejich propojení.

Realizace modelu proběhla pomocí programu Matlab s využitím toolboxu Neural Networks. Během testování bylo použito také specializovaných modulů SOMToolbox pro auditivní a multimodální vrstvu a RecSOMToolbox pro vizuální subsystém. Většinu zmíněných nástrojů jsem kvůli lepší funkčnosti modifikoval.

9.4 Výsledky

Testování probíhalo obdobně jako u Modelu 1. Z předchozích výsledků víme, že model fungoval od velikosti 10x10 neuronů. Protože je auditivní vrstva nezměněna a v budoucnu uvažuji o propojení Modelu 1 a 2 do jednoho, nastavil jsem její počáteční velikost obdobně. Pro rekurzivní vrstvu probíhalo testování od velikosti mřížky 2x2 neurony, jelikož nepožadujeme, aby měl systém stejný počet neuronů citlivých na statickou polohu objektů a na pohyb.

Způsob nastavování a práce s modelem jsou v mnohém shodné s předchozím. Trénování auditivní vrstvy je zcela totožné, pouze jsou použita slovní označení s odlišnými fonetickými vlastnostmi. Opět vznikne čtvercová mřížka, ve které jsou neurony citlivé na dané prostorové označení. Vítězné neurony (WTM) posílají informaci o své aktivitě do multimodální vrstvy.

V případě vizuální mřížky probíhá trénování podle principu popsaných v předchozím odstavci. Při tvorbě sítě je také potřeba nastavit podobné parametry jako u SOM architektury. Jedná se o *rychlost učení*, *velikost a tvar okolí* a také *pokles rychlosti učení* v závislosti na počtu epoch.

Během testování se ukázalo, že pro správnou kategorizaci je nutné, aby velikost vizuálního vstupu byla alespoň 10x10 neuronů, a pokud chceme, aby model pracoval efektivně, je potřeba mřížka o velikosti 13x13 neuronů. Při těchto hodnotách dokáže model správně kategorizovat testovací vstupy se stoprocentní úspěšností. Jedná se o podobné hodnoty jako v případě Modelu 1, což nám dává velkou výhodu pro jejich společnou integraci, jelikož velikost primární auditivní mřížky můžeme ponechat stejnou pro obě části vizuálních subsystémů. Podobně jako u předchozího modelu prokázal větší efektivitu při kategorizaci vstupů do multimodální vrstvy algoritmus Neural Gas oproti klasické 1D SOM a je vhodný pro aplikaci v jednotném modelu. Na závěr uvádím tabulku obsahující konkrétní údaje týkající se výsledného modelu, schopného zvládnout úlohu dle popsaného zadání. Hodnoty neuronů v jednotlivých vrstvách jsou naddimenzované, abychom eliminovali působení náhodných vlivů.

audio vstup	shluky 2D vstupů bez překrytu
video vstup	trajektorie pohybu z 10 souřadnic 2D
audio síť	SOM 2D o velikosti 14x14 neuronů
video síť	RecSOM 2D o velikosti 14x14 neuronů
výstupní funkce	WTM
multimodální vstup	plně propojen s auditivním a vizuálním výstupem
multimodální síť	Neural Gas
výstup	WTA (pro ověření správné kategorizace)

Tab. 6 Přehled jednotlivých částí modelů a způsobů zpracování

9.5 Diskuze

Jak vyplývá z předchozího textu, realizovaný model se od ostatních výzkumů odlišuje v několika ohledech. Pozornost je v něm soustředěna pouze na zpracování prostorové polohy objektu bez přesné identifikace jeho tvaru. Důvodů pro takový postup je několik. Regierův model (1996) obsahuje apriorní předpoklad o centrálním a k němu vztaženém objektu. Vytváří systém, ve kterém je centrální objekt umístěn vždy ve středu a relativní objekt zaujímá prostorovou pozici. Je tedy vytvořen referenční rámec, který není schopen reflektovat změnu polohy senzoru (oka), tzn. přesunu centrálního objektu na periferii. V našem modelu je centrální objekt nahrazen centrem senzoru a změna polohy se projeví změnou prostorového vztahu vzhledem k pozorovateli. Přesto je tento model - podobně jako Regierův - založen na jediném referenčním rámci. Jak již bylo zmíněno (Vavrečka, 2007a), nelze pomocí něj zajistit identifikaci více než jednoho objektu v prostoru (vzhledem k pozorovatelovu centru senzorického pole), a proto je plánováno rozšíření o relativní prostorové moduly, které by umožňovaly identifikaci více objektů.

Důvod pro absenci systému rozpoznání objektů (*what dráha*) je následující. V současné fázi výzkumu není možné zajistit univerzálnost reprezentace těchto vztahů, jelikož modely identifikace objektů nejsou na dostatečné úrovni, abychom zaručili spolehlivost rozpoznání různých tvarů a typů objektů. Většina navrhovaných modelů obsahuje velké množství požadavků na schopnost systému, které lze uskutečnit pouze různými formami redukce.

Proto realizovaný model tyto prostorové vztahy nezahrnuje. Pozornost je v něm soustředěna pouze na základní prostorové vztahy, které je možné realizovat v jediném referenčním rámci bez nutnosti identifikace tvaru objektu. Model nedokáže rozlišit vztahy, při kterých dochází k interakci dvou objektů na úrovni jejich vzájemného propojení či nulové vzdálenosti mezi nimi. Na druhou stranu jsou realizované prostorové reprezentace univerzálně použitelné ve

všech typech prostředí s libovolnými objekty a libovolným typem jazykových označení. Pro samotný výzkum způsobu ukotvení symbolů, přesněji řečeno převodu senzorických informací do abstraktní úrovně, je navrhovaný systém dostačující. Rozšíření modelu nám pouze umožní testovat systém na složitějším typu úloh.

V dalších fázích bude model rozšířen o zmíněný relativní prostorový systém, který využívá informace ze stávajícího absolutního modulu, a také o schopnost reprezentovat dynamické prostorové vztahy prezentované v libovolném počtu kroků. Hlavním cílem, který souvisí s problematikou ukotvení symbolů, je následné zpracování v symbolické rovině a tvorba kompozicionality. Model získá schopnost reprezentovat okolní prostředí a vytvářet o něm závěry ve formě výroků, jejichž pravdivostní hodnota není do systému vkládána externím pozorovatelem, ale je systémem budována a reprezentována na základě součinnosti jeho senzorických vstupů.

10 Model 3 – rostoucí síť

Jelikož se chci přiblížit psychologické plauzibilitě, rozhodl jsem se rozšířit stávající Model 1 o rostoucí síť. Jedná se o architekturu umožňující přidávání či ubírání neuronů pomocí speciálních mechanismů učení na základě vstupních dat, jelikož předchozí model (podobně jako Regierův) obsahoval apriorně daný a fixní počet kategorií pro označení prostorových vztahů, což je v rozporu s *učitelským přístupem* (Ziemke, 1999). Struktura multimodální vrstvy byla predefinována designérem a nereflektovala vstupní data (např. změnu počtu označení). Systém používající rostoucí síť je naopak schopen pracovat flexibilně. V praxi to znamená, že systém se dokáže naučit různé jazyky s rozdílným počtem označení prostorových vztahů. Pokud nejsou výrazy jednoho jazyka používány (změna prostředí), systém se naučí reagovat na jiný jazyk, tzn. přestane používat stávající neurony nebo je přeučí. Řešení je možné použitím architektury GSOM (Growing Self Organizing Maps), založené na přidávání neuronů pro reprezentaci vstupů z prostředí v závislosti na interakci agent-prostředí. V jiné oblasti kognitivních věd byla tato architektura úspěšně odzkoušena při vývoji lexika (Farkaš, 2002). Jelikož je Model 3 v rámci disertace dokončen jen do úrovně testování počtu neuronů, bude následný popis realizace obsahovat pouze základní přehled jeho fungování.

10.1 Realizace

Oproti Modelu 1 a 2 je společná multimodální vrstva tvořena sítí typu GWR – *Growing When Required* (Marsland, 2005). V první fázi učení obsahuje výstupní vrstva pouze dva neurony, přičemž algoritmus umožňuje přidávání nových, pokud síť nedokáže kategorizovat s dostatečnou přesností. Klíčovým parametrem je nastavení *citlivosti* (*insertion threshold*), ovlivňující konečný počet neuronů. *Citlivost* je definována jako míra správnosti kategorizace. Pokud je překročena, síť vygeneruje nový neuron. Závislost změny citlivosti kategorizace na počet neuronů je zachycena v Tab. 7.

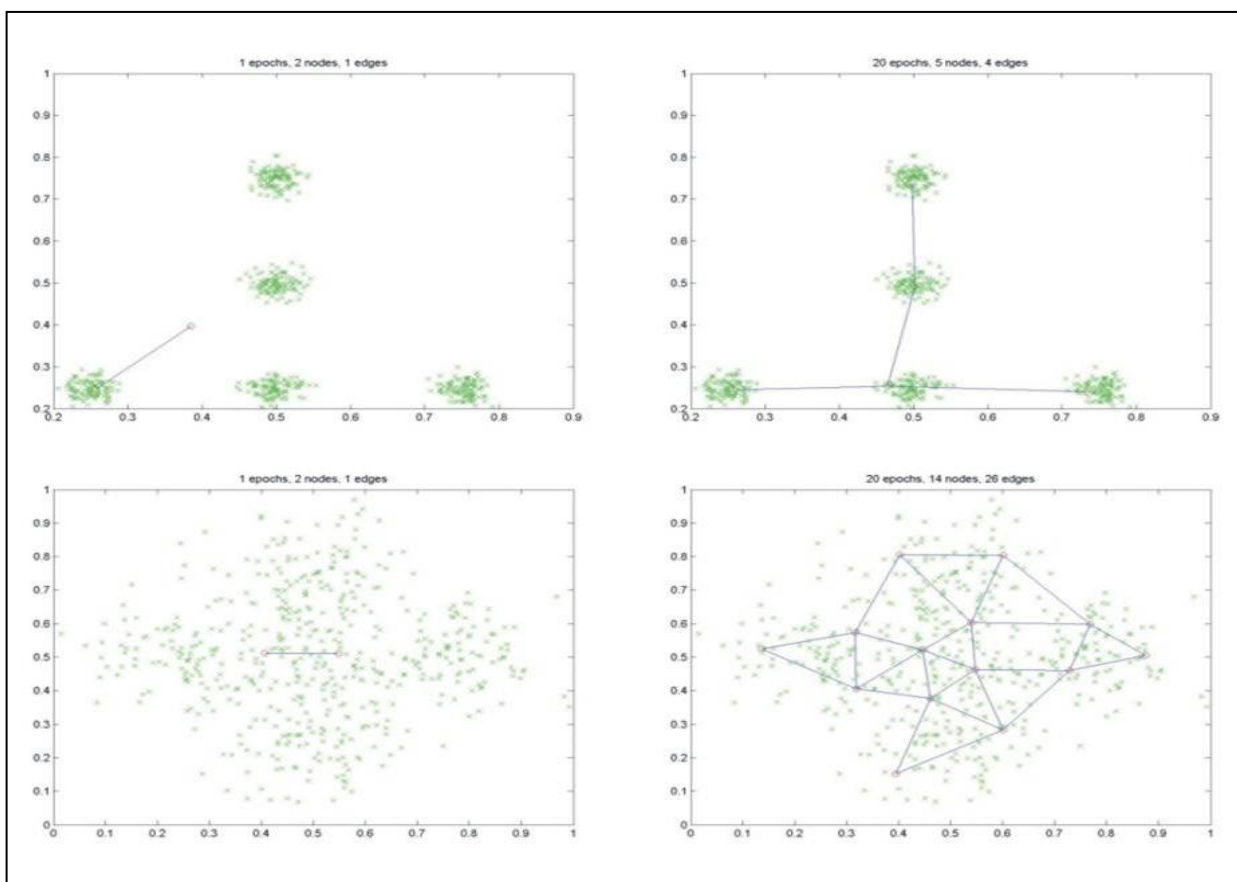
insertion threshold	počet neuronů	insertion threshold	počet neuronů	insertion threshold	počet neuronů
0.00001	3	0.00011	10	0.00021	11
0.00002	3	0.00012	11	0.00022	11
0.00003	3	0.00013	11	0.00023	13
0.00004	5	0.00014	12	0.00024	12
0.00005	5	0.00015	11	0.00025	13
0.00006	5	0.00016	11	0.00026	12
0.00007	8	0.00017	11	0.00027	13
0.00008	8	0.00018	11	0.00028	12
0.00009	9	0.00019	11	0.00029	12
0.0001	10	0.0002	11	0.0003	12

Tab. 7 Závislost počtu přidávaných neuronů na citlivosti algoritmu učení (insertion threshold). Síť byla trénována 20 epoch, learning rate 0.3 pro vítězný neuron a 0.01 pro sousední neurony. V rozmezí hodnot 0.00004 – 0.00006 dokáže síť vytvořit správný počet kategorií pro lokalistickou reprezentaci jednotlivých kategorií. Se zvyšující se hodnotou síť vytváří redundantní počet neuronů, čehož lze využít pro distribuované reprezentace.

Z tabulky lze vyčíst, že změnou parametru *citlivosti* je síť schopna vytvořit adekvátní počet výstupních kategorií na základě vstupních dat, což je ve shodě s *učitelským principem*. Pomocí změny parametru můžeme navíc modifikovat robustnost reprezentace. V případě, že počet neuronů ve výstupní vrstvě odpovídá počtu prostorových kategorií, získáváme lokalistické reprezentace, které můžeme použít jako vstup např. pro následnou symbolickou část modelu (jeden neuron ve výstupní multimodální vrstvě odpovídá jednomu prostorovému vztahu). Případnou změnou parametru a výstupní funkce vytvoříme distribuované reprezentace, tzn. každá prostorová kategorie bude reprezentována pomocí několika neuronů, v jejichž výstupní aktivitě bude zakódována prostorová poloha, což povede k větší odolnosti vůči šumu. V další fázi vývoje modelu je možné měnit nastavení parametru pomocí vyšších vrstev sítě a modifikovat tak robustnost multimodální reprezentace vzhledem k obtížnosti řešené úlohy. Podobný princip nacházíme u sítí typu ART (Carpenter, Grossberg, 2003).

10.2 Výsledek

Již po 20 epochách učení dokázala síť vytvořit optimální počet neuronů pro lokalistickou i distribuovanou multimodální reprezentaci. Model byl schopen vytvořit multimodální reprezentace pěti prostorových kategorií, které jsou obsaženy v trénovací množině. Jelikož vstupy multimodální vrstvy jsou mnohodomenzionální, je na Obr. 34 pro větší přehlednost zachycena geneze GWR sítí pro primární auditivní a vizuální reprezentace, které jsou tvořené na základě dvourozměrných vstupů.



Obr. 34 Ukázka činnosti rostoucích sítí. Na začátku jsou vytvořeny dva náhodné neurony, které jsou trénovány pomocí učení bez učitele. Pokud není kategorizace dostačující, což je řízeno parametrem *insertion threshold*, síť automaticky přidá nové neurony. Vlevo je zobrazena primární auditivní (nahore) a vizuální (dole) reprezentace na začátku učení (pouze dva neurony), vpravo po 20 epochách (auditivní vrstva obsahuje pět neuronů, což odpovídá počtu prostorových označení; vizuální vrstva obsahuje 14 neuronů, které slouží ke kategorizaci vizuálního pole na jednotlivé části). Auditivní vstupy: osa X – délka slova, osa Y – frekvenční křivka slova; vizuální vstupy – poloha bodu v 2D prostoru pomocí souřadnic X a Y . Vstupní vektory jsou označeny zeleným křížkem, neurony červenými kolečky a jejich vzájemné vazby modře.

10.3 Diskuze

Schopnost přidávat nové neurony má několik opodstatnění. Jelikož dosud neznáme přesné funkční propojení vizuální dráhy pro zpracování prostoru, snažíme se napodobit její strukturu pomocí sítě, která se sama vytváří, přičemž jsou jí prezentovány stejné podněty jako lidskému mozku. Síť adaptivně reaguje na vnější podněty a přidáváním nových neuronů hledáme jejich optimální množství pro následné testování na různých typech úloh. Zároveň eliminujeme vliv apriorních předpokladů při konstrukci systémů, jelikož model se během své činnosti modifikuje sám.

11 Závěr

V této části se pokusím na realizovaný model podívat z pohledu teoretických východisek, která byla zmíněna v souvislosti s kognitivní sémantikou, konkrétně s perceptuální teorií kognice a konceptuálními přístupy k reprezentaci znalostí.

Již ze základního popisu modelu je jasné, že je navržen v souladu s požadavky na tvorbu konceptuální úrovně na základě perceptuálních vstupů. Také užití multimodálních reprezentací znamená příklon k teoriím podobným perceptuálnímu symbolickému systému (Barsalou, 1999). Jelikož perceptuální symbol a multimodální reprezentace jsou definovány pouze teoreticky a dosud se v oblasti modelování neobjevil pokus o jejich formalizaci a převod do úrovně simulace, pokusím se jednotlivé části realizovaného systému interpretovat v kontextu těchto teorií, ale s cílem přesněji identifikovat jejich funkci. Zároveň se budu snažit interpretovat model v termínech souvisejících s ukotvením symbolů či Peircovou teorií znaku. Přestože tyto teorie používají odlišné pojmy, je možné nalézt sjednocující prvky.

Z hlediska teorie perceptuálního symbolu (Barsalou, 1999) můžeme v modelu identifikovat pouze multimodální reprezentaci, která společně reprezentuje senzorické vstupy. Mentální simulátory nebyly realizovány, jelikož schopnost produkovat „neomezené“ simulace možných stavů na základě perceptuálních vlastností leží za hranicemi současného vývoje v modelování. Simulátory navíc nejsou ve výchozím článku (Barsalou, 1999) definovány způsobem umožňujícím jejich formalizaci. Proto je třeba nalézt jiné metody, jak s reprezentacemi operovat a jak je využít v procesech vyvozování či myšlení obecně.

Podívejme se nyní na prezentovaný model z hlediska Peircovy teorie znaku. Jestliže Peirce dělí jednotlivé úrovně znaku na *objekt*, *representamen* a *interpretant*, pokusme se najít jejich koreláty v prezentovaném modelu. V případě externího *objektu* máme k dispozici pouze informace o jeho poloze v podobě vektoru dvou souřadnic. Tato vstupní informace je přesto dostatečným zástupcem objektu v externím prostředí, neboť potvrzuje existenci *objektu* a také jeho umístění ve světě, se kterým systém interaguje.

Obtížnější je v našem případě rozlišení *representamen* a *interpretantu*. Podle Peircovy teorie, je významem právě *interpretant* sdružující vlastnosti, které jsou pro externí objekt signifikantní. V konceptuálních teoriích odpovídá danému označení *koncept*, který reprezentuje vlastnosti *objektu*. *Representamen* je symbol, arbitrárně přiřazený *interpretantu*, společně vytvářející znak, který je následně užit v procesech vyvozování. Podle teorie ukotvení symbolů vzniká koncept procesy diskriminace perceptuálních informací z prostředí.

11.1 Diskriminace v symbolické úrovni

V případě hybridního systému jsou procesy diskriminace subsymbolické a identifikace přiřadí subsymbolickému konceptu symbolickou „nálepku“. V takto koncipovaných modelech (Regier, 1996; Harnad, 1990) vzniká apriorní předpoklad, že symbolická úroveň je přiřazována konceptu přímo „uvnitř“ systému, tzn. že pomocí ní designér „označí“ vytvořený koncept. Tyto přístupy vycházejí z předpokladu, že uvnitř systému je s reprezentací pracováno na základech symbolického systému, a proto je klíčovou právě symbolická vrstva.

Faktem zůstává, že při tvorbě obdobných systémů autoři zjednodušují proces akvizice symbolické vrstvy, kterou na základě designérského přístupu vkládají přímo do systému. Pro účely základních metod ukotvení je zmíněná metoda dostačující, ale unikají nám některé důležité aspekty akvizice symbolické roviny pomocí auditivního vstupu. Takový přístup je zároveň v rozporu s *učitelským přístupem*, jelikož symbolickou úroveň do systému apriorně vkládáme, místo aby vznikla na základě učení.

Odklon od podobného způsobu tvorby symbolické úrovně můžeme nejlépe demonstrovat na realizovaném modelu. V první řadě se jedná o akvizici symbolické roviny a její reprezentaci. Jestliže v předchozích modelech dochází k jejímu přiřazení, je v popsaném modelu tento způsob zajištěn pomocí učení ze vstupní vrstvy. Podobný přístup nacházíme u modelu *přenosu ukotvení* (Riga, Cangelosi a Greco, 2004). V jejich případě byl vizuální vstup reprezentován na základě SOM sítě vytvářející subsymbolické reprezentace objektů (koncept). Následně jsou tyto subsymbolické reprezentace asociovány se symbolickým auditivním vstupem pomocí MLP (Multi level perceptron). Tato fáze probíhá na základě *učení s učitelem*. V našem případě je první část architektury shodná, ale auditivní vrstva je trénována na základě *učení bez učitele*. Následné multimodální reprezentace jsou oproti Rigovu modelu opět tvořeny na principu samoorganizace (SOM). Realizovaný model je posunem k *učitelskému přístupu*, jelikož je celý založen samoorganizujícími se sítěmi. Důležité změny jsou také v chápání konceptuální a symbolické úrovně.

V realizovaném modelu je konceptuální úroveň související s reprezentací významu (*interpretant*) uložena v primární vizuální síti. Pomocí ní jsme schopni detekovat polohu objektu v prostoru, tzn. uchovávat konstantní vlastnosti polohy. Jedná se o subsymbolický koncept vznikající na základě diskriminace objektů v prostoru.

V případě auditivní vrstvy je tomu na rozdíl od výše zmíněných přístupů podobně. Přestože auditivní vstup slouží k akvizici symbolické úrovně (jazyková označení), je způsob jeho tvorby stejný jako u vizuálního vstupu. V tom je základní rozdíl oproti klasickým způsobům ukotvení.

Připomeňme si, že základní vlastností symbolů je arbitrárnost, tzn. že nemají žádné podobné vlastnosti s objektem či konceptem. Arbitrárnost ovšem neznamena, že mezi symbolickým označením jednotlivých slov neexistují žádné vzájemné vazby, popřípadě že je nelze reprezentovat v samostatné vrstvě, která dokáže zachytit vztahy podobnosti (v prezentovaném modelu na úrovni zvukové podobnosti, ale je možné zachytit vlastnosti související s reprezentací kontextových informací). Jestliže symbolickou vrstvu do systému nevkládá designér, je potřeba vytvořit způsob její reprezentace. Pro tyto účely slouží v realizovaném modelu primární auditivní vrstva. Přestože jsou symboly arbitrární vzhledem ke konceptům, ke kterým referují, mohou být navzájem propojeny např. podle fonetické blízkosti. V našem případě se jedná o podobnost na úrovni celých slov.

Takový požadavek je založen na psychologických poznatcích. Člověk dokáže volně produkovat slova, která jsou zvukově podobná, aniž by měla jakoukoliv podobnost z hlediska významu (koncepty s podobnými vlastnostmi). Příkladem může být produkce slov „banka, branka, Hanka“, které jsou významově zcela odlišné, ale člověk je dokáže produkovat velmi rychle, což svědčí o způsobu reprezentace nevyžadujícím komplexní sémantické zpracování. Proto je primární auditivní reprezentace vytvořena obdobným způsobem, tzn. že blízká jsou slova, která jsou foneticky podobná. Pokud bychom se chtěli přiblížit psychologické plauzibilitě ještě více, museli bychom symbolické označení reprezentovat na úrovni slabik. Člověk kromě produkce foneticky podobných slov dokáže pomocí kombinace slabik vytvářet také zcela nová označení, která nerefekují k žádnému konceptu ani nejsou podobná známému symbolickému označení, tzn. lze je považovat za nesmyslná. I tato produkce je dostatečně rychlá a automatická, takže svědčí o samostatné reprezentaci symbolické úrovně. Příkladem je schopnost produkovat slova jako „jala, bala, bola“, aniž bychom je někdy předtím slyšeli. Přesto je pro účely modelu dostačující reprezentace v úrovni slov, jelikož se jedná o zjednodušenou demonstraci primární reprezentace symbolické úrovně a základní analýzu procesu ukotvení symbolů.

Hlavním důvodem zvoleného způsobu je schopnost reprezentovat auditivní úroveň samostatně a vytvářet vztahy mezi symbolickými výrazy na základě jejich vzájemné tvarové (zvukové) podobnosti. Dalším důležitým aspektem primární auditivní reprezentace je vzájemná oddělenost jednotlivých slovních označení, tzn. že jsou reprezentována tak, abychom je mohli rozdělit do ohraničených a nepřekrývajících se kategorií. Tato zajistí jednoznačné odlišení

objektů okolního světa a vytvoření znaku (reprezentace obsahující konceptuální a symbolickou úroveň, v našem případě integrovanou v multimodální reprezentaci).

Během testování modelu pro statické prostorové vztahy (Vavrečka, 2007) se objevují pokusy měnit velikost shluků (tedy fonetických rozdílů) auditivního vstupu až do úrovně, kdy se odlišná označení překrývají. Taková situace znamená, že bychom na vstupu přijímali natolik nejednoznačná slova, že nelze určit, do které kategorie spadají. V případě prostorových vztahů by se jednalo o slova jako „vlelu“ nebo „dovo“, vytvořené průnikem kategorie vlevo a dolů, tzn. že auditivní kategorie začínají být „fuzzy“. Pokud bychom během učení přijímali vstupy, které vytvářejí překrývající se shluky v auditivní i vizuální vrstvě, pak není možné naučit systém správně kategorizovat ve vrstvě multimodální. Zdůvodnění nalezneme na základě prosté úvahy. Jestliže máme dva nejednoznačné vstupy, které nelze kategorizovat do oddělených tříd, nelze z nich bez doplňujících informací odvodit jednoznačné kategorie.

Z těchto úvah lze vyvodit několik závěrů týkajících se vztahu mezi jazykem a obrazem (lingvistickou a nelingvistickou reprezentací, symbolickou a konceptuální vrstvou) a způsobu, jakým se navzájem ovlivňují a jaká je jejich vzájemná hierarchie. Jedná se o ustavení vztahu mezi koncepty (nelingvistické) a symboly (lingvistické). Již psychologické studie (Landau, Hofmann, 2005) týkající se Williamsova syndromu (genetická vada ovlivňující vnímání prostoru při zachování jazykové produkce) prokázaly, že jednotlivé subsystemy se mohou během činnosti vzájemně doplňovat a nahrazovat deficit jednoho z nich. Z výsledků tedy vyplývalo, že jeden systém není nadřazen druhému, ale že fungují společně. Pomocí zmíněného experimentu (Landau, Hofmann, 2005) ovšem nelze přesněji určit, jakým způsobem se systémy vzájemně strukturují. Zkusme se nejprve podívat na problematiku z lingvistického hlediska.

V nejobecnější rovině se o vztahu lingvistických a nelingvistických reprezentací vyjadřuje Sapir-Whorfova hypotéza. Jedná se o lingvisticko-antropologickou tezi (Whorf, 1956), podle níž je pojetí reálného světa vystavěno na jazykových zvyklostech konkrétní komunity, jež pak předurčují výběr interpretace reality. To bývá v oblasti reprezentace znalostí a významu interpretováno jako nadřazenost lingvistických reprezentací nad nelingvistickými. Příkladem v oblasti modelování může být mnohokrát zmiňovaný Harnadův model (1990).

Odpůrci Sapir-Whorfovy hypotézy (Gumperz, 1996) argumentují tím, že myšlení není zcela redukovatelné pouze na jazyk a řada významů může být vyjádřena neverbálními či mimojazykovými prostředky, které se mohou podílet při realizaci myšlenkových operací. Jelikož z těchto poznatků opět nedokážeme určit vzájemný vztah těchto reprezentačních vrstev, pokusím se nyní interpretovat realizovaný model a jeho přínos pro danou oblast v kontextu zmíněných teorií.

V první fázi docházíme ke shodnému závěru jako zastánci Sapir-Whorfovy hypotézy, jelikož lingvistické označení v modelu strukturuje nelingvistické koncepty. Systém pracuje na základě ohraničených jazykových vstupů (slovo je přesně identifikovatelné) a překrývajících se prostorových kategorií. Z takových výsledků konkrétního modelu můžeme dojít k závěru, že lingvistický systém strukturuje nelingvistický právě pro svou jednoznačnost a ohraničenost. Přesto nelze podobný závěr přijmout a zobecnit pro celou oblast ukotvení symbolů.

Některá jazyková označení totiž bývají polysémická, tzn. jsou propojená s rozdílnými koncepty (např. klíč jako označení rozdílných předmětů). V takovém případě ovšem jazykové označení přesně nevymezuje koncept, ale stává se nejednoznačným. Přestává platit vztah nadřazenosti lingvistické reprezentace nad nelingvistickou. V dané situaci se systém řídí konceptuální nelingvistickou informací, pomocí které strukturuje způsob propojení se symbolickou rovinou. Konceptuální úroveň je totiž pro polysémická slova jednoznačná (koncepty jsou na rozdíl od identického označení rozdílné a lépe odlišitelné), a proto má během učení větší vliv na kategorizaci. Jelikož neexistují polysémická slova, která by referovala k překrývajícím se konceptům (což je ověřeno v Modelu 1, který nedokáže kategorizovat dva „fuzzy“ vstupy), docházíme k závěru, jenž popírá privilegovanost jednoho ze systémů. Můžeme tedy shrnout:

V oblasti reprezentace významu se na způsobu kategorizace na základě symbolického a konceptuálního vstupu podílí ten, jehož vstupy jsou jednoznačně rozlišitelné, jelikož jednoznačné kategorie determinují způsob tvorby multimodálních reprezentací. Neexistuje předem daná privilegovanost.

Dané tvrzení tvoří jádro navrženého přístupu k reprezentaci významu, založeném na multimodálních reprezentacích pomocí systému dvou primárních reprezentací. Apriorní privilegovanost jednoho systému je v rozporu s psychologickými poznatky o způsobu reprezentace znalostí u člověka. Naopak pomocí systému, který kategorizuje na základě výše zmíněného principu, dokážeme vysvětlit mnoho teoretických poznatků z oblasti experimentální psychologie.

Jestliže panují rozpory o privilegovanosti lingvistických a nelingvistických subsystémů, a to na základě rozdílných výsledků experimentů s odlišně definovanými úlohami (Landau, Hoffman, 2005), lze pro jednotnou interpretaci užít výše popsaného *principu privilegovanosti jednoznačného vstupu*, jelikož dokáže eliminovat vliv typu úlohy a experimentálních podmínek na výsledky výzkumu. Správnost navrženého principu je navíc ověřena konkrétní realizací formou modelu, jehož efektivita fungování verifikuje princip v empirické rovině. Právě provedení a odzkoušení modelu na konkrétní úloze jej činí argumentačně silnější oproti rozpor-

ným teoretickým předpokladům. Na druhou stranu je nutné říci, že model byl zkoušen pouze v konkrétní oblasti vnímání prostoru, ve které se neobjevují jednoznačné vstupy konceptuální reprezentace, jelikož podobný typ úlohy se v této doméně nevyskytuje. Přesto je jasné, že z hlediska (obecné matematické) teorie platí *princip privilegovanosti jednoznačného vstupu* také v nerealizované variantě (konceptuální vstup jednoznačný, kdežto symbolický nejednoznačný), protože obě primární reprezentační vrstvy mají totožnou architekturu a také způsob propojení s multimodální vrstvou (viz Model 1). Daný způsob propojení navíc činí architekturu univerzálně použitelnou. Další výhodou je možnost rozšířit počet modalit, takže je možné společně reprezentovat vizuální, auditivní, motorické a introceptivní informace ve společném multimodálním systému. Pokud zachováme *princip privilegovanosti jednoznačného vstupu*, bude zajištěna tvorba přesně identifikovatelných kategorií.

Můžeme tedy shrnout, že způsob primární reprezentace je odlišný od dosud realizovaných modelů. Harnad (1990), Regier (1996), Siskind (2001) aj. definují symbolickou vrstvu jako apriorně danou a omezují již v počátku možnost systému fungovat podle *principu privilegovanosti jednoznačného vstupu*, jelikož jednoznačný vstup je dán fixně. Systém nedokáže řešit úlohy obsahující nejednoznačně definované symbolické označení.

Systém realizovaný v této práci je svou strukturou blízký Dorffnerovu řešení (1996), založeném na plně konekcionistickém modelu. Již v jeho práci se objevuje odlišné pojmání reprezentace symbolické roviny a procesu diskriminace. Podobně je tomu se způsobem propojení symbolické a konceptuální úrovně do jedné společné vrstvy. Jelikož svůj model realizoval ještě před teorií perceptuálního symbolu (Barsalou, 1999), není společná vrstva nazývána multimodální. Dorffner používá diskriminaci pro tvorbu auditivních (symbolických) kategorií, které mapuje do jednorozměrné sítě SOM, což je ojedinělý přístup k ukotvení. Činnost modelu však interpretuje pouze v kontextu základního principu ukotvení a nepřichází s návrhem *privilegovanosti jednoznačného vstupu*, přestože by ho architektura měla být schopná. Důvod můžeme hledat v omezeném zadání úlohy, které podobný princip pro řešení úlohy nevyžaduje, popřípadě jej teoreticky nedopracoval. Přestože jeho řešení obsahuje latentní schopnost vytvářet kategorie na základě jednoznačného vstupu, potenciál daného přístupu je prezentován až v této práci.

Z hlediska procesu kategorizace tedy můžeme shrnout, že klasické modely ukotvení (Harnad, 1990; Siskind, 2001; Regier, 1996) se ve světle psychologických teorií jeví jako nedostačující, jelikož pomocí nich nedokážeme řešit široké spektrum úloh, kterému je lidský kognitivní aparát během svého života vystaven. Proces diskriminace musí být proveden v obou úrovních znaku (symbolické i konceptuální), abychom mohli pro jejich následnou integraci využít vlastnosti, které jim umožní efektivně vytvářet společné propojení.

11.2 Identifikace

Jestliže jsme podrobněji analyzovali prvotní reprezentaci symbolů a konceptů, můžeme přikročit ke druhé části procesu ukotvení. Nyní pozornost zaměříme na oblast identifikace symbolů a konceptů, která je pro realizovaný model výpočetně nejnáročnější. Pro lepší orientaci bude interpretace probíhat v kontextu teorie znaku a také teorie ukotvení symbolů.

Primární vizuální reprezentace je možné nazírat jako *interpretant* (koncept či konceptuální úroveň), který vznikl na základě procesu diskriminace, tzn. první části procesu ukotvení. Primární auditivní reprezentace je totožná s *representamen* (symbol či symbolická úroveň) a je také vytvořena na základě procesů diskriminace.

Připomeňme si v této chvíli Harnadovo (1990) dělení diskriminace a identifikace. Diskriminace je posouzení, zda jsou dva vstupy stejné, nebo rozdílné, popřípadě jak rozdílné. Jedná se o relativní posouzení, tzn. hledání míry shody. Identifikace je přiřazení arbitrární odpovědi – „jména“ - klasifikující vstupy, které následně tvoří kategorii podle shodného prvku. Identifikace je absolutní posouzení, které rozhodne, zda je vstup členem kategorie, či nikoli (Harnad, 1990).

Pokud se na definice podíváme optikou realizovaného modelu, potřebují pojmy přesnější vymezení. Přestože je diskriminace posouzení míry shody, vytváří ohraničené kategorie již v primární auditivní (symbolické) vrstvě. I když se podle definice diskriminace jedná o relativní posouzení, dostáváme v tomto případě absolutní kategorie. Realizovaný model se začíná odlišovat od Harnadova vymezení. Rozpor můžeme hledat ve způsobu definice pojmů. Ve spojitosti s procesem identifikace je nutné zdůraznit, že pro její provedení je nutné použít **dvou odlišných zdrojů** (reprezentačních úrovní) referujících k totožné entitě či procesu v externím prostředí, přičemž identifikace znamená jejich vzájemné mapování. Rozdíl mezi diskriminací a identifikací je tedy hlavně ve způsobu aplikace. Diskriminace probíhá v rámci **jednoho zdroje** vstupu, který je kategorizován pouze v rámci sebe sama. Pro identifikaci jsou nutné dva vstupní zdroje, u kterých hledáme vzájemné propojení.

V realizovaném modelu tedy probíhá diskriminace samostatně pro vizuální a auditivní vstup a identifikace tvoří propojení mezi jejich primárními reprezentacemi. V Harnadově návrhu se zdůraznění tohoto faktu neobjevuje, přesto je z hlediska realizace klíčový. Pouhé dělení na relativní a absolutní posouzení neimplikuje počet reprezentačních vrstev.

Samotný proces identifikace probíhá v realizovaném modelu dle principu *privilegovanosti jednoznačného vstupu*. V oblasti reprezentace prostoru je vždy jednoznačný auditivní symbolický vstup, přesto je pro obecnou interpretaci nutné připomenout variantu funkce systému,

při které je jednoznačná vizuální kategorie a nejednoznačná auditivní (symbolická) úroveň, tzn. že multimodální reprezentace je strukturována na základě vizuálního vstupu (konceptuální reprezentace).

Můžeme tedy říci, že proces identifikace nemusí probíhat pouze na základě jednoznačné symbolické informace, ale že je možné provádět identifikaci na základě konceptuální informace, přičemž rozhodující je *privilegovanost jednoznačného vstupu*. Oproti Harnadovu způsobu identifikace podle symbolického vstupu je v naší teorii identifikace závislá na tom, ze kterého vstupu přicházejí jednoznačné informace. Proces identifikace není fixně dán, ale je závislý na povaze vstupů z různých modalit. Obecně lze říci, že nezávisí na počtu modalit, jelikož na způsobu identifikace a struktura multimodální vrstvy se podílejí vždy pouze jednoznačné vstupy.

Z hlediska způsobu zpracování informace je proces identifikace jádrem celého systému a nejobtížnější a výpočetně nejnáročnější úlohou. Jedná se o nalezení způsobu mapování mezi konceptem a symbolem. Jestliže je v Harnadově případě tento proces specifikován jako jednoznačné označení kategorie, v realizovaném modelu přistupují k problematice na základě předpokladů zmíněných výše. Harnad provádí symbolické přiřazení jako asociaci řetězce znaků s konceptem reprezentovaným neuronovou sítí (hybridní model). V našem případě vzniká propojení na základě učení a jeho výsledkem je multimodální vrstva. V ní se propojují informace z primárních systémů a jsou na základě učení identifikovány jako jednoznačné kategorie. Rozdíl je v užití metody učení v celém systému. Jedná se o zdůraznění *učitelského přístupu*, který v našem případě přináší výhody v podobě principu *privilegovanosti jednoznačného vstupu*, který v klasických modelech (Harnad, 1990; Regier, 1996) nelze realizovat kvůli *designérskému přístupu* a fixnosti a apriornosti reprezentace symbolického vstupu. Právě samoorganizující se síť umožní systému univerzálně kategorizovat dle rozličných vstupů. Navíc lze systém rozšířit i o další modalit, v čemž spatřuji výhodu a přínos realizovaného modelu v oblasti ukotvení symbolů.

11.3 Kritika navrženého přístupu

Pro komplexní posouzení interpretuji realizovaný model také z hlediska možné kritiky navrženého přístupu. První námitka může zpochybnit způsob reprezentace symbolické roviny pomocí samoorganizující se sítě, která reprezentuje označení v úrovni slov. V první fázi je možno napadnout reprezentaci celých slov pro jejich neschopnost zachytit zvukovou podobu na úrovni fonému. Odpovědí je fakt, že změna způsobu reprezentace z úrovně jednotlivých slov

do úrovně fonémů je pouze technickou záležitostí související s rozšířením architektury modelu. V budoucí verzi je s tímto způsobem reprezentace počítáno, jelikož může odhalit mnoho zajímavého o způsobu tvorby slov. Důvodem redukce na reprezentaci celých slov během současného řešení byla výpočetní náročnost modelu založeného na reprezentaci fonémů, který je v oblasti ukotvení symbolů nadbytečný, jelikož nás v první fázi zajímají podrobnosti týkající se procesu diskriminace a identifikace.

Přechod na reprezentaci v úrovni fonémů by neměl být obtížný, a to díky existujícím modelům, které se danou problematikou zabývají. Způsobem reprezentace symbolických označení na úrovni jednotlivých písmen či fonémů se ve svých pracích zabývá Farkaš (2003, 2007) a užitím principu samoorganizace dosahuje výsledky, které lze užít a propojit s navrženým modelem.

Další námitkou týkající se reprezentace slov pomocí samoorganizujících se map je zbytečná reprezentační mohutnost a nadbytečnost auditivního (symbolického) vstupu. Jestliže se jedná o arbitrární symbolický vstup, stačí jej reprezentovat pomocí sekvence symbolů, která jej bude jednoznačně identifikovat. Použití samoorganizace může být považováno za neefektivní, jelikož nám nepřináší žádné další výhody pro následné zpracování.

Taková námitka je oprávněná, jelikož pro úlohu reprezentace prostorových vztahů, ve které se vyskytují pouze jednoznačné symbolické vstupy, není potřeba jejich reprezentace pomocí tvorby 2D mřížky vytvářející „asociační mapu“. Model však zastupuje obecnější přístup k procesu ukotvení, a proto je navržen tak, aby byl schopen reprezentovat v multimodální vrstvě také nejednoznačné symbolické vstupy, tzn. aby kategorizoval na základě konceptuálního vstupu. Pro lepší orientaci uvádím rozdíl v možných přístupech k problematice.

Harnadovo základní řešení a od něj odvozené způsoby ukotvení vycházejí z předpokladu, že symbolické označení je vždy jednoznačné a dokáže přesně identifikovat koncept, tzn. přiřadit kategorii či prototypu jednoznačnou symbolickou nálepkou, se kterou je následně operováno v symbolickém systému. Nejčastějšími vstupy takové formy ukotvení mohou být jednotlivé objekty externího prostředí, jejichž koncepty vytvořené na základě perceptuálních vstupů a za pomoci procesu diskriminace jsou považovány za překrývající se prototypy (není jasná hranice prototypu kočky a rysa, popřípadě jiného zvířete). Naopak symboly jsou vždy jednoznačné označení (označení kočka, rys, pes apod.), které odliší třídu či koncept (nejednoznačný). Proto je Harnadův model vhodný pro označení jednotlivých objektů, tříd objektů, popřípadě jednoznačných procesů.

Pokud bychom chtěli percentuálně vyjádřit zastoupení takových objektů a procesů během normálního způsobu interakce s okolím, těžko bychom v současných studiích našli odpovídající statistiku. Přesto můžeme říci, že se s danou formou znaků (jednoznačný symbol, nejed-

noznačný koncept) setkáváme ve většině případů. To nás ale neopravňuje říci, že je Harnadův model obecným řešením procesu ukotvení. Lidský kognitivní aparát je totiž schopen reprezentovat také objekty a situace mimo rámec Harnadova modelu.

V prvé řadě se jedná o polysémická slova. Jestliže je symbol identifikován s různými koncepty, nelze o takovém označení říci, že je jednoznačné. Oproti předchozí situaci, kdy byla konceptuální vrstva nejednoznačná, nastává podobná situace v symbolické vrstvě. Rozdíl je také v míře nejednoznačnosti. Jestliže u konceptuální úrovně se kategorie překrývají, u symbolické úrovně můžeme hovořit (v případě reprezentace pomocí fonetických vlastností) o identických kategoriích, jelikož polysémická slova (např. jazyk, klíč, dát apod.) nevykazují v symbolické úrovni žádné rozdíly. Takové vstupy nedokáže Harnadův model zpracovat díky své architektuře a hierarchii úrovní. Harnad postuluje symbolickou rovinu hierarchicky nejvýše, jelikož ta následně slouží pro operace v rámci formálního symbolického systému. Model realizovaný v této práci funguje odlišně. Hierarchicky nejvýše stojí multimodální vrstva, která vytváří propojení mezi konceptuální a symbolickou úrovní. Právě ona dokáže absolutně odlišit jednotlivé znaky (tvořené propojením konceptu a symbolu).

Při porovnání zmíněných přístupů docházíme k překvapivým závěrům. Pokud bychom interpretovali multimodální vrstvu z hlediska Harnadovy teorie, jednalo by se o proces identifikace vedoucí k ukotvení symbolické vrstvy. V našem modelu je multimodální vrstva procesem identifikace a zároveň vrchní vrstvou celého systému, která tvoří podklad pro následnou manipulaci (např. pomocí formálního symbolického systému). To znamená, že navržený model ukotvení obsahuje samostatnou reprezentační úroveň, která identifikuje symboly s koncepty. Identifikace je procesem a produktem této činnosti ve formě multimodální reprezentace. Je schopná vytvářet vždy jednoznačná označení - na rozdíl od symbolické úrovně, která selhává v případě polysémie. Pomocí přidání této třetí (multimodální vrstvy) vytváříme systém, který je univerzálně použitelný v procesu ukotvení. Změnám v nazírání reprezentace významu užitím navrženého modelu se budeme věnovat v samotném závěru.

Podívejme se mezitím podrobněji na další aspekty polysémie. Kritika může namítat, že polysémická slova tvoří pouze malé procento výrazu z celkového slovníku a že je možné tyto výrazy zanedbat. Tímto krokem však zavíráme cestu k napodobení lidského kognitivního aparátu, jelikož Harnadův systém není schopen podobné úlohy zvládnout. Na tento argument může kritika namítat, že polysémická slova je možné rozlišit na základě analýzy kontextu dané promluvy, ze které zjistíme, který koncept je potřeba symbolickému označení přiřadit. Takové argumentaci lze oponovat v několika úrovních.

V elementární úrovni můžeme použít následující příklad. Jestliže jsou nám během fáze učení prezentována pouze jednoslovná symbolická označení objektu (např. jazyk) a ostenze objektu,

nelze kontextuální analýzu použít, jelikož symbolická (auditivní) informace obsahuje pouze tuto promluvu. Systém nedokáže provést jednoznačnou identifikaci konceptu se symbolem. Pokud použijeme architekturu realizovanou v této disertaci a modifikovanou pro oblast reprezentace objektů, lze pomocí konceptuální úrovně identifikovat koncept se symbolem na základě odlišnosti konceptuálních kategorií (např. při promluvě je mi ukázán klíč od bytu nebo klíč na utahování), jejichž perceptuální vlastnosti a následně tvořené koncepty jsou natolik odlišné, že jsou klíčové pro proces identifikace v multimodální vrstvě. Identifikace závisí na jednoznačnějším, nikoli na apriorně privilegovaném vstupu, a proto jsou oba vstupy v multimodální vrstvě zareprezentovány odlišně.

Případ, ve kterém nelze provést rozlišení na základě analýzy kontextu, je ale příliš elementární a můžeme se s ním setkat pouze během raného vývoje dítěte. Proto se pokusíme podrobněji analyzovat vliv kontextu, který nám umožní komplexnější pohled na fungování modelů ukotvení. Nejprve si načrtneme způsob zpracování (polysémických) slov v kontextu promluvy. Před jejím začátkem je potřeba specifikovat, jak definujeme kontext. Z našeho pohledu můžeme brát kontext jako vliv vzájemného působení časově blízce prezentovaných označení na význam, který v promluvě mají. Laicky řečeno, společný výskyt slov tvoří kontext, pomocí kterého lze specifikovat význam jednotlivých označení. Jedná se o elementární formu kontextu, při které nebereme v potaz subjektivitu, intencionální postoje mluvčího a významy slov v úrovni obecnějších lingvistických kategorií (téma, žánr, styl). Jedná se nám o analýzu kontextu v úrovni jednoduchých promluv, pomocí které je možné upřesnit významy (polysémických) slov. Pokusme se nalézt rozdíly v analýze kontextu za použití Harnadova a vlastního modelu.

V případě prvního zmíněného máme reprezentovány polysémická slova pomocí symbolického označení, identifikující více rozdílných konceptů, které slouží k reprezentaci subsymbolických vlastností, a odvozeně tedy i významu. Jestliže potřebujeme určit, ke kterému konceptu referujeme, můžeme analyzovat slova související s daným symbolickým označením, vyskytující se v promluvě společně. V symbolické rovině provedeme logickou analýzu určující platnost výroků v dané promluvě či soustavě promluv (výroků) a přiřadíme jim pravdivostní hodnoty. V Harnadově modelu se nejedná o aplikaci symbolického systému dle pravidel formální sémantiky (pravdivost je dána na základě nerozpornosti a kompletnosti), ale o aplikaci sémantiky, která bere v potaz konceptuální vlastnosti spojené s daným symbolem. Kontext nám slouží jako způsob testování pravdivosti pro různé koncepty, se kterými je polysémické slovo spojeno. V symbolickém systému dochází k ověřování a identifikaci konceptu, který se pro dané symbolické označení hodí nejlépe. Pro tyto účely již nestačí užití pravdivostních hodnot pravda/nepravda, ale na základě odlišných kombinací vlastností konceptů je potřeba vyjádřit

míru konzistence výroku na zvolené škále, nejčastěji (0;1). Jedná se o posun směrem k *fuzzy logice*.

V případě užití Harnadova modelu se můžeme setkat v oblasti zpracování kontextu s obtížemi, které jsme analyzovali již dříve. Jestliže jsou symboly definovány jako fixní a apriorní označení, znamená analýza kontextu výpočetně velmi náročný proces, jelikož je potřeba analyzovat vliv kontextu na význam slov „testováním“ pomocí míry pravděpodobnosti, určující koncepty, které jsou pro daná symbolická označení nositeli významu. To je dáno způsobem reprezentace symbolické roviny a užitím pouze dvou reprezentačních úrovní, ze kterých je jedna brána jako fixní a absolutní z hlediska jednoznačnosti. V případě, že bychom použili Harnadem navrženou symbolickou vrstvu jako vrchní úroveň, která díky své jednoznačnosti slouží k manipulaci s reprezentacemi, znamenalo by každé polysémické slovo pro systém extrémní výpočetní zátěž, jelikož by nebylo možné tyto výjimky systém naučit. Nedostatek vyplývá z fixní symbolické roviny a způsobu jejího propojení s konceptuální úrovní a je v rozporu s poznatky z oblasti psychologie, protože člověk dokáže zpracovávat odlišné významy slov velmi rychle a automaticky.

Způsob zpracování v realizovaném modelu je jednodušší. Hlavní přínos spočívá v užití multimodální vrstvy, která se dokáže naučit odlišnosti slov na základě jak vizuálního, tak auditivního vstupu. Právě přidání vrstvy sloužící k jednoznačné kategorizaci je jádrem systému schopného univerzálně reprezentovat okolní prostředí. Výhody systému se objevují i v případě potřeby rozlišit polysémické slovo na základě analýzy kontextu promluvy. Opět využijeme schopnosti systému vytvářet reprezentace pomocí metod učení umožňující reprezentovat společný výskyt slov a následně určit význam jednotlivých slov v promluvě. Oproti fixní symbolické rovině u Harnada můžeme pomocí primární auditivní (symbolické) reprezentace zachytit nejen fonetické vlastnosti slov, ale také vytvářet asociační mapy na základě jejich vzájemného výskytu. Jedná se pouze o přidání dalších vlastností, které během učení reprezentujeme v primární auditivní vrstvě. Stačí použít další dimenze, které slouží ke sdružování podobných slov dle blízkosti v promluvě. Tyto vlastnosti pak slouží v procesu identifikace s konceptuální rovinou, jelikož dokážou strukturovat tvorbu multimodální vrstvy a odlišit polysémická slova na základě kontextu. Rozdíl oproti Harnadovu modelu je ve způsobu činnosti modelu. Není nutné pokaždé provádět analýzu kontextu na základě testování vhodných konceptů, ale je možné reprezentovat časté společné výskyty slov (kontext) v primární auditivní vrstvě pomocí metod učení a následně tyto informace použít k tvorbě jednoznačných kategorií v multimodální vrstvě odlišující polysémická slova. Hlavním rozdílem je opět užití metod učení oproti fixním mechanismům.

11.4 Sémantika modelu ukotvení

Skrze ukázky odlišnosti modelů ukotvení se dostáváme k nejobtížnější pasáži týkající se samotné reprezentace významu. Abychom dokázali spolehlivě určit, co pojem reprezentace významu znamená, bude třeba provést odlišení na základě identifikace rozdílů mezi Harnadovým přístupem a realizovaným řešením. Podívejme se nejdříve na prvé zmíněné. V jeho pojetí dochází ke tvorbě peircovského znaku, který je tvořen na základě konceptu a symbolu. Způsobem propojení je proces identifikace jednoho s druhým. Vzniká systém uspořádaných dvojic, přičemž první část zachycuje subsymbolické konstantní vlastnosti objektu v externím prostředí a druhá je tvořena arbitrárním shlukem symbolů, jednoznačně určující daný objekt. Takto vybudovaný systém je následně připraven vykonávat základní myšlenkové operace na základě manipulace s ukotvenými reprezentacemi. Význam je zde reprezentován v konceptuální úrovni (*interpretant*) a zajišťuje korespondenci interního systému reprezentací s externím prostředím.

Pokud přistoupíme k následné tvorbě myšlenkových operací, nenacházíme u Harnadova přístupu (1990) příliš mnoho informací o jeho funkci. Ve svém hybridním systému navrhuje realizovat myšlenkové procesy pomocí pravidel inference, se kterými jsme se setkali v oblasti formálního symbolického systému. Základem jsou pravidla vyvozování, která ze vstupních výroků vytvoří závěr. Harnad blíže neurčuje, jakým způsobem proces probíhá. Ve svém návrhu zdůrazňuje nutnost formálního symbolického systému jako jediného způsobu manipulace s interními reprezentacemi.

„Sémantická interpretovatelnost, aby byla symbolická, musí být propojena s explicitní reprezentací, syntaktickou manipulovatelností a systematicitou. Žádné z těchto kritérií není arbitrární, a pokud je zlehčíme, ztratíme možnost pracovat s pojmy jako přírodní kategorie a také možnost operovat na základě formální teorie výpočtu, čímž se ze slova „symbolický“ stává nevysvětlená metafora (která se liší interindividuálně). Proto budeme přemýšlet pouze o klasickém formálním systému (tak, jak je definován), chceme-li ukotvovat symboly.“ (Harnad, 1990, s. 6)

Tuto pasáž lze považovat za klíčovou. Hovoří o tom, že neexistuje alternativa k formálnímu systému, kterou bychom mohli použít jako podklad pro myšlenkové operace. Problematické je určit, jaký druh sémantiky měl Harnad na mysli v souvislosti se svým hybridním modelem. Jestliže se jedná o pravdivostní hodnoty, které podporují konzistentnost a úplnost systému výroků, nepotřebuje systém konceptuální úroveň zajišťující korespondenci s externím světem.

Takový systém se nijak neliší od čistě symbolických systémů. Harnadův model ovšem konceptuální úroveň obsahuje.

Rozdílnost spočívá ve způsobu, jakým přiřadíme pravdivostní hodnotu výrokům, na jejichž základě chceme provést inferenci. Můžeme použít konceptuální úroveň, tudíž nepotřebujeme systému interpretovat význam a pravdivost výroků externě. Systém vyvozuje na základě pravdivosti získané z perceptuálních dat a reprezentovaných konceptuálně. Použitím obdobného systému se ale dostáváme do obtíží. Jestliže má být systém konzistentní a úplný, je potřeba, aby tuto vlastnost měly i koncepty, které přiřazují pravdivost výrokům. Jelikož víme, že perceptuálně získané koncepty obsahují velké množství nepřesností, redukcí a mnohoznačností, je tato konzistence nerealizovatelná. Přestože Harnad tvrdí, že symbolická rovina je jednoznačná a přesně identifikuje kategorie, je nutné si uvědomit, že význam je uložen v ne jednoznačné konceptuální rovině (viz intenzionální vágnost). Symbolický systém manipuluje s významy konceptuální roviny, ale není na ní nezávislý. Naopak se jedná o roviny, které jsou ve vzájemné korespondenci. Proto není možné nekonzistenci a nepřesnost konceptuální roviny pomocí symbolického systému redukovat. Jestliže jsou kategorie nejednoznačné, překrývající se a obsahují chybějící informace, nelze předpokládat, že v symbolickém systému lze tyto vlastnosti redukovat a dospět k logicky jednoznačnému závěru *pravda/nepravda*. Přestože Harnad tvrdí, že procesem identifikace získáme jednoznačné kategorie, zjistili jsme, že tato jednoznačnost neplatí pro všechny případy (polysémie, kontext). V případě překrývajících se konceptuálních kategorií je možné, že systém identifikuje nejednoznačné vstupy s nesprávným symbolickým označením, a nemůže být úplný a konzistentní. Máme zde tedy tři faktory (polysémie, kontext, nejednoznačnost konceptů), které způsobují, že použití klasické sémantiky pravdivostních hodnot *pravda/nepravda* je pro Harnadův model nerealizovatelné.

Alternativou je možnost užití *fuzzy logiky* v procesech vyvozování. V takových systémech je konzistentnost systému posunuta do roviny pravděpodobnosti, výroky mohou nabývat pravdivosti na škále hodnot (nejčastěji 0;1) a systém je schopen vytvářet inference, pro které nebude problematická nekonzistentnost konceptuální úrovně. Logická konsistence je zde nahrazena pravděpodobností. Díky fixnímu způsobu reprezentace symbolické roviny si však podobný systém nedokáže poradit s polysémií či vlivy kontextu.

Shrňme tedy, že v Harnadově přístupu je význam reprezentován v konceptuální vrstvě, která slouží jako podklad pro následnou pravdivostní hodnotu symbolických označení. Pomocí pravidel vyplývání jsou vytvářeny myšlenkové operace a jejich závěry. Dosud jsme se však nezmínili (Harnad tuto informaci neuvádí), jakým způsobem je reprezentován význam produktu myšlenkových operací. Abychom zajistili zpětnou korespondenci s konceptuální úrovní, je opět nutné do ní závěr inference namapovat. V případě, že odpovídající koncept neexistuje, je

nutné jej vytvořit na základě stávajících konceptů. V případě, že koncept existuje, je nutné závěr v symbolické rovině porovnat s vlastnostmi v rovině konceptuální. Tím ale narážíme na problém zpětného převodu symbolické roviny do konceptuální. Jestliže je výrok pravdivý na základě pravidel vyplývání, jaké to má implikace pro konceptuální úroveň?

Přestože celá disertační práce byla věnována procesu propojení konceptuální úrovně se symbolickou, neobjevily se zde žádné informace o navazující části, tzn. zajištění způsobu převodu logických operací do úrovně konceptů. Jelikož se jedná o oblast, která není přesněji popsána ani v úrovni teorie, byly by další vývody v úrovni spekulace. Přesto této oblasti přiřkládám prioritní význam a budu se jí věnovat v souvislosti s rozvojem modelu reprezentace prostorových vztahů.

Vraťme se nyní k původnímu tématu, kterým je porovnání Harnadova a vlastního modelu z hlediska reprezentace významu. První zmíněný jsme již popsali, takže můžeme přikročit k definici rozdílů. I v případě vlastního modelu je naplněna podstata Peircovy teorie znaku, jelikož systém obsahuje symbolickou a konceptuální úroveň, která je založena na informaci z externího prostředí. Kromě zmíněných dvou však systém obsahuje také multimodální vrstvu, která nekoresponduje s klasickou teorií znaku. Multimodální vrstva slouží k propojení konceptuální a symbolické úrovně. Takové propojení není u Peirce přesněji specifikováno, ani neobsahuje žádný název, pokud jej ovšem nepovažujeme za samotný znak. U Harnada můžeme multimodální vrstvu ztotožnit s procesem identifikace. Narážíme na rozdíl, jelikož v modelu realizovaném v této práci se nejedná o pouhý proces, ale o další reprezentační vrstvu, která tento proces provádí a samostatně reprezentuje. Jestliže Harnad vidí identifikaci jako mechanickou proceduru, jejíž výsledky jsou uloženy ve způsobu propojení a vzájemného mapování symbolů a konceptů, v našem případě jsme vytvořili novou reprezentační úroveň. Ta vytváří jednoznačně identifikovatelné kategorie, které jsou nositeli společných vlastností. Jestliže u Harnada jsou jednoznačné kategorie pouze v symbolické rovině, v realizovaném modelu jsou jednoznačné kategorie v rovině multimodální. Proces identifikace se stává samostatnou částí systému, a to se schopností reprezentovat a jednoznačně (absolutně) kategorizovat. Opodstatněnost nového prvku v modelu ukotvení byla prokázána na základě předchozí argumentace, tzn. že Harnadův model není schopen vytvářet jednoznačné kategorie v případě vlivu polysémie či kontextu, jelikož obsahuje fixní symbolickou rovinu. Užití třetí reprezentační vrstvy se nabízí jako logický krok směrem k univerzálnímu reprezentačnímu systému. Proces identifikace je prováděn na základě učení, což zajišťuje jeho schopnost přizpůsobit se vstupním informacím. Jelikož je identifikace realizována pomocí samostatné vrstvy, umožní nám reprezentovat kategorie na základě *privilegovanosti jednoznačného vstupu*, čímž eliminuje vlivy polysémie. V případě rozšíření vlastností primární symbolické vrstvy a způsobu

učení v multimodální vrstvě můžeme zajistit reprezentaci kontextuálních informací v multimodální vrstvě, což sníží výpočetní náročnost spojenou s nutností testování možných variant - stejně jako v případě Harnadova modelu. Další výhodou je možnost napojení ostatních modalit na tuto vrstvu, přičemž identifikace bude vždy probíhat na základě jednoznačného vstupu zajišťujícího tvorbu jednoznačných kategorií.

Podobně jako u Harnadova modelu nás zajímá způsob, jakým můžeme s takto vybudovaným systémem provádět myšlenkové operace. Opět budeme hovořit pouze o elementární formě, která je spojena s logickými principy vyplývání. Již v začátku můžeme odmítnout přístup založený na formální sémantice a také přístup založený na konzistenci a úplnosti výroků, které získaly pravdivost na základě konceptuální úrovně. Jedná se o přístupy používající pravdivost v úrovni *pravda/nepravda*. Jak bylo zmíněno výše, je význam v systému reprezentován pomocí konceptuální roviny, která obsahuje velkou míru nejasnosti a nekonzistence, takže podobný typ posuzování pravdivosti výroků není možný.

V této chvíli je nutné zdůraznit, že právě způsob reprezentace významu pomocí konceptuální úrovně neumožňuje systému naplnit Harnadův předpoklad, tzn. fungování dle pravidel formálního systému a jeho sémantiky. Přestože jsou jevy v externím prostředí na úrovni elementárních vztahů zcela deterministické (jednoznačné fungování fyzikálních zákonů, pokud nebereme v potaz názory kvantové fyziky), není perceptuální aparát podobný lidskému schopen zachytit a reprezentovat tato fakta a jevy jednoznačným a konzistentním způsobem (viz intenzionální vágnost). Pokud chceme s obdobnými reprezentacemi následně pracovat v rovině abstraktních a jednoznačných kategorií, projeví se nekonzistence konceptuálního pokladu (reprezentujícího význam) v rozporných výsledcích inference v úrovni symbolické. Je proto nutné zdůraznit, že nejenom formální sémantika není vhodným nástrojem pro reprezentaci významu v oblasti umělé inteligence, ale příliš hrubým nástrojem reprezentace významu je také sémantika postavená na konceptuálních základech operující s pravdivostí v úrovni *pravda/nepravda*.

Tento fakt zní poněkud neintuitivně v kontextu systému, který je v disertaci vytvořen. Jestliže multimodální vrstva vytváří jednoznačné kategorie na základě *privilegovanosti jednoznačného vstupu*, jedná se o totožnou reprezentační vrstvu, jakou je symbolická v případě Harnadova modelu - s tím rozdílem, že je odolná proti vlivům polysémie.

Přesto se mohou během procesu akvizice symbolických a konceptuálních reprezentací objevit šumy a nekonzistence promítající se do reprezentací v multimodální vrstvě a v rámci následných operací bude na základě inference docházet k rozporným výsledkům. Takové chyby na základě architektury, kterou používáme pro reprezentaci významu, nemůžeme eliminovat, a proto je potřeba hledat alternativní způsob práce v úrovni inference. Jak již bylo zmíněno vý-

še, alternativu může představovat *fuzzy logika*, u které je konzistence vyjádřena mírou pravděpodobnosti. Takový systém je daleko bližší lidskému způsobu operací s reprezentacemi a nabízí možnost eliminovat nekonzistentnost konceptuální úrovně. Nalezení vhodného způsobu implementace zmíněného logického aparátu do stávajícího systému je dalším krokem jeho vývoje, který potvrdí správnost směřování a univerzálnost navrhovaného řešení.

12 Literatura

- Aleman, A., van Lee, L., Mantione, M. H., Verkoijen, I.G., de Haan, E. H. (2001). Visual imagery without visual experience: evidence from congenitally totally blind people. *Neuroreport*, 12(11): 2601-4.
- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11): 832–843.
- Andersen, R.A., Snyder, L.H., Bradley, D.C., Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, 20:303-30.
- Anderson M. L. (2006). Cognitive science and epistemic openness (preprint, published version). *Phenomenology and the Cognitive Sciences*, 5(2):125-154.
- Arditi, A., Holtzman, J. D., Kosslyn, S. M. (1988). Mental imagery and sensory experience in congenital blindness. *Neuropsychologia*, 26:1-12.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon.
- Baddeley, A.D., Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, volume 8, pages 47-89). New York: Academic Press.
- Bailey, D., Feldman, J., Narayanan, S., & Lakoff, G. (1997). Embodied lexical development. *Proceedings of the Nineteenth Annual Meeting of the Cognitive Science Society* (19-25). Mahwah, NJ: Erlbaum.
- Barsalou, L.W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, M.A. Conway (Eds.), *Theories of memories* (29-101). London: Erlbaum.
- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577-609.
- Barsalou, L.W., Kyle, A.K., Wilson, C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7(2):84-91.
- Barsalou, L.W., Prinz, J.J. (1997). Mundane creativity in perceptual symbol systems. In T.B. Ward, S.M. Smith, & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery, and change* (pp. 267-307). Washington, DC: American Psychological Association.
- Barsalou, L.W., Yeh, W., Luka, B.J., Olseth, K.L., Mix, K.S., & Wu, L. (1993). Concepts and meaning. In K. Beals, G. Cooke, D. Kathman, K.E. McCullough, S. Kita, D. Testen (Eds.), *Chicago Linguistics Society 29: Papers from the parasession on conceptual representations* (pp. 23-61). University of Chicago: Chicago Linguistics Society.
- Bavin, E. (1990). Locative terms and Warlpiri acquisition. *Journal of Child Language*, 17:43-66.
- Behrmann, M. (2000). Spatial reference frames and hemispatial neglect. In New cognitive science.
- Bellugi, U., Sabo, H., Vaid, J. (1988). Spatial deficits in children with Williams syndrome. In J. Stiles-Davis, M. Kritchewsky, & U. Bellugi (Eds.), *Spatial Cognition, Brain Bases, and Development* (pp. 273-298). Hillsdale: Lawrence Erlbaum.
- Beni, De R., Cornoldi, C. (1988): Imagery limitations in totally congenitally blind subjects. *J Exp Psychol Learn Mem Cogn*, 14(4):650-5.
- Beni, De R., Moe, A. (2003). Presentation modality effects in studying passages. Are mental images always effective. *Applied Cognitive Psychology*, 17(3):309-324.
- Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*, Cambridge, MA: MIT Press

- Bridgeman, B., Gemmer, A., Forsman, T., Huemer, V. (2000). Processing spatial information in the sensorimotor branch of the visual system. *Vision Research*, 40: 3539-3552.
- Brooks, R.A. (1993). The engineering of physical grounding. Proceedings of the fifteenth annual meeting of the cognitive science society, 153-154. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, J.H., Johnson, M.H., Paterson, S.J., Gilmore, R., Longhi, E., Karmiloff-Smith, A. (2003). Spatial representation and attention in toddlers with Williams syndrome and Down syndrome. *Neuropsychologia*, 1553:1-10.
- Brunswik, E. (1952). *The Conceptual Framework of Psychology*. Chicago: University of Chicago Press.
- Cangelosi A., Parisi D. (2004). The processing of verbs and nouns in neural networks: Insights from synthetic brain imaging. *Brain and Language*, 89(2):401-408.
- Cangelosi, A. (2001). Evolution of communication and language using signals, symbols and words. *IEEE Transactions on Evolutionary Computation*, 5(2):93-101.
- Cangelosi, A. (2002). Grounding language in sensorimotor and cognitive categories. *AISB Quarterly*, 115: 5,8.
- Cangelosi, A. (2006). The Grounding and Sharing of Symbols'. *Pragmatics and Cognition*, 14(2):275-285.
- Cangelosi, A., Greco, A., Harnad, S. (2000). From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories. *Connection Science*, 2:143-162.
- Cangelosi, A., Greco, A., Harnad, S. (2002). Symbol grounding and the symbolic theft hypothesis. In: *Simulating the Evolution of Language*. (A. Cangelosi and D. Parisi, Eds.), London, Springer,,: 191-210.
- Cangelosi, A., Harnad, S. (2001). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language perceptual categories. *Evolution of Communication, special issue on Grounding Language*, 4:117-142.
- Cangelosi, A., Parisi, D., editors (2001). *Simulating the Evolution of Language*. London: Springer Verlag.
- Cappa, S.F., Perani, D., Schnur, T., Tettamanti, M., Fazio, F. (1998). The effects of semantic category and knowledge type on lexical-semantic access: A PET study. *Neuroimage*, 8:350-359.
- Carlson, L.A., West, R., Taylor, H.A., Herndon, R.W. (2002). Neural correlates of spatial term use. *Journal of Experimental Psychology: Human Perception and Performance*, 28:1391-1408.
- Carlson-Radvansky, L.A., Irwin, D.E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, 46:223-244.
- Carlson-Radvansky, L.A., Jiang, Y. (1998). Inhibition accompanies reference-frame selection. *Psychological Science*, 9(5):386-391.
- Carpenter, G., Grossberg, S. (1986). Adaptive resonance theory; stable selforganization of neural recognition codes in response to arbitrary lists of input patterns. In Proceedings of the 8th Annual Conference of the Cognitive Science Society , pp. 45-62
- Carpenter, G. A., Grossberg, S. (2003). Adaptive Resonance Theory, In M.A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*, Second Edition, Cambridge, MA: MIT Press S, 87-90
- Carpenter, G., Grossberg, S. (1990). ART3: Hierarchical search using chemical transmitters in self-organising pattern-recognition architectures. *Neural Networks*, 3(2):129-152.
- Carroll, L. (1960). *The annotated Alice: Alice's adventures in wonderland and Through the looking glass*. New York: Bramhall House.
- Carson-Radvansky, L. A., Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, 7:56-60.

- Carson-Radvansky, L.A., Logan, G.D. (1997). The influence of functional relations on spatial template construction. *Journal of Memory & Language*, 37:411-437.
- Clark, A. (1993): *Associative Engines*. Cambridge, MA : MIT Press.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, H.H., Clark, E.V. (1977). *Psychology and Language*. Harcourt, Brace, Jovanovich, Inc.
- Cohn, A.G. (1996). *Calculi for Qualitative Spatial Reasoning*. In Calmet, J., Campbell, J.A., and Pfalzgraph, J. (Eds.) *Artificial Intelligence and Symbolic Mathematical Computation*. LNCS 1138. Springer Verlag, 124-143.
- Colby, C.L., Goldberg, M.E. (1999). *Space and attention in parietal cortex*. *Annual Review of Neuroscience*, 22:319-49.
- Cole, D (2004).: The Chinese Room Argument, *The Stanford Encyclopedia of Philosophy (Fall 2004 Edition)*, (Edward N. Zalta Ed.), URL = <http://plato.stanford.edu/archives/fall2004/entries/chinese-room/>.
- Comrie, B. (1981). *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press.
- Cooper, G. S. (1968). A semantic analysis of English locative prepositions. *Bolt, Beranek and Newman report*, No. 1587
- Coradeschi, S., Saffioti, A. (2003). An Introduction to the Anchoring Problem. *Robotics and Autonomous Systems*, 43:85-96.
- Cornoldi, C., Bertuccelli, B., Rocchi, P., Sbrana, B. (1993): Processing capacity limitations in pictorial and spatial representations in the totally congenitally blind. *Cortex*, 29(4):675-89.
- Cottrell, G., Bartell B., Haupt, C. (1990). Grounding meaning in perception_ In Proceedings of the German Workshop for AI GWAI Heidelberg Springer
- Coventry, K. (2004). *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions (Essays in Cognitive Psychology)*. Psychology Press.
- Coventry, K.R., Mather, G. (2002). The real story of 'over'?. In K. R. Coventry & P. Olivier (Eds.), *Spatial Language: Cognitive and Computational Aspects*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Crawford, L. E., Regier, T., Huttenlocher, J. (2000). Linguistic and non-linguistic spatial
- Damasio, A. R. (2000). *The Feeling of What Happens: Body, Emotions and the Making of Consciousness*. London: Vintage.
- Damasio, A.R. (1989). Time-locked multiregional retroactivation: A systems level proposal for the neural substrates of recall and recognition. *Cognition*, 33:25-62. Dartmouth College, Hanover, New Hampshire.
- Davidson, D. (1984). 'First-Person Authority', *Dialectica*, 38, reprinted in Davidson, 2001b.
- Davidson, D. (2001). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- Davidsson, P. (1993). Toward a General Solution to the Symbol Grounding Problem: Combining Machine Learning and Computer Vision. In: *AAAI Fall Symposium Series, Machine Learning in Computer Vision: What, Why and How?*:157-161.
- De Jong, E.D. (2000). Autonomous Formation of Concepts and Communication. Ph.D. thesis, Vrije Universiteit Brussel

- Denis, M., Cocude, M. (1992). Structural properties of visual images constructed from poorly or well-structured verbal descriptions. *Memory and Cognition*, 20: 497–506.
- Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dorffner, G., Prem, E. (1993). Connectionism, Symbol Grounding, and Autonomous Agents. In: Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society, 144-148.
- Dorffner, G., Prem, E., Trost, H. (1993). Words, Symbols, and Symbol Grounding, Oesterreichisches Forschungsinstitut fuer Artificial Intelligence, Wien, TR- 93-30.
- Dorffner G., Hentze M., Thurner G. (1996). A Connectionist Model of Categorization and Grounded Word Learning, in *Koster C., Wijnen F. (eds.): Proceedings of the Groningen Assembly on Language Acquisition (GALA '95)*.
- Farkas, I. (2003). Lexical acquisition and developing semantic map. *Neural Network World*, 13(3):235-245.
- Farkas, I., Crocker, M. (2007). Systematicity in sentence processing with a recursive self-organizing neural network. *Proceedings of the 15th European Symposium on Artificial Neural Networks (ESANN'07)*, Bruges, Belgium, (Ed.) M. Verleysen, pp. 49-54,.
- Farkas, I., Li, P. (2002). Modeling the development of lexicon with a growing self-organizing map. In H.J. Caulfield et al. (Eds.), *Proceedings of the 6th Joint Conference on Information Sciences*, Research Triangle Park, NC, pp. 553-556,.
- Fauconnier, G. (1985). *Mental spaces*. Cambridge, MA: MIT Press.
- Fauconnier, G. (1997) *Mappings in thought and language*. New York: Cambridge University Press.
- Fauconnier, G. (1997). *Mappings in Thought and Language*. Cambridge: Cambridge University Press.
- Fauconnier, G., Turner, M. (2002). *The Way We Think*. New York: Basic Books.
- Feldman, A., Fanty, M.A., Goddard, N.H., Lynne, K.J. (1988). Computing with structured connectionist networks. *CACM*, 31(2):170-187.
- Ferenz, K. S. (2000). The role of nongeometric information in spatial language. PhD Thesis,
- Fillmore, Ch. (1982). Frame semantics. In *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea, 111-137. Seoul: Hanshin
- Fodor, J. A. (1975). *The Language of Thought*. New York: Crowell.
- Fodor, J.A., Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28:3-71.
- Frege, G. (1892). "Über Sinn und Bedeutung" in *Zeitschrift für Philosophie und philosophische Kritik* 100: 25-50. Translation: "On Sense and Reference" in Geach and Black (1980).
- Friederici, A.D., Levelt, W.J.M. (1990). Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception & Psychophysics*, 47:253-266.
- Garrod, S., C., Ferrier, G., Campbell, S. (1999). 'In' and 'on': investigating the functional geometry of spatial prepositions. *Cognition*, 72:167-189.
- Gibson, J. J. (1979). *The Ecological Approach to Perception*. Boston: Houghton-Min.
- Glasgow, J.I. (1993). The imagery debate revisited: A computational perspective. *Computational Intelligence*, 9:309-333.

- Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork. *Cognition*, 52:125--157.
- Gopnik, A. (1984). Conceptual and semantic change in scientists and children: Why there are no semantic universals. *Linguistics*, 20:163-179. Also in B. Butterworth and O. Dahl (Eds.) (1984). *Linguistic universals: internal and external explanations*. The Hague: Mouton.
- Granlund, G. (2005). A Cognitive Vision Architecture Integrating Neural Networks with Symbolic Processing. *Künstliche Intelligenz*, 2:18-24.
- Grabowski, J., Miller, G.A. (2000). Factors Affecting the Use of Spatial Prepositions in German and American English: Object Orientation, Social Context, and Prepositional Patterns. *Journal of Psycholinguistic Research*, 29:517-553.
- Greenberg, J.H. (1963/1966). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of grammar*, ed. Joseph H. Greenberg, 2nd edition, 73-113. Cambridge, Mass: MIT Press. (Reprinted in [204], 40-70.)
- Gumperz, J. (1996). *Rethinking Linguistic Relativity*. Edited by John Gumperz & Stephen Levinson. Cambridge University Press.
- Haber, R.N., Haber, L.R., Levin, C.A., Hollyfield, R. (1993): Properties of spatial representations: data from sighted and blind subjects. *Percept Psychophys.*, 54(1):1-13.
- Harnad, S. (1987). (Ed.): *Categorical Perception: the Groundwork of Cognition*. New York: Cambridge University Press, 287-300,
- Harnad, S. (1990). The Symbol Grounding Problem. *Physica D*, 335-346.
- Harnad, S. (2001). Grounding Symbols in the Analog World with Neural Nets – a Hybrid Model. *Psychology*, 12:12-78.
- Harris, R. (1987). *Reading Saussure: A Critical Commentary on the 'Cours de linguistique générale'*. London: Duckworth
- Hawkes, T. (1977). *Structuralism and Semiotics*. London: Routledge.
- Hayward, W. G., Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, 55(1):39-84.
- Heller, M.A., Calcaterra, J.A., Tyler, L.A., Burson, L.L. (1996): Production and interpretation of perspective drawings by blind and sighted people. *Perception*, 25(3):321-34.
- Herskovits, A. (1988). Spatial Expression and the Plasticity of Meaning. In: Rudzka-Ostyn, B. *Topic in Cognitive Linguistics*. John Benjamins Publishing Company, Amsterdam.
- Hiles, D. (2001). *Rethinking Paradigms of Research in Psychology*. Brno: Lecture Series.
- Höfdding H. (1891). *Outlines of psychology*. New York: Macmillan.
- Hume, D. (1739/1978). *A Treatise of Human Nature*. Oxford: Oxford University Press.
- Huttenlocher, J. (1976). Language and intelligence. In L. Resnick (Ed.), *New approaches to intelligence*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Huttenlocher, J. (1973). Language and thought. In G. Miller (Ed.), *Language and communication*. New York: Basic Books.
- Huttenlocher, J. (1973). Language and thought. In G.A. Miller (Ed.), *Communication, language, and meaning: Psychological perspectives* (172-184). New York: Basic Books.

- Huttenlocher, J. (1976). Language and intelligence. In L.B. Resnick (Ed.), *The nature of intelligence* (261-281). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Huttenlocher, J., Hedges, L., Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98:352-376.
- Huttenlocher, J., Smiley, P., Charney, R. (1983). The emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, 90, 72-93.
- Chalmers, D. J. (1992). Subsymbolic Computation and the Chinese Room. In: *The Symbolic and Connectionist Paradigms: Closing the Gap*, (J. Dinsmore, Ed.), Hillsdale: Lawrence Erlbaum, 25-48
- Chandler, D. (1995): *The Act of Writing: A Media Theory Approach*. Aberystwyth: University of Wales.
- Chandler, D., (2007). *Semiotics: the basics, 2nd Edition*. Routledge, London, New York.
- Choi, S., Bowerman, M. (1992). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41:83-121.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: The MIT Press.
- Jackendoff, R. (1987). On beyond zebra: The relation of linguistic and visual information. *Cognition*, 26:89-114.
- Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in english. *Natural Language and Linguistic Theory*, 14:305-354.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17:187-201.
- Johnson, M. (1987). *The body in the mind: The bodily basis of reason and imagination*. Chicago: University of Chicago Press.
- Johnson-Laird, P.N., Miller, G.A. (1976). *Language and Perception*. Belknap Press.
- Johnston, J. R., Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, 6:529-545.
- Jolicoeur, P., Gluck, M.A., Kosslyn, S.M.(1984). Pictures and names: Making the connection. *Cognitive Psychology*, 16: 243-275.
- Just, M.A., Newman, S. D., Keller, T.A., McEleney, A., Carpenter, P.A. (2004). Imagery in sentence comprehension: An fMRI study. *NeuroImage*, 21:112-124.
- Kanizsa, G. (1979). *Organization in vision: Essays in Gestalt perception*. New York: Praeger Press.
- Kant, E. (1781). *Critique of Pure Reason*, Hackett, Indianapolis / Cambridge.
- Kant, I. (2001): *Kritika čistého rozumu*. Praha: Oikomenh.
- Kay, P., McDaniel, C. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 54(3):610-646.
- Klauer, K. C., Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General*, 133:355-381.
- Knauff, M., May, E. (2006): Mental Imagery, Reasoning, and Blindness. *Quarterly Journal of Experimental Psychology. Section A Human Experimental Psychology*, 59:161-177.
- Kohonen, T.(1989). *Self-organization and associative memory*. New York: Springer.

- Kosslyn, S. M. (1994): *Image and Brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
- Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology*, 22:333-347.
- Kosslyn, S. M. (2005): Mental images and the brain. *Cognitive Neuropsychology*, 22:333-347.
- Kosslyn, S. M., Ganis, G., Thompson, W. L. (2003): Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Science*, 7:109-111.
- Kosslyn, S.M.(1980). *Image and Mind*. Cambridge: Harvard University Press.
- Křemen, J. (2007). *Modely a systémy*. Praha: Academia ČMT.
- Kuhn, T.(1996). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Kuipers, B. (2007). An intellectual history of the Spatial Semantic Hierarchy In Margaret Jefferies and Albert (Wai-Kiang) Yeap (Eds.), *Robot and Cognitive Approaches to Spatial Mapping* To appear, Springer Verlag,
- Labov, W. (1973). The boundaries of words and their meanings. In C. J. Baily and R. Shuy, editors, *New ways of analyzing variation in English*, pages 340–373. Washington D.C: Georgetown Univ. Press
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G. (1988). Cognitive semantics. In U. Eco, M. Santambrogio, & P. Violi (Eds.), *Meaning and mental representations* (pp. 119-154).Bloomington, IN: Indiana University Press.
- Lakoff, G., Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.
- Landau, B., Hoffman, J.E. (2005): Parallels between spatial cognition and spatial language: Evidence from Williams syndrome. *Journal of Memory and Language*, 53(2):163-185.
- Landau, B., Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition, *Behavioral and Brain Sciences*, 16: 217-265.
- Langacker, R. (1999). *Foundations of Cognitive Grammar: Volume II: Descriptive Application*. Stanford University Press.
- Langacker, R.W. (1986). An introduction to cognitive grammar. *Cognitive Science*, 10:1-40.
- Langacker, R.W. (1987). *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R.W. (1987). *Foundations of cognitive grammar: volume I.: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R.W. (1991). *Foundations of cognitive grammar, volume II: Descriptive application*. Stanford, CA: Stanford University Press.
- Langacker, R.W. (1997). Constituency, dependency, and conceptual grouping. *Cognitive Linguistics*, 8:1-32.

- Langer, S.K. (1951): *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite and Art*. New York: Mentor.
- Langston, W., Kramer, D.C., Glenberg, A.M. (1998). The representation of space in mental models derived from text. *Memory & Cognition*, 26(2): 247-262.
- Law, D., Miikkulainen, R. (1994). Grounding Robotic Control with Genetic Neural Networks. Tech. Rep. AI94-223. Austin: Dept. of Computer Sciences, The University of Texas
- Lechte, J. (1994): *Fifty Key Contemporary Thinkers: From Structuralism to Postmodernity*. London: Routledge.
- Lem, S. (1964). *Summa technologiae*. Krakow: Wyd. lit.
- Levinson, S.C. (1996). Frames of Reference and Molyneux's Question: Crosslinguistic Evidence. In: *Language and Space*. P. Bloom, et al. Cambridge, MA, MIT Press: 109–170.
- Li, P., Farkas, I., MacWhinney, B. (2004). Early lexical acquisition in a self-organizing neural network. *Neural Networks*, 17(8-9):1345-1362.
- Locke, J. (1690/1979). *An essay concerning human understanding*. P. H. Nidditch, ed. Oxford: Oxford University Press.
- Logie, R.H., Zucco, G.M., Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, 75:55-74.
- Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, 99(4):587-604.
- Marsland, S. (2002). A self-organising network that grows when required. *Neural Networks*, 15:1041–1058.
- Martinetz, T. M., Schulten, K. J. (1991). A neural gas network learns topologies. In T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, editors, *Artificial Neural Networks*, pages 397-402. North-Holland, Amsterdam.
- Materna, P. (1995). *Svět pojmů a logika [The World of Concepts and Logic]*. Praha: Filozofie AV ČR. (second edition in 1998).
- Materna, P. (2000). Svět pojmů a logika. 2. opr. vyd. Praha : Filosofia, 2000. 135 s. Vydal: Filosofický ústav AV ČR. .
- Mayo, M. (2003). "Symbol Grounding and its Implication for Artificial Intelligence", in *Twenty-Sixth Australian Computer Science Conference*, pp. 55-60.
- McCloskey, M. (2001). Spatial representation in mind and brain. In B. Rapp (Ed.), *What deficits reveal about the human mind/brain: A handbook of cognitive neuropsychology* (pp. 101-132). Philadelphia: Psychology Press.
- McCloskey, M., Rapp, B. (2000). Attention-referenced visual representations: Evidence from impaired visual localization. *Journal of Experimental Psychology: Human Perception and Performance*, 26:917-933.
- McNamara, T. P. (1986). Mental representation of spatial relations. *Cognitive Psychology*, 18:87-121.
- McNamara, T.P., Diwakhar, V.A.(1997). Symmetry and assymetry of human spatial memory, *Cognitive Psychology*, 34:160-190.
- Michael, L.A. (2006). Internal supports for external symbol grounding. In *External Symbol Grounding Workshop*, University of Plymouth.
- Michotte, A. (1963). *The Perception of Causality*. New York: Basic Books.
- Millner, A.D., Goodale, M.A. (1995). *The Visual Brain in Action*. Oxford University Press.

- Minsky, M. (1974). *A Framework for Representing Knowledge. A. I. Memo 306*. Cambridge, MA: Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
- Munnich, E., Landau, B., Doshier, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, 81:171-207.
- Munroe, S., Cangelosi, A. (2002). Learning and the evolution of language: the role of cultural variation and learning cost in the Baldwin Effect. *Artificial Life* 8:311-339
- Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Nersessian, N.J. (1992). In the theoretician's laboratory: Thought experimenting as mental modeling. *Proceedings of the Philosophy Association of America*, 2, 291-301.
- Newell, A. (1973). You can't play 20 questions with nature and win. *Visual information processing*, 135–183.
- Newell, A., Simon, H.(1972). *Human Problem Solving*. Prentice Hall.
- Newell, A., Simon, H.A. (1963): GPS: A Program that Simulates Human Thought, In: *Feigenbaum, E.A. and Feldman, J. (eds.): Computers and Thought*, McGraw-Hill, New York.
- Newton, N. (1996). *Foundations of understanding*. Philadelphia: John Benjamins.
- Paivio, A.(1986). *Mental representation: A dual coding approach*. New York: Oxford.
- Peacocke, C. (1992). *A study of concepts*. Cambridge, MA: MIT Press.
- Pecher, D., Zwaan, R.A. (2005). *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge University Press.
- Peirce, C.S.(1958). *Collected Papers of Charles Sanders Peirce*, vols. 1–6, Charles Hartshorne and Paul Weiss (eds.), vols. 7–8, Arthur W. Burks (ed.), Harvard University Press, Cambridge, MA, 1931–1935,.
- Perani, D., Cappa, S.F., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A., Fazio, F. (2003). A fMRI study of word retrieval in aphasia *Brain Lang.*; 85(3):357-68.
- Peregrin, J.: Co je Myšlenka?, In : Kognice 2007, Praha, v tisku
- Perenin, M.T., Vighetto, A. (1988). Optic ataxia: a specific disruption in visuomotor mechanisms. I. Different aspects of the deficit in reaching for objects. *Brain*, 111: 643-674.
- Pfeifer, R., Scheier, C.(1999). *Understanding Intelligence*. MIT Press, Cambridge, USA
- Piaget, J., Inhelder, B.(1967). *The Child's Conception of Space*, New York: Norton.
- Plunkett, K., Sinha, C., Moller, M., Strandsby, O.(1993). Symbol grounding or
- Prinz, J. (2003). "Consciousness, Computation, and Emotion," In S. C. Moore and M. Oaksford (Eds.), *Emotional Cognition: From Brain to Behaviour*. John Benjamins Publishing Company
- Prinz, J.J. (1997). *Perceptual cognition: An essay on the semantics of thought*. Doctoral dissertation. University of Chicago.
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann, & W. Prinz, *Relations between perception and action: Current approaches* (pp. 167–201). Berlin: Springer
- Pulvermüller F., Mohr, B., H.Schleichert (1999). Semantic or lexico-syntactic features: What determines word-class specific activity in the human brain? *Neuroscience Letters*, 275:81-84.
- Pylyshyn, Z.W. (1984). *Computation and Cognition: Towards a Foundation for Cognitive Science*. MIT Press,

- Pylyshyn, Z.W. (2002): Stalking the Elusive Mental Image Screen (reply to commentators). *Behavioral and Brain Sciences*, 25:216-237.
- Pylyshyn, Z.W. (2003) Stalking the Elusive Mental Image Screen (reply to commentators). *Behavioral and Brain Sciences*, 25, (216-237.
- Quillian, M.R. (1967). Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities", *Behavioral Science* 12: 410-430.
- Regier, T. (1992). The Acquisition of Lexical Sémantice for Spatial Terms: A Connectionist Model of Perceptual Categorization. *PhD Thesis / Tech. Rep.* TR-92-062. Berkeley: Dept. of Computer Science, University of California at Berkeley.
- Regier, T.(1996). *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: MIT Press.
- Regier, T., Carlson, L.A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, 130:273–298.
- Regier, T., Zheng M. (2003). An attentional constraint on spatial meaning. In R. Alterman and D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society*.
- Riga, T., Cangelosi, A., Greco, A. (2004). Symbol grounding transfer with hybrid self-organizing/supervised neural networks. In *IJCNN04 International Joint Conference on Neural Networks*. Budapest.
- Richards, I. A. (1932): *The Philosophy of Rhetoric*. London: Oxford University Press.
- Rosenstein, M., Cohen, P. R. (1998). Concepts from time series. *Proceedings of the Fifteenth National Conference on Artificial Intelligence (739-745)*. Menlo Park, CA: AAAI Press.
- Rosch, E. (1978). Principles of categorization. In: Rosch, E., Lloyd, B. (Eds.), *Cognition and Categorization*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 27-48.
- Roy, D. (2002). Learning Words and Syntax for a Visual Description Task. *Computer Speech and Language*, 16(3).
- Roy, D. (2004). Grounding Language in the World: Schema Theory Meets Semiotics. URL: <http://web.media.mit.edu/~dkroy/papers/pdf/aij.pdf>
- Roy, D.(2005). Grounding words in perception and action: computational insights. *Trends in Cognitive Sciences*, 9(8):389-396.
- Roy, D.(2005). Semiotic schemas: a framework for grounding language in action and perception. *Artificial Intelligence*, 167(1-2):170-205.
- Saussure, F. de (1965). *Course In General Linguistics*. McGraw-Hill Humanities/Social Sciences/Languages.
- Saussure, F.de (1916/1983): *Course in General Linguistics*. London: Duckworth
- Seabra, L., Chauhan, A. (2006). One-Class Lifelong Learning Approach to Grounding, *Workshop on External Symbol Grounding. Book of Abstracts and Papers*, 15-23.
- Searle, J. R.(1980). Minds, brains, and programs. *Behavioural and Brain Sciences*, 3:417-457.
- Searle, J.R. (1974). Chomsky's revolution in linguistics. In: Harman, G. (Ed.), *On Noam Chomsky: Critical Essays*. Doubleday, New York. pp. 2–33

- Sharkey, N. E., Jackson, S. A. (1994). Three Horns of the Representational Trilemma. In: *Symbol Processing and Connectionist Models for Artificial Intelligence and Cognitive Modeling: Steps towards Integration*, (V. Honavar and L. Uhr, Eds.), Cambridge, MA: Academic Press, 155-189.
- Sharkey, N.E., Jackson, S.A. (1996). Grounding Computational Engines. *Artificial Intelligence Review*, 10:65-82.
- Schills, K. (2006). Hybrid architecture for the sensorimotor representation of spatial configurations, *Cogn Process*, 7 (Suppl. 1):S90–S92.
- Schyns, P., (1991). A modular neural network model of concept acquisition. *Cognitive Science*, 15: 461-508.
- Siskind, J.M. (2001). Grounding the Lexical Semantics of Verbs in Visual Perception using Force Dynamics and Event Logic. *J. Artif. Intell. Res.*, 15:31-90.
- Sloman, A. (1962). Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth, Oxford University DPhil Thesis, (digitised 2007, Oxford Research Archive)
- Sloman, A. (2002). Getting Meaning off the Ground: Symbol Grounding vs Symbol Attachment/Tethering Invited seminar presentation at MIT Media Lab.
- Slotnick, S. D., Thompson, W. L., Kosslyn, S. M. (2005): Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex*, 15:1570-1583.
- Snyder, L.H., Grieve, K.L., Brotchie, P., Andersen, R.A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, 394:887--891.
- Steels, L. (1999). *The Talking Heads Experiment: Volume I. Words and Meanings. Pre-Edition*. Antwerpen: VUB Artificial Intelligence Laboratory, LABORATORIUM.
- Sun, R. (2000). Symbol Grounding: A New Look at an Old Idea. *Philosophical Psychology*, 13:149-172.
- Sun, R., Zhang, X. (2002). Top-down versus Bottom-up Learning in Skill Acquisition in *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, pp. 63-89
- Tadeo, M., Floridi, L. (2005). The Symbol Grounding Problem: a Critical Review of Fifteen Years of Research. *Journal of Experimental and Theoretical Artificial Intelligence*, 419- 445.
- Takáč, M. (2005). Návrh kognitívnej architektúry pre jazykové experimenty. In: Kelemen, J., Kvasnička, V., Pospíchal, J. (eds.): *Kognice a umělý život V*. Slezská univerzita, Opava, , pp. 549-562.
- Takáč, M. (2007). Konštrukcia významov a jej dynamika v procese iterovaného učenia. In: Kelemen, J., Kvasnička, V., Pospíchal, J. (eds.): *Kognice a umělý život VII*. Slezská univerzita, Opava, pp. 341-347.
- Takáč, M. (2006). Kognitívna sémantika rozlišovacích kritérií [Cognitive semantics of identification criteria]. In: Kelemen, J., Kvasnička, V. (Eds.), *Kognice a umělý život VI*. Slezská univerzita, Opava, pp. 363-372.
- Talmy, L (2003). *Toward a Cognitive Semantics, Volume 1: Concept Structuring Systems; Volume 2: Typology and Process in Concept Structuring*, MIT Press, Language, Speech, and Communication series.
- Talmy, L. (1988). *Force dynamics in language and cognition*. Cogn. Sci.
- Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press,.
- Tarski, A. (1944). The Semantic Conception of Truth: and the Foundations of Semantics'. *Philosophy and Phenomenological Research*, 4(3):341-376.
- Taylor, H.A., Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31: 61-292.

- Thagard, P. (1992). *Conceptual revolutions*. Princeton, NJ: Princeton University Press.
- the emergence of symbols vocabulary growth in children and a connectionist net. *Connection Science*,
- Thomas, N. J. T. (1999): Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, 23:207-245.
- Thomas, N. J. T. (2003): Imagining Minds. *Journal of Consciousness Studies*, 10:79-84.
- Tichý, P. (1996). *O čem mluvíme? Vybrané stati k logice a sémantice*. Praha: Filosofie.
- Touretzky, D. S., Hinton, G. E. (1985). Dymbols Among the Neurons Details of a Connectionist Interference Architecture. In: Proceedings of the 9th International Joint Conference on Artificial Intelligence (IJCAI-85) Los Angeles, Kalifornia
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84 (4):327–352.
- Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, 119: 63-76.
- Ullman, S. (1996). *High-level Vision: Object Recognition and Visual Cognition*. Cambridge, MA: MIT Press.
- Ungerleider, L.G., Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M.A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549-586). Cambridge, MA: MIT Press.
- Vandeloise, C. (1994). Methodology and Analyses of the Preposition. *Cognitive Linguistics*, 5 (2):157-184
- Vanlierde, A., Wanet-Defalque, M.C. (2004): Abilities and strategies of blind and sighted subjects in visuo-spatial imagery. *Acta Psychol*, 116(2):205-22.
- Varela, F., Thompson, E., Rosch, E. (1991). *The Embodied Mind - Cognitive Science and Human Experience*. Cambridge, MA: MIT Press / Bradford Books.
- Vavrečka, M. (2006). Vliv dlouhodobého působení textových stimulů na produkci a prožívání vizuálních mentálních představ. In *Kognice 2006*. 1. vyd. Praha : Psychologický ústav AVČR, 150-158
- Vavrečka, M. (2006a). Ukotvení symbolů v kontextu nulového sémantického závazku. In Kelemen, J., Kvasnička, V., *Kognice a umělý život VI*. Opava : Slezská univerzita, 401-411
- Vavrečka, M. (2007). Ukotvení prostorových vztahů, In J. Kelemen, V. Kvasnička, P. Trebatický: *Kognice a umělý život VII*. Bratislava.
- Vavrečka, M. (2007a). Model reprezentace prostorových vztahů. In J. Lukavský, R. Šíkl (Eds.). *Kognice*. Praha : Psychologický ústav AVČR, v tisku
- Vavrečka, M. (2008). Reprezentace významu pomocí multimodálních reprezentací. *Kognice a umělý život VIII* (VŠE Praha), v tisku
- Vecchi, T. (1998): Visuo-spatial imagery in congenitally totally blind people. *Memory*, 6(1):91-102.
- Voegtlin, T. (2002). Recursive self-organizing maps. *Neural Networks*. 15(8-9):979-991.
- Vogt, P. (2003). Anchoring of semiotic symbols. *Robotics and Autonomous Systems*, 43(2): 109-120.
- Vogt, P. (2003). Grounded lexicon formation without explicit reference transfer: who's talking to who? In W. Banzhaf and T. Christaller and P. Dittrich and J. T. Kim and J. Ziegler (Eds.) *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL)* Springer-Verlag
- Whorf, B. (1956). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. MIT Press
- Wood, F.T. (1979). *English Prepositional Idioms*. London: Macmillan.

Worgan, S.F., Damper, R.I. (2006). *Grounding symbols in the physics of speech communication*. In: External Symbol Grounding Workshop 2006, Plymouth, UK.

Zentall, T.R. (1986). Retrospective coding in pigeons' delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, 12:69-77.

Ziemke, T.(1999). Rethinking Grounding. In: *Understanding Representation in the Cognitive Sciences*, (A. Riegler, M. Peschl and A. von Stein, Eds.), New York: Plenum Press, 177-190.

Zimler, J., Keenan, J.M. (1983): Imagery in the congenitally blind: how visual are visual images? *J Exp Psychol Learn Mem Cogn*, 9(2):269-82.