

**FILOZOFICKÁ FAKULTA UNIVERZITY PALACKÉHO
V OLOMOUCI**



Katedra Obecné lingvistiky

Disertační práce

Sémantické rysy a jejich úloha v pojmové reprezentaci u dětí

STUDIJNÍ OBOR: ***Obecná jazykověda a teorie komunikace***

VYPRACOVALA: **Mgr. Kristýna Konečná**

VEDOUCÍ PRÁCE: **doc. PhDr. Filip Smolík, Ph.D., DSc.**

2021

Prohlašuji, že jsem práci vypracovala samostatně a uvedla všechny použité prameny.

V Olomouci, 10. 11. 2021

podpis

Děkuji především prof. PhDr. Filipu Smolíkovi, PhD., za konzultace, rady a připomínky, které mi během psaní disertační práce poskytl. Děkuji také všem dalším kolegům, kteří mne inspirovali a věnovali mi spoustu cenných rad.

Obsah

1.	Úvod.....	- 6 -
2.	Přehledová část	- 8 -
2.1	Psycholingvistika v českém prostředí a motivace vzniku naší práce	- 8 -
2.2	Sémantické rysy, pojmy a jejich kategorizace	- 9 -
2.2.1	Pojem, slovo, slovní tvar a mimojazyková realita	- 9 -
2.2.2	Pojmy a kategorizace.....	- 11 -
2.3	Rysové reprezentace pojmů – jejich podstata, využití i nedostatky	- 21 -
2.4	Využití rysových reprezentací v rychlých mentálních úvahách o pojmech	- 23 -
2.5	Priming a sémantický priming	- 25 -
2.6	Osvojování pojmů u dětí	- 26 -
3.	Empirická část	- 28 -
3.1	Úvod	- 28 -
3.2	Výběr kategorií a pojmů do databáze sémantických rysů	- 29 -
3.3	Pilotní studie	- 32 -
3.3.1	Cíl pilotní studie.....	- 32 -
3.3.2	Metoda pilotního sběru dat	- 33 -
3.3.3	Statistické zpracování dat a výsledky pilotního výzkumu.....	- 36 -
3.3.4	Vyhodnocení pilotní studie – výběr vhodné metody sběru dat	- 40 -
3.4	Hlavní sběr dat.....	- 40 -
3.4.1	Cíl hlavního sběru dat	- 40 -
3.4.2	Metoda hlavního sběru dat.....	- 41 -
3.4.3	Statistické zpracování dat	- 47 -
3.4.4	Výsledky zpracování databáze a diskuse	- 50 -
3.4.5	Revize databáze sémantických rysů	- 57 -
3.5	Rychlá sémantická úloha: ověření vlivu míry rysové podobnosti na sílu primingového efektu	- 59 -
3.5.1	Cíl experimentu	- 59 -
3.5.2	Metoda experimentu.....	- 60 -
3.5.3	Výsledky	- 65 -
3.5.4	Diskuse	- 68 -
4.	Závěr	- 72 -
5.	Seznam použité literatury.....	- 74 -
6.	Přílohy	- 78 -
6.1	Příloha 1 (Přehled kategorií a pojmů a frekvence jejich výskytu ve vybraných korpusech)	- 78 -
6.2	Příloha 2 (Informační dopis ředitelům vybraných základních škol – hlavní sběr dat).....	- 80 -
6.3	Příloha 3 (Informační dopis a dotazník pro rodiče – hlavní sběr dat)	- 81 -
6.4	Příloha 4 (Informační dopis a dotazník pro rodiče – sémantická úloha).....	- 83 -
6.5	Příloha 5 (Zadání úkolu – pilotní sběr dat).....	- 84 -

6.6	Příloha 6 (Zadání úkolu – hlavní sběr dat)	- 86 -
6.7	Příloha 7 (Zadání úkolu – sémantická úloha)	- 87 -
6.8	Příloha 8 (Příklady popisů pojmů dětmi)	- 88 -
6.9	Příloha 9 (Sémantické rysy uvedené dětmi ve věku 8–10 let pro dva vybrané pojmy)	- 89 -
6.10	Příloha 10 (Srovnání sémantických rysů uvedených dětmi a slovníkové definice)	- 90 -
6.11	Příloha 11 (Srovnání sémantických rysů uvedených českými dětmi a anglickými mluvčími)	- 92 -
6.12	Příloha 12 (Přehled pojmových párů a míry sémantické podobnosti mezi pojmy v páru)	- 95 -
7.	Anotace	- 97 -

1. Úvod

V této práci představujeme databázi sémantických rysů a experiment, který si klade za cíl ověřit využití rysových reprezentací v sémantických úlohách založených na rychlém rozhodování. Databáze byla vytvořena pro vybrané objektové pojmy na základě dat získaných od českých dětí nižšího školního věku. V navazujícím experimentu, který byl prováděn s participanty stejné věkové kategorie, byla využita data z této databáze.

Databáze sémantických rysů byla vytvořena pro 104 substantivních pojmů z 13 různých sémantických kategorií. Děti ve věku 8–10 let měly za úkol zapisovat významy vybraných pojmů a jejich odpovědi byly posléze rozděleny na nejmenší významové jednotky, tzv. sémantické rysy. Systematickým kódováním odpovědí získaných od dětí byl vytvořen přehledný soubor dat, který byl následně zpracován pomocí vhodných statistických metod. Cílem vzniku takové databáze není důsledná charakterizace a jednoznačné definování významů vybraných pojmů pomocí jejich sémantických rysů. Cílem je představit mentální rysové reprezentace pojmů u dětí na základě empirických dat. Tato data byla získána od 381 dětských probandů. V naší práci vycházíme z předpokladu, že takovýto popis pojmů není zdaleka vyčerpávající a konečný a že daná data nepředstavují komplexní mentální reprezentaci vybraných pojmů, ale pouze již zmíněnou rysovou reprezentaci.

Materiál pro naši databázi byl získán od českých dětí. Tím se významně liší od naší vzorové studie, ve které byla data získávána od dospělých anglických mluvčích. Inspirovali jsme se konkrétně především anglosaskými databázemi sémantických rysů, které byly sestaveny K. McRaem a jeho kolegy (studie *Semantic feature production norms for a large set of living and non-living things* z roku 2005) a D. P. Vinsonem ve spolupráci s G. Vigliocco (studie *Semantic feature production norms for a large set of objects and events* z roku 2008).

Databáze sémantických rysů slouží především jako soubor dat, který lze využít pro další psycholingvistické experimenty. Domníváme se však, že by rovněž mohla být inspirativním zdrojem informací pro učitele či rodiče, kteří se chtějí blíže seznámit s dětskou řečí a dětským myšlením (z databáze lze získat určitou představu o osvojování významů slov).

Navazující primingový experiment je inspirován úlohou, která byla představena ve studii *On the Nature and Scope of Featural Representations of Word Meaning* (z roku 1997), jejímiž autory jsou K. Mcrae, M. S. Seidenberg a V. R. de Sa. Experiment vychází z myšlenky, že sémantické rozhodování v úlohách, ve kterých je vyžadována odpověď na dané otázky ve velmi krátkém časovém intervalu, je postaveno především na rysových reprezentacích pojmů, tj. není nutné provádět komplikovanější úvahy a k vyřešení takového typu úlohy jsou využity především základní mentální procesy. Opakem jsou takové úkoly, ve kterých je poskytnut účastníkům dostatek času k přemýšlení a je možné do procesu usuzování zapojit komplexnější mentální procesy a využít hlubší znalosti.

Pojmy z databáze byly spárovány na základě rysové podobnosti, abychom se pokusili ověřit tuto hypotézu, která říká, že v rychlé sémantické úloze míra rysové podobnosti dvou pojmů

ovlivňuje primingový efekt – čím podobnější si jsou rysové reprezentace dvou pojmů (primu a targetu), tím rychlejší je reakce na cílové slovo.

V úvodu teoretické části práce zasazujeme náš výzkum do širšího psycholingvistického kontextu a objasňujeme motivaci ke vzniku této práce. předkládáme přehled teorií o pojmech a jejich kategorizaci, v rámci kterých je sémantický rys zpravidla považován za základní stavební jednotku pojmů. Tomuto tématu se věnuje řada odborníků napříč disciplínami – vedle lingvistiky (sémantiky) je v centru zájmu také ve filozofii a psychologii. Dále se věnujeme postavení rysových reprezentací v kontextu mentálních reprezentací a využití rysových reprezentací v sémantických úlohách, ve kterých je očekávána rychlá reakce participanta. S ohledem na metodu aplikovanou v představovaném experimentu definujeme primingový efekt a sémantický priming. V závěru teoretické části práce uvádíme poznatky z vývojové psychologie a psycholingvistiky související s naším tématem.

2. Přehledová část

2.1 Psycholingvistika v českém prostředí a motivace vzniku naší práce

Psycholingvistika vznikla v 50. letech 20. století – jako samostatný obor je vnímána od roku 1951, kdy se na Cornellově univerzitě uskutečnil seminář, na kterém lingvisté a psychologové formovali své společné cíle. O tři roky později vyšel sborník věnující se psycholingvistickým tématům. V tomto sborníku byl představen teoretický základ této nové vědní disciplíny a byly etablovány metody výzkumu.¹

Kořeny psycholingvistiky lze hledat v teorii informace (tento obor významně ovlivnil psycholingvistiku především v jejích počátcích), ve filozofii, psychologii i v lingvistice, přičemž interdisciplinarita oboru naznačuje, že i v dnešní době se mnohá témata, která psycholingvisty zajímají, překrývají s tématy studovanými v jiných vědních disciplínách. Svědčí o tom i samostatná kapitola *Hranice psycholingvistiky* v Dějinách lingvistiky J. Černého, který uvádí: „Vymezit hranice psycholingvistiky, tj. zjistit co všechno patří do okruhu jejích zájmů, je úkol nesnadný.“² Černý dále uvádí, že témata, kterými se kdy psycholingvisté zabývali, se dotýkají v podstatě všech otázek lingvistických.³

Přesto není v českém prostředí psycholingvistický výzkum příliš rozšířený, ačkoliv by specifika češtiny mohla zásadně přispět při hledání odpovědí na otázky, kterým se psychologie jazyka věnuje. Platí to i pro zkoumání osvojování jazyka, které je v centru zájmu i naší práce: „Ačkoliv je osvojování jazyka předmětem systematického výzkumu již od padesátých let 20. století, stále ještě trpí nedostatkem základních popisných informací. Tím spíš to platí o jazycích, které byly mimo hlavní proud výzkumu, tj. prakticky všech mimo angličtinu. Jakákoli systematická práce dokumentující průběh osvojování češtiny je přínosná.“⁴

Podobně jako F. Smolík popisuje českou psycholingvistiku v o něco mladší publikaci i J. Průcha: „V současné české vědě se neprovádí systematický výzkum řeči, ani neexistuje institucionální základna pro vývojovou psycholingvistiku.“⁵

Ještě téhož roku (2011) ale začala v České republice fungovat LABELS – Laboratoř behaviorálních a experimentálních lingvistických studií, která vznikla jako sdílené pracoviště Filozofické fakulty Univerzity Karlovy a Psychologického ústavu Akademie věd České republiky. Zde se mimo jiné zabývají osvojováním řeči u dětí.⁶ Je zřejmé, že se situace týkající se psycholingvistického výzkumu v českém prostředí zlepšuje a vznikají nové výzkumy a práce, které obohacují nejen tuzemskou, ale i celosvětovou psycholingvistiku o nové poznatky.

Ačkoliv je naše práce z velké části založena na reprodukci anglosaských studií, věříme, že může přispět k rozšíření škály poznatků o jazyku ve vztahu k mysli a vývoji dětské řeči. A to především z toho důvodu, že námi provedený výzkum předkládá data získaná od dětí mladšího školního věku z českého prostředí (v originální studii pracují autoři s dospělými anglickými mluvčími),

¹ Černý, J.: *Dějiny lingvistiky*, 1996. S. 351.

² Tamtéž, s. 353.

³ Tamtéž.

⁴ Smolík, F.: *Psycholingvistika a čeština: některá slibná témata*, 2009. S. 246.

⁵ Průcha, J.: *Dětská řeč a komunikace. Poznatky vývojové psycholingvistiky*, 2011. S. 13.

⁶ Dostupné z: www.labels.ff.cuni.cz [cit. 10. října 2021]

což považujeme za významný přínos do české psycholingvistiky vzhledem k výše zmíněné situaci.

Navíc by náš výzkum mohl mít význam i mimo vědecké bádání – předpokládáme, že kromě možného využití v dalších psycholingvistických experimentech by mohla databáze představovat také užitečný zdroj informací o dětském myšlení a jazyce pro učitele a rodiče.

2.2 Sémantické rysy, pojmy a jejich kategorizace

Předmětem našeho zájmu jsou především sémantické rysy, pojmy a kategorie. V centru se z našeho hlediska nachází právě pojem. Výchozí otázka tedy zní: „Proč disponujeme pojmy, k čemu nám slouží?“ Odpovědi na tuto otázku budeme hledat především ve filozofii a v kognitivní psychologii.⁷ Představíme významné základní teorie týkající se kategorizace pojmů a v souvislosti s kategorizací se budeme zabývat také nejmenšími významovými jednotkami, tedy sémantickými rysy⁸, které v těchto kategorizacích v různé podobě figurují, neboť právě ony jsou považovány za spojovací či rozlišovací prvek mezi jednotlivými pojmy. Vzhledem k primárně lingvistickému zaměření této práce se zmíníme také o rozkladu pojmu (slova) na nejmenší významové jednotky ze sémantického hlediska a objasníme rozdíl mezi tzv. komponentovou analýzou a „kognitivním“ přístupem, který uplatňujeme při tvorbě databáze my.

2.2.1 Pojem, slovo, slovní tvar a mimojazyková realita

Pokud budeme operovat termínem pojem (koncept) v lingvistických kruzích, velice pravděpodobně se zastavíme u F. de Saussura, Ch. S. Peirce, či u Ch. K. Ogdena a I. A. Richardse, kteří se věnovali teorii jazykového znaku, jehož součástí je právě pojem.

V pojetí F. de Saussura se jazykový znak skládá ze dvou částí, které jsou psychické povahy a které tvoří jeden celek „v našem mozku asociativním svazkem sjednocovaný“, jak uvádí sám F. de Saussure⁹. První část takového znaku se nazývá „signifiant“; v českém překladu se běžně uvádí termín „označující“. Jde o tzv. akustický obraz (tj. o zvukovou stránku, resp. jedná se o psychický zvukový otisk v mysli a nemusí nutně dojít k vyslovení nahlas; F. de Saussure zde mluví také o „reprezentaci“¹⁰). Druhá část znaku, „signifié“, se týká „označovaného“ – tedy představy, obrazu dané věci v naší mysli, která je ještě abstraktnější než výše zmíněný akustický obraz (jde tedy v podstatě o koncept; v našem případě pak můžeme použít český termín pojem).¹¹

⁷ Ačkoliv je tato práce prezentována primárně v lingvistickém prostředí, má jednoznačně mezioborový charakter, a proto věříme, že výsledky našeho bádání budou uplatnitelné nejenom v jazykovědných, ale i v dalších oborech, např. v psychologii, filozofii či pedagogice.

⁸ V naší práci používáme vedle termínu „rys“ také synonymní termín „atribut“, případně „vlastnost“.

⁹ Saussure, F. de: *Kurs obecné lingvistiky*, 1996. S. 95.

¹⁰ Tamtéž, s. 96.

¹¹ Tamtéž, s. 95–97. (Srov. také např. Černý, J.: *Dějiny lingvistiky*, 1996. S. 140–141.)

V případě idejí F. de Saussura není pro jazykový znak podstatná vedle zvukové realizace ani mimojazyková realita.¹² Avšak u jiných autorů se ve vztahu k těmto dvěma částem se zpravidla mluví také o nejazykové části znaku, kterou je reálný předmět ve skutečném světě:

„Proces pojmenovávání se netýká pouze dvou složek (slovo a věc), ale má tři základní složky: slovní tvar se pojí s určitým pojmem a spolu – jako jazykový znak o dvou nerozlučných stránkách – se vztahují k příslušné části světa, který nás obklopuje.“¹³

Tato mimojazyková realita je brána v potaz v triadických znakových modelech, např. v případě dalšího významného vědce věnujícího se jazykovému znaku, Ch. S. Peirce. Podobně ji do svého tzv. sémiotického trojúhelníku zařazuje také Ch. K. Ogden a I. A. Richards.¹⁴

Z této triády, kterou pro naše potřeby můžeme zjednodušeně pojmenovat pomocí výrazů slovní tvar – pojem – věc, se v naší práci zaměřujeme především právě na pojem, který je uložen v naší mysli. Považujeme za vhodné už nyní zmínit, že pojmy, které jsme vybrali do naší databáze a našeho experimentu, nejsou abstraktní. Je možné je vyjádřit pomocí slov, nebo dokonce ukázat na věci, ke kterým odkazují ve světě kolem nás. Rovněž je nezbytné uvést, že v naší práci jsou pojmy, které nás zajímají, participantům předkládány v písemné podobě. Jedná se o pojmy/slova z oblasti základních sémantických okruhů (námi zkoumané kategorie jsou zvířata, rostliny, nábytek, dopravní prostředky apod.) a vždy jsou to substantiva.

V naší práci tedy mluvíme především o pojmech, ale také o slovech – respektive o slovních tvarech, které zastupují jednotlivé pojmy, přičemž těmito slovními tvary odkazujeme k daným pojům. Pojmy jsou však zároveň úzce svázány s reálnými objekty (viz k tomu níže o subjektivnosti a objektivnosti pojmů v případě dětského myšlení) a je těžké vytyčit hranici jasně určující, kdy mluvíme jen o pojmu a kdy o skutečné věci nacházející se ve světě fyzicky. G. L. Murphy v souvislosti s tím mluví o problematice terminologické, kdy se autoři děl pojednávajících o pojmech snaží reálné objekty a pojmy odlišit, avšak ne vždy je to snadné, či dokonce možné¹⁵: „V mnoha případech je zřejmé, zda se mluví o pojmu nebo o věci (...). Jindy zase může být jejich nejednoznačné rozlišení záměrné. (...) Představte si například, že se naučíte, že psi štěkají. Naučili jste se, že psi štěkají, nebo že pojem PES¹⁶ má vlastnost „štěká“? V určitém smyslu jste se dozvěděli něco nového o tom, jak funguje skutečný svět, a naučili jste se tedy něco o skutečných psech. Z jiného úhlu pohledu jste na základě tohoto nového zjištění

¹² **Karlík, P., Nekula, M. a Pleskalová, J. (eds.):** *Nový encyklopedický slovník češtiny* [online], 2012–2020, heslo *Znak*. Dostupné z: www.czechency.org/slovník/ZNAK#bibitem8 [cit. 10. září 2021]

¹³ **Černý, J.:** *Úvod do studia jazyka*, 1998. S. 148.

¹⁴ **Černý, J. a Holeš, J.:** *Sémiotika*, 2004. S. 45–46.

¹⁵ Záměrně jsme zde vynechali tu část textu, která se v souvislosti s tématem zabývá také otázkou typografickou, protože nás zde zajímá primárně samotná otázka (ne)možnosti odlišení pojmů od věcí vyskytujících se v reálném světě.

¹⁶ V naší práci respektujeme v rámci citací typografické označení pojmů či kategorií tak, jak jej uvádějí citovaní autoři, přičemž věříme, že toto rozlišení bude dostatečně zřejmé z kontextu. V ostatních částech práce kategorie vyznačujeme tučným písmem v uvozovkách, pojmy tučným písmem, sémantické rysy kurzívou a reálné odpovědi probandů (příp. úseky jejich odpovědí) kurzívou v uvozovkách – domníváme se, že v praktické části, ve které popisujeme vlastní experiment, je takové značení poměrně jednoznačné; problematické může být spíše v teoretické části či při interpretacích a diskusích.

aktualizovali informace o pojmu PES a upravili si jeho mentální reprezentaci. S největší pravděpodobností jste se však naučili obojí.“¹⁷

Je důležité podotknout, že pojmy nemusí nezbytně zastupovat pouze věci, které se v našem světě skutečně vyskytují: „Máme tendence přemýšlet o kategoriích jako o něčem, co existuje ve skutečném světě, a o pojmech jako o jejich mentálních reprezentacích, ale taková analýza není správná. Je zavádějící, protože pojmy nepotřebují mít skutečné, ve světě existující protějšky (např. jednorozec) a protože lidé mohou spíše utvářet než objevovat strukturu světa.“¹⁸

Domníváme se, že takový pohled na věc koresponduje s významnou Lakoffovou tezí, ve které se mluví o tzv. pojmovém vtělesnění. Lakoff vysvětluje, že pojmové vtělesnění je „Myšlenka, že vlastnosti jistých kategorií jsou ovlivněny biologickými předpoklady člověka a jeho zkušeností z existence ve fyzickém a společenském prostředí. Toto se staví do protikladu k představě, že pojmy existují nezávisle na tělesné povaze myslících bytostí a jejich zkušenosti.“¹⁹

2.2.2 Pojmy a kategorizace

Jednotlivé objekty vyskytující se v našem světě nejsou zcela jedinečné v tom smyslu, že je (zpravidla) možné najít další, jim podobné či příbuzné předměty. Proto je možné v našich myslích utvářet určitá zobecnění v podobě pojmů, které je dále možné sdružovat do kategorií, což se děje obvykle právě na základě určité podobnosti či příbuznosti. Přestože je náš pes Alík zcela originální a je to pro nás naprosto jedinečná bytost, stále jej můžeme zařadit do kategorií jako je „zvíře“, „pes“ či „domácí mazlíček“. „Pro naše myšlení, vnímání, jednání a řeč neexistuje nic centrálnějšího, než je kategorizace. Kategorizujeme vždy, když nějakou skutečnost vnímáme jako *druh* něčeho, například strom. Kategorie používáme pokaždé, když uvažujeme o *typech* věcí, jako jsou židle, národy, nemoci, emoce, prostě cokoli. Pokaždé když záměrně provádíme nějakou aktivitu, třeba něco tak běžného jako psaní tužkou, zatloukání kladivem nebo žehlení, používáme kategorie.“²⁰

I v tomto případě však G. L. Murphy poukazuje na možný terminologický chaos, jelikož je mnohdy těžké určit hranici mezi pojmem a kategorií – například **pes/„pes“** může zastávat funkci jak pojmu, tak kategorie, do které můžeme zařadit všechny individuální psy, kteří se na

¹⁷ **Murphy, G. L.:** *The big book of concepts*, 2002. S. 6. (překlad autorky)

„In many cases, it is simply obvious whether one is talking about the concept or the object (...). In other cases, though, the ambiguity between the concept and the thing is intentional. (...) For example, suppose you learn that dogs bark. Have you learned that dogs bark or that DOG (the concept) has the property „barks“? From one perspective, you have learned about a property in the world, and so you have learned something about actual, lower-case dogs. From another perspective, you have changed your concept, and so you have modified your mental representation, that is, DOG. In fact, you have learned probably both.“

¹⁸ **Medin, D. L.:** *Concepts and conceptual structure*, 1989. S. 1469. (překlad autorky)

„It is tempting to think of categories as existing in the world and of concepts as corresponding to mental representations of them, but this analysis is misleading. It is misleading because concepts need not have real world counterparts (e.g. unicorns) and because people may impose rather than discover structure in the world.“

¹⁹ **Lakoff, G.:** *Ženy, oheň a nebezpečné věci*, 2006. S. 25.

²⁰ Tamtéž, s. 19.

světě vyskytují: „(...) ať už disponuji jakýmkoliv pojmem, vždy existuje skupina/kategorie věcí, která by tímto pojmem mohla být označena.“²¹

Otázkou, proč disponují lidské bytosti schopností utvářet pojmy a slučovat je do kategorií, se zabývá velké množství filozofů či psychologů. Ať už demonstrují své teorie na příkladech, nebo je popisují v obecnějším smyslu, většina se v odpovědích na tuto zásadní otázku shoduje – ve společenském smyslu je podstatné, že bez schopnosti utvářet si pojmy by spolu lidé nemohli komunikovat; na individuální úrovni jde o určitou redukci informací a proces úspornosti, abychom lépe zvládali v každodenním životě nakládat se světem kolem nás, protože naše mentální kapacita není nekonečná:

„Potřebujeme disponovat znalostmi o věcech kolem nás, abychom s nimi mohli zacházet a účelně jednat v prostředí, které nás obklopuje. V nejobecnějším smyslu představuje naše vědění veškeré informace, které jsme zdělili nebo jsme se je naučili. (...) Proč si ale potřebujeme znalosti organizovat? Nestačí pouze zkušenost získat a uchovat. Zkušenost potřebujeme také uspořádat tak, aby byla úsporná a informativní. (...) Vytváříme si kategorie objektů, jako jsou psi, z části kvůli tomu, abychom se vyhnuli nutnosti pamatovat si každého jednotlivého psa, kterého jsme kdy viděli (...). Náš paměťový systém vyžaduje jistou kognitivní úspornost („cognitive economy“) v organizování zkušenosti. (...) musíme abstrahovat od konkrétní zkušenosti a vyvíjet obecné pojmy.“²²

„Pojmy jsou vazivo, které udržuje náš mentální svět pohromadě. Ať už vstupujeme do pokoje, zkusíme novou restauraci, jdeme do obchodu nakupovat potraviny či k lékaři, čteme nějaký příběh – ve všech takových situacích se musíme spoléhat na pojmy, které nám pomáhají rozumět tomu, co se děje ve světě kolem nás. (...) Ačkoliv jsem nikdy předtím neviděl rajče zcela identické tomu, které teď držím v ruce, můžu předpokládat, že i toto rajče je stejné jako všechna předešlá rajčata, a tudíž je jedlé. Jestliže si vytvoříme pojem (mentální reprezentaci) odpovídající dané kategorii (tj. třídě objektů nacházejících se ve světě), tento pojem nám pomáhá chápat novou entitu náležící této kategorii a adekvátně na ni reagovat. Pojmy jsou typ mentálního lepidla, jsou pojítkem mezi naší předchozí zkušeností a aktuální interakcí se světem a zároveň jsou součástí naší širší vědomostní struktury.“²³

Specifičtější příklad toho, proč je důležité umět „zevšeobecňovat“ a nechápat striktně individuálně jednotlivé věci či situace vyskytující se kolem nás, uvádí D. L. Medin: „K čemu jsou dobré kategorie? Kategorizace umožňuje zacházet se dvěma či více odlišnými entitami jako se sobě rovnými, takže je možné čerpat z dříve nabytých vědomostí a vytvářet predikce.

²¹ **Murphy, G. L.:** *The big book of concepts*, 2002. S. 5. (překlad autorky)

„(...) whatever my concept is, there is a category of things that would be described by it.“

²² **Eysenck, M. W. a Keane, M. T.:** *Kognitivní psychologie*, 2008. S. 342–343.

²³ **Murphy, G.:** *The big book of concepts*. 2002. S. 1. (překlad autorky)

„Concepts are the glue that holds our mental world together. When we walk into a room, try a new restaurant, go to the market to buy groceries, meet a doctor, or read a story, we must rely on our concepts of the world to help us understand what is happening. (...) Although I've never seen this particular tomato before, it is probably like other tomatoes I have eaten and so is edible. If we have formed a concept (a mental representation) corresponding to that category (the class of objects in the world), then the concept will help us understand and respond appropriately to a new entity in that category. Concepts are a kind of mental glue, then, in that they tie our past experiences to our present interactions with the world, and because the concepts themselves are connected to our larger knowledge structures.“

Vezměme si jako příklad psychodiagnostické kategorie. Je nutné mít přístup k relevantním znalostem – to vysvětluje, proč kliničtí psychologové nepřístupují (respektive, neměli by přistupovat) ke každému klientovi individuálně. (...) Kliničtí psychologové musejí nějakým způsobem přenést své znalosti a zkušenosti na aktuálně řešený problém; v takovém případě jistě ocení, že se nová situace v mnohém podobá jiné, se kterou se už dříve setkali. Přestože kliničtí psychologové můžou (ale nemusí) využít konkrétní systém kategorizace, vždy musí najít výchozí body spojující předchozí a aktuální situaci; to znamená, že musejí kategorizovat. Diagnostické kategorie pomáhají klinickým psychologům předvídat efektivnost jednotlivých typů léčby a pomáhají jim sdílet zkušenosti s dalšími terapeuti.²⁴

Zdá se, že otázku, proč lidé ve své mysli umějí vytvářet pojmy a kategorizovat je, není tak těžké zodpovědět. Avšak porozumět tomu, jakým způsobem pojmy (respektive svět) kategorizujeme, je podle všeho mnohem komplikovanější. Touto problematikou se učenci a badatelé zabývali už od dob Aristotela. Budeme však přesnější, když řekneme, že přes dvě tisíciletí panovalo přesvědčení o tom, že pojmy sdružujeme do kategorií jednoduše na základě společných rysů a všechny tyto pojmy jsou stejně kvalitními členy dané kategorie:

„Od dob Aristotela až po pozdní práce Wittgensteinovy byly kategorie považovány za neproblematické, za něco, čemu dobře rozumíme. Byly považovány za abstraktní nádoby a věci byly buď v nich, nebo mimo ně. Věci byly řazeny do stejné kategorie pouze tehdy, pokud měly určité vlastnosti společné. A vlastnosti, které měly společné, byly považovány za vlastnosti definující tuto kategorii.“²⁵

Proces kategorizování byl po dlouhou dobu poměrně přirozeně spojován s definicemi: cílem bylo najít pro pojmy (i abstraktního charakteru) definici, která dostatečně vystihne daný pojem a zároveň jej odlišuje od pojmů jiných. Pokud byl posléze objeven nesoulad mezi daným pojmem a jeho definicí (tj. skutečnost ve světě odpovídala pojmu, ale vymykala se definici), musela být tato definice přehodnocena.²⁶

G. L. Murphy zároveň zdůrazňuje, že „(...) ne všichni tito autoři explicitně zmiňovali, že užívají definiční teorii pojmů. Spíše tento přístup považovali za samozřejmý a zaměřovali se na to, jak se lidé učí nové pojmy (tj. jak se učí tyto definice) na základě zkušeností.“²⁷

²⁴ **Medin, D. L.:** *Concepts and conceptual structure*, 1989. S. 1469. (překlad autorky)

„What good are categories? Categorization involves treating two or more distinct entities as in some way equivalent in the service of accessing knowledge and making predictions. Take psychodiagnostic categories as an example. The need to access relevant knowledge explains why clinical psychologists do not (or could not) treat each individual as unique. (...) Clinicians need some way to bring their knowledge and experience to bear on the problem under consideration, and that requires the appreciation of some similarity or relationship between the current situation and what has gone before. Although clinical psychologists may or may not use a specific categorization system, they must find points of contact between previous situations and the current context; that is, they must categorize. Diagnostic categories allow clinicians to predict the efficacy of alternative treatments and to share their experience with other therapists.“

²⁵ **Lakoff, G.:** *Ženy, oheň a nebezpečné věci*, 2006. S. 20.

²⁶ **Murphy, G. L.:** *The big book of concepts*, 2002. S. 11–12.

²⁷ Tamtéž, s. 12. (překlad autorky)

„(...) these writers did not always explicitly say, “I have a definitional theory of concepts.” Rather, they took such an approach for granted and then went about making proposals for how people learned concepts (i.e., learned these definitions) from experience.“

Dnes se pro tento dlouhodobě implicitně uznávaný přístup ke kategorizaci pojmů zpravidla používá název „klasická teorie“ (toto pojmenování zavedli Smith a Medin²⁸ v roce 1981). O tom, že vlastně nešlo o žádnou mysliteli předloženou teorii, ale v podstatě o dogmatickou myšlenku, píše Lakoff: „Klasická teorie nebyla výsledkem empirických studií. Nebyla dokonce ani předmětem větší debaty. Byla to filozofická pozice, k níž se dospělo na základě apriorních spekulací. Během staletí se jednoduše stala součástí nevyřčených předpokladů, považovaných ve většině vědeckých disciplín za dané. (...) Byla vyučována ve většině disciplín ne jako empirická hypotéza, ale jako nezpochybnitelná pravda, která slouží k definování dalších pravd. Během podivuhodně krátkého období se toto všechno změnilo. Kategorizace se přesunula z pozadí na výsluní díky empirickým studiím z celé řady oborů.“²⁹

Klasická teorie je tedy postavena na několika přesvědčeních: význam kategorie je v mysli uložen ve formě definice vystihující všechny členy, které jsou si sobě rovné; neexistují případy, ve kterých se pojmy pohybují na hranici mezi více kategoriemi; každý člen kategorie musí obsahovat atributy, které jsou jednotlivě nezbytné a společně dostačující na to, aby člen mající tyto charakteristiky mohl patřit do dané kategorie.

Už Wittgenstein v rámci svých filozofických úvah však došel k závěru, že taková představa kategorizace není aplikovatelná na všechny typy kategorií. Jako vzorový příklad problematické kategorizace se i dnes uvádí kategorie her: „Míním deskové hry, karetní hry, míčové hry, bojové hry atd. Co ty všechny mají společného? – Neříkej: „Něco společného mít *musejí*, jinak by se jim neříkalo ‚hry‘“ – nýbrž *podívej se*, jestli je tu něco, co je společné jim všem. – Neboť když se na ně podíváš, neuvidíš sice něco, co by bylo *všem* společné, ale uvidíš všelijaké podobnosti, příbuznosti, a sice řadu takových podobností a příbuzností. (...) Nemohu tyto podobnosti charakterizovat lépe než slovem „rodové podobnosti“; neboť takto se překrývají a kříží různé podobnosti vyskytující se u členů nějaké rodiny (...).“³⁰

Nicméně G. L. Murphy (s odkazem na M. Ramscara, 1997³¹) poukazuje na to, že Wittgensteinův přístup k problematice pojmů byl především filozofický, a nelze jej proto řadit mezi přímé předchůdce moderních psychologických teorií o kategorizaci pojmů. Wittgenstein se domníval, že obecně není možné analyzovat pojmy a popsat je za pomoci jejich dílčích atributů. Tudiž jeho ideje sice roztříštily předchozí představy o možnosti definování pojmů, zároveň ale ani neumožňovaly přistupovat k analýze pojmů a jejich kategorizaci ani tak, jak to ve své teorii představily E. Roschová a C. B. Mervisová, které dále pracovaly s Wittgensteinovým termínem „family resemblance“ (tj. rodinná podobnost/příbuznost) a přitom na tohoto filozofa explicitně odkazovaly^{32, 33}

Postupně byly předkládány další a další příklady, které byly v rozporu s klasickým přístupem ke kategorizaci. Vezměme si jako příklad kategorii „psi“: můžeme s jistotou říci, že ne všichni psi *mají ocas* či *srst*, ale také to, že není pravda, že všichni psi *štěkají*. Podle klasické teorie

²⁸ Smith, E. E. a Medin, D. L.: *Categories and concepts*, 1981. S. 1.

²⁹ Lakoff, G. L.: *Ženy, oheň a nebezpečné věci*, 2006. S. 20.

³⁰ Wittgenstein, L.: *Filozofická zkoumání*, 1993. S. 45–46.

³¹ Ramscar, M.: *Wittgenstein and the representation of psychological categories*, 1997.

³² Rosch, E. a Mervis, C. B.: *Family Resemblances: Studies in the Internal Structure of Categories*, 1975. S. 574–575.

³³ Murphy, G. L.: *The Big Book of Concepts*, 2002. S. 499–500.

nemohou tedy být tyto rysy součástí definice, protože výše uvedené vlastnosti nemají všichni členové kategorie. Přesto podvědomě tušíme, že *mit srst, ocas* a především pak *štěkot* patří mezi dobré znaky, pomocí kterých můžeme poznat zvíře, které patří do kategorie „psi“, a odlišit jej od jiných živočišných druhů).³⁴

Předmětem dalších diskusí byla intuitivní představa, že ne všichni zástupci kategorie jsou si rovni a že někteří zástupci jsou nějakým způsobem jednoznačně lepší než jiní (mezi „ptáky“ se řadí **kos**, ale i **tučňák**; zároveň většina z nás jistě považuje z těchto dvou právě **kosa** za lepšího zástupce ptačí říše). Původní myšlenka, že typičtějšími představiteli kategorie jsou ti, kteří jsou více frekvencovaní, byla postupně zavrhnuta. Uvedme opět příklad z kategorie „ptáci“. **Slepice** je zvíře, které se hojně vyskytuje na naší planetě, a přesto je zřejmé, že většina z nás ji nepovažuje za tradičního zástupce kategorie „ptáci“.³⁵

Jak už bylo naznačeno výše, zásadní vliv na změnu chápání kategorizace, a to především v oblasti psychologie, je přisuzován E. Roschové a C. B. Mervisové, které v 70. letech představily přelomovou studii, ve které empiricky (tj. na základě několika experimentů) dokazovaly nesprávnost klasického přístupu, oživily Wittgensteinův pojem rodinné podobnosti a předložily důkazy o prototypických jevech.

Roschová a Mervisová pomocí svých experimentů doložily, že někteří členové kategorie jsou typičtější než ostatní (v experimentu byly analyzovány základní sémantické kategorie jako jsou „nábytek“, „oblečení“ či „ovoce“) a že existuje korelace mezi typičností a počtem sdílených rysů: typičtější zástupci mají více společných rysů s dalšími (ne nezbytně všemi) zástupci kategorie než ti, kteří jsou považováni za méně typické.³⁶ Takové zjištění je v souladu s hypotézou o rodinné podobnosti, kterou Roschová s odkazem na Wittgensteina vysvětluje takto: „Rodinná příbuznost je dána vztahem, který může být definován jako AB, BC, CD, DE. To znamená, že každá položka (člen kategorie) má přinejmenším jeden, avšak spíše několik prvků (rysů) společných s jedním nebo více členy kategorie, ale žádný, nebo jen několik prvků je společných všem členům kategorie.“³⁷

Zároveň je nutné poznamenat, že ačkoliv se zpravidla odkazuje na první práce Roschové, ona sama v průběhu let změnila přístup k interpretaci vlastních experimentů, na což poukazuje například G. Lakoff: „Není pochyb o tom, že takové prototypové jevy opravdu existují. Interpretace těchto výsledků se však stala předmětem sporů a nepochopení. (...) Ale prototypové jevy samy o sobě nekonstituují žádnou konkrétní alternativní teorii mentální reprezentace. Jako zodpovědná experimentátorka Roschová vždycky důsledně rozlišovala mezi výsledky pokusů a teoriemi, kterými by je bylo možné vysvětlit.“³⁸

³⁴ Volně podloženo příkladem uvedeným v **Murphy, G. L.**: *The big Book of Concepts*, 2002. S. 30–31.

³⁵ Volně podloženo příkladem uvedeným v **Murphy, G. L.**: *The big Book of Concepts*, 2002. S. 31–32.

³⁶ **Rosch, E. a Mervis, C. B.**: *Family resemblances: Studies in the Internal Structure of Categories*, 1975.

³⁷ Tamtéž, s. 575. (překlad autorky)

„A family resemblance relationship consists of a set of items of the form AB, BC, CD, DE. That is, each item has at least one, and probably several, elements in common with one or more other items, but no, or few, elements are common to all items.“

³⁸ **Lakoff, G.**: *Ženy, oheň a nebezpečné věci*, 2006. S. 53.

Navzdory výše zmíněnému vývoji postoje samotné autorky k výsledkům této práce se dnes v publikacích věnujících se kategorizaci pojmů uvádí teorie prototypů jako nástupce klasické teorie: podle této teorie má každá kategorie svého prototypického člena, zástupce, který se vyznačuje tím, že obsahuje všechny charakteristické (nikoliv definující) vlastnosti. Podle některých interpretací může jít o nejtypičtějšího reálně existujícího člena kategorie, ale také o ideálního zástupce, který ve skutečném světě neexistuje – je pouze souhrnem všech charakteristických vlastností. Členové kategorie se pak s tímto prototypickým představitelem shodují v různé míře (sdílejí či nesdílejí charakteristické rysy). Z toho vyplývá, že všichni členové si nejsou rovni, a tudíž je možné je seřadit podle míry prototypičnosti (**kos** je prototypičtější zástupce kategorie „**ptáci**“, než je například **slepice**, které mimo jiné chybí ptákům typická vlastnost *létání*). Navíc neexistuje jasná hranice, kde končí jedna kategorie a kde začíná druhá (lidé se často dohadují, i v českém prostředí³⁹, zda je **rajče** „**ovoce**“ či „**zelenina**“, podobně je tomu v případě **melounu**).⁴⁰

Další teorie, známá jako teorie exemplářů, přináší méně schematické a více praktické řešení: v paměti si uchováváme zástupce kategorie, se kterými se ve svém životě setkáváme. Není tedy potřeba vytvářet mentální prototyp, který shrnuje všechny charakteristické rysy pro danou kategorii (a který ve skutečném světě ani nemusí existovat). To, že považujeme některé zástupce kategorií za typičtější než jiné, je dáno naší zkušeností s jednotlivými zástupci kategorie: „Za předpokladu, že jsme v minulosti potkali mnoho drozdů, je pravděpodobné, že v naší paměti bude uchováno mnohem víc zástupců drozdů než tučňáků. Proto bude nějaký zástupce drozda vyvolán v paměti mnohem rychleji než nějaký zástupce tučňáka (...).“⁴¹ Takový přístup nabízí určitou flexibilitu při kategorizaci pojmů – jelikož zde nedochází k posuzování členství v kategorii na základě jednoho nejlepšího zástupce, můžou být v paměti vyvolávání různí zástupci podle aktuální potřeby.

Zdá se, že za pomoci teorie exemplářů je možné vysvětlit velké množství příkladů, které byly uváděny jako doklady teorie prototypů, a někdy je dokonce možné pomocí tohoto přístupu objasnit jevy, pro které teorie prototypů vysvětlení nenachází⁴². Na druhou stranu lze na základě mnohých experimentů doložit jevy, které dokládají existenci prototypů a v rámci teorie exemplářů pro ně nelze najít odpověď.⁴³ Kromě specifických experimentů, které dokládají, či vyvracejí teorii exemplářů, lze sledovat nedostatky této teorie i v případě některých základních otázek: „Pokud se například snažíme odpovědět na otázku, zda je tvrzení „Všichni ptáci jsou tvorové“ pravdivé, zdá se, že spoléháme spíš na naši obecnou znalost než na specifické příklady. Teorie exemplářů neposkytuje žádné uspokojivé vysvětlení toho, jak si takové obecné znalosti konstruujeme, protože bere v potaz pouze úroveň jednotlivých konkrétních příkladů.“⁴⁴

³⁹ Zde narážíme na možnou rozdílnost kategorizace u různých národů; k tomuto tématu se ještě vrátíme.

⁴⁰ Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008. S. 354–355.

⁴¹ Tamtéž, s. 358–359.

⁴² Viz například doklady teorie exemplářů v Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008. S. 358–359.

⁴³ Viz například výhody prototypů v Murphy, G. L.: *Big book of concepts*, 2002. S. 54–58.

⁴⁴ Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008. S. 359–360.

Ani jedna z těchto teorií ovšem nedokázala vysvětlit další typy kategorizace. Jedná se například o tzv. goal-derived kategorie (tzn. kategorie, které jsou vytvářeny podle účelu) a ad hoc kategorie, které uvádí Barsalou: „Existuje nekonečné množství goal-derived kategorií, například *jídlo, které jíme, když máme dietu, oblečení, které nosíme, když malujeme pokoje, potravinový obchod, ve kterém se prodávají čerstvé byliny (...)* atd. Velké množství těchto kategorií vzniká ad hoc a nejsou pevně uloženy v paměti, ale jsou utvářeny improvizovaně za účelem dosažení aktuálního, nového cíle.“⁴⁵ Barsalou tyto kategorie srovnává se standardními taxonomickými kategoriemi, o kterých jsme mluvili výše (např. kategorie „**ptáci**“), a poukazuje na jejich odlišné struktury: „Jestliže tyto dva typy kategorií mají odlišný původ, jejich kognitivní struktura se může lišit. V mnoha předchozích studiích bylo řečeno, že běžné taxonomické kategorie mají prototypickou strukturu, přičemž některé exempláře jsou typičtější než jiné. Naproti tomu složení tzv. goal-derived kategorií, které jsou založeny na pojmových kombinacích⁴⁶ a ve kterých nejsou přístupné žádné exempláře, neumožňuje abstrahovat prototypickou informaci. Zmíněné pojmové kombinace navíc mohou být lépe charakterizovány definicemi než prototypy.“⁴⁷

Také G. L. Murphy a D. L. Medin zpochybnili ve své studii (publikované v roce 1985) dostatečnost dosavadních teorií – ty podle nich nedokážou vysvětlit nestandardní kategorizace. Poukazují přitom na biblické rozdělení zvířat na ta, která jsou čistá, a ta, která jsou nečistá. Je zřejmé, že pojmy v takovýchto kategoriích nejsou sdružovány pouze na základě podobnosti tak, jak o tom mluví dříve zmíněné teorie. Murphy a Medin předpokládají, že pojmy jsou úzce svázány s našimi (byť laickými) znalostmi o světě a jsou součástí komplexnějších vědomostí. Předkládají proto nový přístup, avšak zdůrazňují, že se nejedná o novou teorii, která by tvořila kontrast s těmi předchozími: „(...) nenavrhujeme nový model pojmové reprezentace. Spíše představujeme teorii, která objasňuje, co je tím tmelem, který drží pojem pohromadě (...).“⁴⁸

Nemožnost vysvětlit některé principy kategorizace a typy kategorií pomocí klasického přístupu nebo teorií prototypů a exemplářů nakonec vedla ke vzniku přístupu, který je označován jako „knowledge approach“, někdy také jako „theory view“, či dokonce „theory theory“.

Podle této teorie nelze pojmy, které jsou součástí našeho každodenního myšlení, vnímat jako izolované jednotky. K izolaci zpravidla dochází pouze v experimentech, které se zabývají pojmy

⁴⁵ Barsalou, L. W.: *Deriving categories to achieve goals*, 1991. S. 1. (překlad autorky)

„An infinite number of goal-derived categories exist, including *foods to eat on a diet, clothing to wear while house painting, grocery stores that sell fresh herbs (...)*, and so forth. Many of these are ad hoc categories, not established in memory but derived impromptu to achieve a current and novel goal.“

⁴⁶ Komplexní pojmy či pojmové kombinace (např. **zralé ovoce**) jsou dalším důležitým tématem týkajícím se kategorizace pojmů. Toto téma však přesahuje rámec naší práce. Stručné shrnutí této problematiky lze najít např. v Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008. S. 363.

⁴⁷ Barsalou, L. W.: *Deriving categories to achieve goals*, 1991. S. 2. (překlad autorky)

„If these two types of categories have different origins, then their cognitive structures may differ. Much previous work has found that common taxonomic categories exhibit prototype structure, with some exemplars being more typical than others. (...) In contrast, the formulation of goal-derived categories through conceptual combination in the absence of exemplars should preclude the abstraction of prototypical information. Moreover, the conceptual combination that underlies goal-derived categories may produce definitions rather than prototypes to represent these categories.“

⁴⁸ Murphy, G. L. a Medin, D. L.: *The Role of Theories in Conceptual Coherence*, 1985. S. 290.

a jejich kategorizací. Ve skutečném světě jsou ale pojmy součástí širších znalostí, které se učíme: „Když se učíme pojmy týkající se zvířat, informace je spojena s naší obecnou znalostí o biologii, o chování a s dalšími relevantními oblastmi (možná také s vařením, ekologií, klimatem atd.). Vztah je oboustranný: pojmy jsou ovlivňovány tím, co už víme, zároveň však nové pojmy mění naši obecnou znalost.“⁴⁹

V českém vydání již citované *Kognitivní psychologie* amerických autorů M. W. Eysencka a M. T. Keana je tento přístup nazýván jako „teorie pojmů založené na explanaci“ nebo „teorie pojmů založené na znalostech“. Uvádí se zde, že: „Teorie pojmů založená na explanaci vidí pojmy jako entity, které zahrnují víc než pouhé seznamy atributů – pojmy nesou také znalosti o kauzálních vztazích a další základní vědomosti (...).“⁵⁰

Tento přístup, jak už bylo naznačeno, nepopírá existenci jednotlivých vlastností (atributů, rysů) pojmů, o kterých se mluví v rámci předešlých teorií. Pouze zdůrazňuje, že tyto rysy jsou vzájemně svázány různými vztahy – uveďme opět příklad pro kategorii „**ptáci**“: například *křídla*, *lehké kosti* a *peří* jsou tři provázané vlastnosti, které úzce souvisejí s dalším atributem, a to *létáním*. Předpokládá se tedy, že struktura pojmů je komplexnější, mezi jednotlivými vlastnostmi pojmů existují vztahy a zároveň jsou pojmy součástí našich obecnějších znalostí, což zajišťuje jejich určitou koherenci a smysluplnost. Navíc je pro pojmy a jejich kategorizaci příznačná dynamika: nejde o staticky uložené znalosti, ale dochází k přizpůsobování pojmů a kategorií podle aktuální potřeby či situace, což dokládají například výše zmíněné ad hoc a tzv. goal-derived kategorie.⁵¹

Murphy zdůrazňuje, že „Je pravděpodobné, (...) že tento přístup musí být kombinován s některým z dalších přístupů, aby mohla být utvořena kompletní teorie pojmů. A nakonec, teoretici zastávající teorii exemplářů by mohli souhlasit s tím, že musí existovat také nějaká obecnější znalost vedle znalosti exemplářů, která ovlivňuje pojmy a jejich užívání (ačkoliv je pravdou, že většina z nich se tímto tématem vůbec nezabývala).“⁵²

V předchozích odstavcích jsme nastínili vývoj přístupu ke kategorizaci. Od původního přesvědčení, že kategorie jsou postaveny na jednoznačných definicích, došlo v uplynulých letech k výraznému posunu; dnes je zřejmé, že problematika kategorizace je velice komplexní. Zdá se, že aktuálně mohou další zásadní poznatky přicházet především z oblasti psychologie a neurověd, ve kterých je možno empiricky dokazovat různé teorie a mapovat neurofyziologické procesy pomocí moderních technologií.

⁴⁹ **Murphy, G. L.:** *The Big Book of Concepts*, 2002. S. 60 (překlad autorky)

„When we learn concepts about animals, this information is generated with our general knowledge about biology, about behavior, and other relevant domains (perhaps cuisine, ecology, climate, and so on). This relation works both ways: Concepts are influenced by what we already know, but a new concept can also effect a change in our general knowledge.“

⁵⁰ **Eysenck, M. W. a Keane, M. T.:** *Kognitivní psychologie*, 2008. S. 360.

⁵¹ Tamtéž, s. 360–361.

⁵² **Murphy, G. L.:** *Big book of concepts*, 2002. S. 65. (překlad autorky)

„It is likely, (...), that this view will have to be combined with one of the other views in order to form a complete theory of concepts. Finally, exemplar theorists might also agree that there must be a level of general knowledge that is separate from exemplar knowledge and that affects concepts and their use (though in fact most have not addressed this issue.“

Zmiňme jen na okraj, že zajímavý příspěvek do této diskuse může přinést také antropologický výzkum. Mohlo by se zdát, že principy kategorizace, ať už jsou ve skutečnosti jakékoliv, jsou univerzální a víceméně shodné u všech lidských bytostí. Například Lakoff však ukazuje na australském domorodém jazyce dyirbalu, že tomu tak není, a dodává, že „Je nepopíratelné, že lidé po celém světě kategorizují věci způsobem, jenž zaráží západní rozum a vyvádí z konceptu západní lingvisty a antropology.“⁵³

V našem případě nás ovšem více než kulturní odlišnosti zajímají rozdíly věkové – vývoji pojmů u dětí se věnujeme v podkapitole *Osvojování pojmů u dětí*.

Z triády sémantické rysy – pojmy – kategorie, kterou jsme zmiňovali na začátku této podkapitoly, se nyní zaměříme na její poslední část, tj. na sémantické rysy. Protože jsou sémantické rysy nedílnou součástí teorií kategorizací, byly v kontextu této podkapitoly zmiňovány již mnohokrát. Pojem bývá zpravidla chápán jako soubor těchto sémantických rysů (i když většina teorií se s chápáním pojmu jakožto pouhým souborem sémantických rysů nespokojí). Ačkoliv jsme používali různé názvy ve shodě s různými obory a autory, vždy jsou atributy, vlastnosti, či (sémantické) rysy pojmů chápány totožně, tj. jako nejmenší sémantické jednotky.

Ve strukturální sémantice se pro tyto nejmenší jednotky používá termín sém. Dříve se navíc užívaly termíny jako sémantická figura, složka či sémantický komponent⁵⁴. V psychologických a filozofických pracích se mluví také o attributech pojmů či o jejich vlastnostech⁵⁵. Ve studiích, ze kterých vychází praktická část naší práce, se pak nejčastěji setkáváme s označením „semantic feature“: „V mnoha vlivných teoriích o slovním významu, o pojmech a kategorizacích se pracuje se sémantickými rysy jakožto se základní jednotkou.“⁵⁶

Začněme nyní opět v lingvistických kruzích. Už v 50. letech 20. století Luis Hjelmslev vyslovil myšlenku, že nejen „označující“ složka znaku, ale také „označované“ (tedy význam) se skládá z menších významových částí. Na základě této myšlenky došlo k pokusům o definování znaků za pomoci jejich nejmenších významových rysů. Rozklad významu na tyto jednotky se nazývá složková analýza.

Jak taková složková analýza vypadá? „Vezměme si nějaký jednoduchý systém znaků. Například slovo *muž, žena, chlapec, dívka*. Každou jednotku se nám musí podařit jednoznačně definovat pomocí několika, pokud možno co nejméně, významových rysů. Ani u dvou jednotek nesmí být rysy totožné – museli bychom hledat další odlišující rys. Pokud bychom takový rys již nenašli, muselo by se jednat o dokonalé synonymum.“⁵⁷

⁵³ Lakoff, G.: *Ženy, oheň a nebezpečné věci*, 2006. S. 102.

⁵⁴ Černý, J. a Holeš, J.: *Sémiotika*, 2004. S. 119.

⁵⁵ Viz např. Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008, a Lakoff, G.: *Ženy, oheň a nebezpečné věci*, 2006.

⁵⁶ Mcrae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 547. (překlad autorky)

„In many of the most influential theories of word meaning and of concepts and categorization, semantic features have been used as their representational currency.“

⁵⁷ Černý, J. a Holeš, J.: *Sémiotika*, 2004. S. 119–120.

Taková analýza se pak zpravidla zobrazuje pomocí symbolů plus a mínus v přehledných tabulkách – viz příklad uvedený v tabulce 1.

	„osoba“	„dospělá“	„mužského rodu“
<i>muž</i>	+	+	+
<i>žena</i>	+	+	-
<i>chlapec</i>	+	-	+
<i>dívka</i>	+	-	-

Tab. 1. Složková analýza – muž, žena, chlapec, dívka⁵⁸

Asi mezi nejznámější komponentové analýzy patří analýza nábytku sloužícího k sezení, kterou provedl v 60. letech B. Pottier⁵⁹. Další zajímavou prací je pak například analýza rodinných vztahů, jak ji uvádí Erhart⁶⁰, ve které se odráží jazykový vývoj a rovněž rozdílnost mezi jazyky (národy) – některé sémy, které se v této analýze objevují, již v dnešní době v některých jazycích nepovažujeme za podstatné a jiné sémy naopak ve výčtu z našeho pohledu chybí.

Nejmenší významovou jednotku, ke které lze dojít sémantickým rozbořem, tedy sém, definuje E. Lotko jako „distinktivní minimální významový rys n. prvek schopný odlišit jedno slovo, lexém od jiného (význam lexému se chápe jako určitá konfigurace sémů, např. muž = + osoba, – ženské pohlaví, + dospělost)“⁶¹.

Samotná definice sémů, jak je uvedena například v *Sémiotice* J. Černého a J. Holeše, tj. „Nejmenší významové rysy, které již nemají samostatnou existenci (...)“⁶², je ve své podstatě totožná s definicí sémantického rysu, jak jej chápeme my v naší práci.

Výše zmíněná komponentová analýza se do značné míry odlišuje od databází sémantických rysů, které bývají v současné době sestavovány v rámci psycholingvistického výzkumu. Je zřejmé, že cílem složkové analýzy je rozložit znaky na co nejmenší množství sémů, které znaky definují, a zároveň je pomocí těchto sémů jednoznačně odlišit od ostatních znaků. Databáze sémantických rysů (či tzv. norem produkce sémantických rysů) nejsou však utvářeny za účelem provedení akurátní analýzy pojmů s cílem co nejefektivněji definovat jednotlivé pojmy. Pojmy v databázích nejsou pouhým výčtem nejdůležitějších a diferenčních rysů. Tyto databáze jsou soubory sémantických rysů vybraných pojmů produkovaných velkým množstvím jedinců, kteří se sběru dat účastní. K. McRae a kol. vysvětlují, že „hlavním cílem sestavování norem produkovaných sémantických rysů je vytvořit empiricky odvozené pojmové reprezentace, které mohou být využity při testování teorií týkajících se sémantických reprezentací a výpočtů.“⁶³

⁵⁸ Tabulka byla převzata z Černý, J. a Holeš, J.: *Sémiotika*, 2004. S. 120.

⁵⁹ Pottier, B.: *Vers une sémantique moderne*, 2009. S. 108–139.

⁶⁰ Erhart, A.: *Úvod do obecné a srovnávací jazykovědy*, 1973. S. 77–78.

⁶¹ Lotko, E.: *Slovník lingvistických termínů pro filology*, 2005. S. 101.

⁶² Černý, J. a Holeš, J.: *Sémiotika*, 2004. S. 119.

⁶³ Mcrae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 547. (překlad autorky)

„The major purpose of collecting semantic feature production norms is to construct empirically derived conceptual representations that can be used to test theories of semantic representation and computation.“

2.3 Rysové reprezentace pojmů – jejich podstata, využití i nedostatky

Z hlediska ukládání informací do dlouhodobé paměti jsou faktické znalosti uloženy v sémantické paměti, která je – vedle epizodické paměti, ve které se uchovávají události – součástí deklarativní (explicitní) paměti. Sémantická paměť není individuální, ukládají se v ní obecně platné informace, které jsou součástí sociokulturního kontextu. Způsob, jakým jsou tyto poznatky zakódovány v paměti, může být verbální, sensorický, či motorický v závislosti na povaze daného poznatku.⁶⁴

Zaměříme se nyní detailněji na toto zakódování poznatků v paměti. V roce 1986 píše Paivio: „Jakým způsobem reprezentujeme informace v mysli a jak využíváme tyto informace, abychom interagovali se světem vhodným způsobem? Tato otázka je stará a přetrvává. (...) Otázka přetrvává, protože je nesmírně těžká, možná vůbec nejtěžší v celé vědě.“⁶⁵

I dnes jsou pro zkoumání těchto procesů užívány alternativní metody a je možné, že přímé empirické metody, díky kterým bychom mohli pozorovat mentální reprezentace našich poznatků, nebudou nikdy objeveny.⁶⁶

Samotný výraz reprezentace pochází z latinského verba praesentare, které lze přeložit jako „zpřítomnit, představovat, ukazovat“, a z prefixu re- vyjadřujícího opakování.⁶⁷ A právě to je charakteristické pro lidskou mysl. Ta nám – mimo jiné – umožňuje „vytáhnout“ z paměti a zpětně zpřítomnit či zobrazit individuální (ve světě nebo ve fantazii existující) objekty, ale také v průběhu našeho života utvářené pojmy zastupující celou skupinu příbuzných objektů: „(...) reprezentace je jakýkoliv záznam, znak nebo množina symbolů, které nám zpětně představují nějakou věc. (...) Tato věc je zpravidla nějakým aspektem vnějšího světa nebo objektem naší imaginace (tj. našeho vnitřního světa).“⁶⁸

V naší práci se zaměřujeme na (sémantické) reprezentace objektových pojmů. Ty jsou součástí vnitřních, tedy mentálních reprezentací. Vnitřní reprezentace jsou protipólem reprezentací vnějších, za které jsou považovány například obrázky (reprezentace obrazová, například) či písmo (reprezentace jazyková).⁶⁹ Podobným způsobem dělíme vnitřní reprezentaci na analogovou (například zrakové či sluchové představy) a abstraktní, na lidských smyslech nezávislou reprezentaci propoziční, která do jisté míry připomíná jazyk – tento jazyk je však univerzální a bývá označován termínem „mentálština“. V rámci propoziční reprezentace jsou rozlišovány objektové pojmy (tj. např. **pes**) a relační pojmy (tj. např. **praštit**), tyto pojmy se dále v souladu s určitými pravidly kombinují a vytvářejí schémata.⁷⁰

Již bylo zmíněno, že teorie zabývající se otázkou, jakým způsobem lidé kategorizují pojmy (potažmo objekty), stojí zpravidla na předpokladu existence rysů (atributů). Rysy a rysové

⁶⁴ Vágnerová, M.: *Základy psychologie*, 2004. S. 64–65.

⁶⁵ Paivio, A.: *Mental Representations. A Dual coding approach*, 1990. S. 3. (překlad autorky)

„How do we represent information mentally and how do we use that information to interact the world in adaptive ways? The problem is old and persistent. (...) The problem persists because it is extraordinarily difficult, perhaps the most difficult one in all of science.“

⁶⁶ Sternberg, R. J.: *Kognitivní psychologie*, 2009. S. 243.

⁶⁷ Podle M. Sedlákové: *Vybrané kapitoly z kognitivní psychologie*, 2004. S. 43.

⁶⁸ Eysenck, M. W. a Keane, M. T.: *Kognitivní psychologie*, 2008. S. 298.

⁶⁹ Tamtéž, s. 297–298.

⁷⁰ Tamtéž, s. 301–303.

reprezentace pojmů jsou běžně považovány za součást (konceptuálních) reprezentací: „Sémantické rysy umožňují zkoumat mnoho behaviorálních jevů týkajících se pojmů, kategorizace a sémantické paměti v případě dospělých jedinců, ale i dětí a pacientů s neuropsychologickými problémy. Velké množství teorií a modelů v této oblasti je založeno na reprezentaci a komputaci za pomoci sémantických rysů. Proto normy sémantických rysů, které jsou výsledkem přirozené produkce, hrály a stále hrají významnou roli v těchto oborech.“⁷¹

Je zřejmé, že konceptuální reprezentace jsou mnohem komplexnější a není možné je zaměňovat s pouhými rysovými reprezentacemi. To ostatně přiznávají i samotní autoři databází sémantických rysů: „(...) nedomníváme se, že sémantické znalosti jsou v mozku skutečně reprezentovány jako seznam verbalizovaných rysů. Je pravděpodobnější, že participanti produkují sémantické rysy v normované úloze z reprezentací, které si vytvořili na základě opakovaného vícesmyslového prezentování členů dané kategorie a interakcí s těmito exempláři.“⁷²

Předpokládá se, že jen v některých typech mentálních procesů týkajících se pojmů a jejich kategorizace jsou využívány jako výchozí zdroj informací rysové reprezentace – o tom, z jakých zdrojů čerpáme informace, když řešíme sémantické úlohy, které vyžadují rychlé (nebo naopak pomalé) úsudky, mluvíme v následující podkapitole.

Databáze sémantických rysů a z nich odvozené rysové reprezentace nám nabízejí možnost prozkoumat obsahy vybraných pojmů a alespoň částečně nahlédnout do jejich vnitřní struktury. Zároveň je možné popsat vztahy mezi jednotlivými pojmy a kategoriemi i jevy týkající se taxonomizace. Obsah databáze je však do jisté míry závislý na individuálních znalostech participantů, kteří se zúčastní sběru dat, a na jejich přístupu k úkolu definování pojmů. V případě odvozování rysových reprezentací z dat nashromážděných v databázi není získán pouze malý soubor rysů podstatných pro charakterizaci významu pojmu a jeho diferenciaci od pojmů ostatních. Participanti nemají za úkol cíleně porovnávat pojmy, jako se o to pokoušejí badatelé v komponentové analýze. Participanti mají za úkol popisovat vždy jeden vybraný pojem; není ale vyloučeno, že se zcela přirozeně ve své mysli uchýlí k porovnávání pojmu, jehož význam popisuje, s dalšími podobnými pojmy.

⁷¹ **Mcrae, K. et al.:** *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 547. (překlad autorky)

„Semantic features have provided insight into numerous behavioral phenomena concerning concepts, categorization, and semantic memory in adults, children, and neuropsychological populations. Numerous theories and models in these areas are based on representations and computations involving semantic features. Consequently, empirically derived semantic feature production norms have played, and continue to play, a highly useful role in these domains.“

⁷² Tamtéž, s. 549. (překlad autorky)

„(...) we do not believe that semantic knowledge is represented in the brain literally as a list of verbalize features. Rather, when participants produce features in a norming task, they directly exploit representations that have developed through repeated multisensory exposure to, and interactions with, exemplars of the target category.“

Kromě statistických analýz dat z databází a jejich interpretací je možné využít tyto databáze sémantických rysů jako podkladový materiál pro mnohé psychologické a neurologické výzkumy, které se zabývají otázkami lidské mysli.

Domníváme se, že data získaná od dětských participantů mohou také přispět k výzkumu dětské řeči (nabízí se například srovnání s daty z databází vytvořených z materiálu získaného od dospělých mluvčích, v českém prostředí viz například práci M. Nagye⁷³).

2.4 Využití rysových reprezentací v rychlých mentálních úvahách o pojmech

Současný výzkum pojmů se zaměřuje především na komplikovanější a komplexnější znalosti, tzv. znalosti vyšší úrovně (*higher level knowledge*), přičemž se předpokládá, že právě tyto znalosti mají významný vliv na obsah pojmů. Takový přístup je chápán jako alternativní k tradičním přístupům, které považují za základní pojmovou jednotku sémantický rys. K. McRae se svými kolegy však poukazuje na to, že tyto dva přístupy se s velkou pravděpodobností nevylučují: první zmiňovaný přístup se uplatňuje spíše v situacích, kdy řešíme komplikované sémantické úlohy, zatímco ten druhý souvisí s každodenními rutinními procesy, kdy rozpoznáváme objekty ve svém okolí a patřičně na ně reagujeme.⁷⁴

Cílem níže předkládaného experimentu je potvrdit následující hypotézu: na základě výše zmíněné úvahy je možné předpokládat, že v rychlých sémantických úlohách dochází při rozhodování k využití rysových reprezentací.

Pokud však úloha nevyžaduje rychlé odpovědi a nabízí možnost důkladnějšího promýšlení, do hry se zapojují i další, komplexnější kognitivní schopnosti a znalosti: „S úlohami, které [na pomyslné škále] spadají do [využití znalostí] vyšší úrovně, je významně spjato hlubší přemýšlení. Vyvozování závěrů vychází z mnoha zdrojů informací, jedná se například o smyslové vnímání předmětů, znalost jazyka, přehled v oblasti biologie, či o zkušenost s psychologickými experimenty, tj. o znalost, jak tyto experimenty fungují a jak v nich reagovat.“⁷⁵

K. McRae se svými kolegy⁷⁶ odkazuje na příklad se zebrou a koněm ze studie F. C. Keila⁷⁷. Jedná se o úlohu, ve které mají děti rozhodovat, zda zvíře, které se z důvodů uvedených v úloze změnilo na zvíře jiné, stále je nebo už není považováno za původní zvířecí druh. Dětem v této úloze byla představena situace tohoto typu:

⁷³ Nagy, M.: *Produkce sémantických rysů u dospělých mluvčích*, 2014.

⁷⁴ McRae, K., Sa, V. R. de a Seidenberg, M. S.: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 99.

⁷⁵ Tamtéž, s. 100. (překlad autorky)

„Tasks situated toward the high-level end are those that demand considerable reasoning (...). Performing (...) judgments requires integrating a number of information sources, such as the perceptual features of objects, knowledge of one's language, general knowledge of biology, and knowledge of the pragmatics of answering questions in psychology experiments.“

⁷⁶ Tamtéž.

⁷⁷ Keil, F. C.: *Concepts, kinds, and cognitive development*, 1989.

„Nedaleko Drydenu žijí na farmě zvířata, která vypadají takto (je ukázán obrázek mývala). Často se plíží kolem domů, občas snědí nějaké odpadky a moc rádi omývají své jídlo v potůčcích. No, a nějací vědci se vydali tato zvířátka pečlivě prostudovat pomocí mikroskopů a podobných nástrojů, a když se na ně podívali zblízka a prozkoumali je i zevnitř, zjistili, že vlastně nejsou vůbec podobní mývalům. Tito vědci zjistili, že ta zvířata mají skunčí krev, kosti i mozek a vůbec všechny vnitřnosti jsou spíše skunčí než mývalí. Pak ještě zjistili, že maminky a tatínkové těchto zvířat byli taky skunci, a když tato zvířata měla děti, zase to byli skunci. Takže maminky a tatínkové vypadali takhle (je ukázán obrázek skunka) a jejich děti taky vypadaly takhle (opět je ukázáno na skunka). Co myslíš, co jsou tedy tato zvířata ve skutečnosti zač, jsou to mývalové nebo skunkové?“⁷⁸

Úlohy se zúčastnily děti předškolního a mladšího školního věku. V odpovědích lze vidět, že reakce na takovou otázku je spjata se složitějšími mentálními procesy:

„C: Řekl bych že skunk... Ano, skunk.

E: Proč?

C: Protože má všechno jako skunk, jen nemá... jen navenek tak nevypadá. A jeho děti byli skunkové a jeho rodiče byli skunkové.

E: A to nevdá, že vypadá takhle (ukazuje na obrázek mývala) nebo že se hrabe v odpadcích a že nesmrdí?

C: Ani moc ne...“⁷⁹

Úloha, kterou dětem předkládáme my, je časově omezená a vyžaduje rychlý úsudek – probandi odpovídají na jednoduchou sémantickou otázku pouze „ano“ nebo „ne“ pro jednotlivé pojmy. Takový typ úlohy patří na opačný konec pomyslného spektra úrovní znalostí: „Na opačném konci [pomyslné škály] se nacházejí úlohy spjaté se znalostmi nižší úrovně, jako je rychlé pojmenovávání objektů, ověřování rysů, posuzování podobnosti a odlišnosti a sémantické rozhodování (například „je to skutečný předmět?“). Výkon v takovýchto typech úloh může významně záviset na rysových reprezentacích a jen minimálně jej můžou ovlivňovat procesy spojené s hlubším přemýšlením a zdůvodněním.“⁸⁰

⁷⁸ Keil, F. C.: *Concepts, kinds, and cognitive development*, 1989. S. 176–177. (překlad autorky)

„There are animals that live out on a farm near Dryden and they look just like this (shows picture of raccoon). They prowl around houses a lot and sometimes eat garbage, and they like to wash their food in streams. Well, some scientists went and studied these animals very carefully with microscopes and other sorts of stuff, and when they looked at them and their insides, they found out that they weren't like most raccoons at all. The scientists found out that they had the blood of skunks, the bones of skunks, the brains of skunks, and all the other inside parts of skunks rather than those of raccoons. And they found out that these animals' mommies and daddies were skunks and that when they had babies their babies were skunks. So their mommies and daddies look just like this (points to picture of skunk) and their babies look just like this (points to picture of skunk). What do you think they really are, raccoons or skunks?“

⁷⁹ Tamtéž, s. 180. (překlad autorky)

„C: I'd say a skunk...yes, a skunk.

E: Why?

C: Because it has everything that a skunk has only it doesn't...only the outside parts of it doesn't look like that. And its babies were a skunk and its parents were a skunk.

E: Does it matter that it looks like that (points to picture of raccoon) or that it paws through your trash or that it doesn't stink?

C: Not really, not really.“

⁸⁰ McRae, K., Sa, V. R. de a Seidenberg, M. S.: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 100. (překlad autorky)

2.5 Priming a sémantický priming

Součástí předkládané práce je experiment, který má za cíl ověřit, zda má míra rysově podobnosti dvou pojmů vliv na sílu primingového efektu v úloze rychlého sémantického rozhodování (detailněji byla tato teorie představena v předchozí podkapitole a navazujeme na ni také v praktické části této práce).

Priming je mentální jev, který G. T. M. Altmann vysvětluje jako „otevření scénáře, které usnadňuje například vybavení slova v paměti. Mentální zpracování určitého druhu podnětu je usnadněno tím, že následuje po podnětu stejného typu.“⁸¹

V psycholingvistice se primingového efektu využívá při zkoumání jevů na různých jazykových úrovních – od fonologické až po konceptuální. Jako příklad se zpravidla uvádí priming v úlohách lexikálního rozhodování (detekce slova). Pokud se zkoumanému participantovi prezentuje (vizuálně či auditivně) jako primingové slovo ‚muž‘, jeho reakce na cílové slovo (v angličtině se používá označení „target“) je rychlejší, pokud jsou oba výrazy ze stejného okruhu slovní zásoby. Target tedy může být v takovém případě například slovo ‚chlapec‘. Prezentujeme-li však jako target slovo ‚kniha‘, k tomuto zrychlení nedochází. Tento efekt se projevuje například v úlohách, ve kterých posuzujeme, zda je zobrazené cílové slovo skutečné, či se jedná o tzv. pseudoslovo – takové slovo má zpravidla strukturu, která připomíná skutečná slova daného jazyka, nic však neznamená, nenese žádný význam, nebo neslovo (v takovém případě slovo neexistuje a navíc porušuje také fonetická pravidla odpovídající danému jazyku).⁸²

Uvedme příklad se slovy ‚člun‘, ‚lod‘, ‚led‘ a s pseudoslovem ‚slun‘. Účastníkům nějakého experimentu je představeno slovo ‚člun‘, nebo pseudoslovo ‚slun‘. Participant se musí rozhodnout, zda je zobrazené slovo skutečné, či nikoliv. Toto rozhodování je jednoduchá a velmi rychlá záležitost – pokud je však participantům v předchozím kroku zobrazeno primové slovo ‚lod‘, dochází k urychlení rozhodování, to se však nestane v případě nepříbuzného slova ‚led‘.⁸³

Za účelem ověření předkládané hypotézy vytváříme v našem experimentu podmínky pro vznik tzv. sémantického primingu. Sémantický priming je takový aktivizační mechanismus, kdy „odpověď na cílový pojem (např. pes) je rychlejší, jestliže mu předchází sémanticky příbuzný prime (např. kočka); při nepříbuzném primu (např. čůčka) tento efekt nenastává. Sémantický priming může nastat díky primu, který částečně aktivuje příbuzná slova nebo pojmy a urychlí jejich pozdější zpracování nebo rozpoznání.“⁸⁴

„Toward the low-level end are speeded object naming, feature verification, same-different judgements, and semantic decisions (e.g., „is it a concrete object?“). Performance on these tasks may depend largely on featural representations, with minimal influence from higher level reasoning processes.“

⁸¹ Altmann, G. T. M.: *Výstup na babylonskou věž*, 2005. S. 18.

⁸² Karlík, P., Nekula, M. a Pleskalová, J. (eds.): *Nový encyklopedický slovník češtiny* [online], 2012–2020, heslo *Psycholingvistické metody*.

Dostupné z: www.czechency.org/slovník/PSYCHOLINGVISTICK%C3%89%20METODY [cit. 13. října 2020]

⁸³ Altmann, G. T. M.: *Výstup na babylonskou věž*, 2005. S. 19.

⁸⁴ *Lexical decision tasks, semantic priming, and reading* [online], 2014. (překlad autorky)

Dostupné z: www.apa.org/pubs/highlights/peeps/issue-33 [cit. 1. 6. 2021]

Priming je zpravidla automatický proces⁸⁵, ve kterém primový vjem ovlivňuje reakci na target podvědomě. Chceme-li v sémantické úloze záměrně vyvolat primingový efekt tak, aby jej účastníci experimentu neodhalili a nepokoušeli se úkol řešit alternativními způsoby (např. předběžně generovat pojmy, které jsou s primovým slovem významově příbuzné, případně zpětně kontrolovat při zobrazení cílového slova jeho vztah s primem⁸⁶), je důležité zvolit správnou metodu, která eliminuje tyto nežádoucí postupy.

2.6 Osvojování pojmů u dětí

Téma osvojování slov postupem času zaujalo právem důležité místo ve vědě. Jak zmiňuje např. Paul Bloom, „Učení se slovům a zvláště jmenům pro věci se rozhodně *zdá* jako jednoduchý proces, přinejmenším badatelům, kteří se na výzkumu této problematiky nepodílí.“⁸⁷ Z přibývajících výzkumů je však zřejmé, že nejen osvojování slovní zásoby, ale veškeré učení se rodné řeči a schopnostem komunikace pomocí rodného jazyka je ve skutečnosti velice komplexní a komplikovaný kognitivní proces.

Seznam zásadních otázek týkajících se této problematiky lze najít například u již výše jmenovaného P. Blooma: vědce zajímá, jakým způsobem se dítě učí chápat vztah mezi slovem a označovaným předmětem a jak si utváří pojem, který označuje všechny podobné předměty a nikoliv jen jednu konkrétní viděnou věc, jak se učí rozumět pojmům, které není možné vnímat zrakovými smysly a není tedy na ně možné poukazovat přímo, jaký podíl má na osvojování záměrná aktivita rodičů (ve smyslu odkazování na věci, vysvětlování, opravování).⁸⁸

Psycholingvistický výzkum v oblasti dětské řeči na sémantické úrovni je nicméně ve většině současných případů zaměřen spíše na rozšiřování slovní zásoby než na hloubku slovníku ve smyslu osvojování významů slov: „Zatímco osvojování lexika dětmi je často objasňováno se zaměřením na kvantitativní stránku procesu, tj. rozšiřování slovní zásoby v jednotlivých fázích věku, kvalitativní stránka procesu, tj. osvojování slov jakožto sémantických kategorií, je objasněna mnohem méně. Některé poznatky svědčí o tom, že řeč matek usnadňuje dětem osvojovat si významy slov specifickým způsobem: Tím, že autosémantická slova (plnovýznamová), tj. podstatná jména, slovesa, příslovce jsou v řeči matek (různých jazyků) často vyslovována opakovaně, izolovaně či se silným zdůrazněním, s poukazováním na objekty, s nimiž je slovo spojeno.“⁸⁹

V centru zájmu vědců ale není jen učení se slovům – osvojování sémantické/lexikální kompetence je jen jednou z jazykových úrovní, která jsou zkoumány. Badatele zajímají všechny jazykové roviny a (dětská) řeč jako celek. Samotný proces osvojování řeči je přitom

„(...) a response to a target (e.g., dog) is faster when it is preceded by a semantically related prime (e.g., cat) compared to an unrelated prime (e.g., car). Semantic priming may occur because the prime partially activates related words or concepts, facilitating their later processing or recognition.”

⁸⁵ *Lexical decision tasks, semantic priming, and reading* [online], 2014.

Dostupné z: www.apa.org/pubs/highlights/peeps/issue-33 [cit. 20. 6. 2021]

⁸⁶ Tamtéž. [cit. 20. 6. 2021]

⁸⁷ Bloom, P.: *Jak se děti učí významu slov*, 2015. S. 15.

⁸⁸ Bloom, P.: *Jak se děti učí významu slov*, 2015.

⁸⁹ Průcha, J.: *Dětská řeč a komunikace*, 2011. S. 72.

nejrychlejší a nejzásadnější v předškolním věku, proto jsou hledány metody, které nabízejí možnost zkoumat komunikační schopnosti dítěte ještě dříve, než se narodí a než začíná vůbec mluvit. Např. G. Altman zmiňuje ve své knize *Výstup na babylonskou věž* experimenty J. Mehlera⁹⁰ či A. DeCaspera⁹¹ a jejich kolegů, kteří už v 80. letech zkoumali schopnost dětí rozlišit od sebe různé jazyky a známé příběhy od neznámých už v prenatálním a novorozeneckém období.⁹²

Velké množství výzkumů věnujících se vývoji dětské řeči se zaměřuje na děti mladšího (předškolního) věku. To je pochopitelné především z toho důvodu, že v tomto předškolním období se řečová schopnost vyvíjí nejzásadněji. V této životní fázi lze pozorovat řadu důležitých změn v tělesném a psychickém vývoji, tedy právě i ve vývoji řeči.

Naše databáze byla sestavena z dat získaných od dětí ve věku 8–10 let. Experiment byl prováděn se stejnou věkovou kategorií. Pro výběr dětí nižšího školního věku jsme se rozhodli z několika důvodů. Ve chvíli, kdy dítě nastupuje do školy, je sice splněna velká část jazykového a komunikačního vývoje, přesto je z hlediska vývoje sémantické jazykové složky období nižšího školního věku zajímavé: „Celkový vývoj poznávacích procesů se projeví i změnou způsobu, jakým děti uvažují o slovech, co pro ně určitý slovní výraz (např. pomeranč) znamená. Malí školáci je obvykle definují ve vztahu ke konkrétnímu kontextu (máme je vždycky u babičky), popisem jeho vnějších znaků nebo aktivitou, která je s ním spojena (jsou k jídlu). Postupně se charakteristika slova stává obecnější, v průběhu školního věku se zvyšuje tendence užít k takové charakteristice nadřazenou kategorii, tj. dedukce (ovoce) nebo definovat slovo pomocí vztahu tohoto objektu k jiným objektům, resp. pojmům stejné úrovně, tj. analogií (např. pomeranč je jako citron, event. protikladnými analogiemi, např. kyselý je opak sladkého).“⁹³

Při výběru účastníků výzkumu byla nastavena spodní věková hranice 8 let také z praktických důvodů: bylo nezbytné, aby účastníci disponovali potřebnými kompetencemi: sběr dat do databáze sémantických rysů vyžadoval schopnost umět se písemně se vyjadřovat⁹⁴ a pro účast v experimentu bylo nutné umět rychle číst.

⁹⁰ Mehler, J. et al.: *A precursor of language acquisition in young infants*, 1988.

⁹¹ DeCasper, A. a Spence, M. J.: *Prenatal maternal speech influences newborns' perception of speech sounds*, 1986.

⁹² Altmann, G.: *Výstup na babylonskou věž*, 2005. S. 24–28.

⁹³ Vágnerová, M.: *Kognitivní a sociální psychologie žáka základní školy*, 2001. S. 110.

⁹⁴ Více k výběru metody sběru dat viz níže v praktické části (v podkapitole *Pilotní studie*).

3. Empirická část

3.1 Úvod

Primárním cílem této práce bylo vytvoření komplexní databáze sémantických rysů pro vybrané pojmy z různých sémantických kategorií a následná analýza a interpretace získaných dat. Sémantické rysy pojmů byly získány pomocí empirické metody – to znamená, že tyto rysy byly abstrahovány z definic pojmů, které sepisovaly děti ve věku 8–10 let (tedy děti mladšího školního věku), které se zúčastnily úkolu, jenž byl součástí našeho výzkumu. Kompletní databáze a výsledky jejího statistického zpracování jsou zpřístupněny online na internetové adrese www.childrenfeatures.webnode.cz. Tato databáze může sloužit jako zdroj dat pro další psycholingvistické experimenty. Získaná data však nejsou určena pouze vědecké komunitě – domníváme se, že vedle využití v psycholingvistických výzkumech věnujících se dětské řeči může být databáze využita například také jako podnětný studijní materiál pro pedagogy, či pro rodiče.

Před vytvořením vlastní databáze byl připraven pilotní experiment, ve kterém jsme ověřovali několik možných metod týkajících se postupů při sběru dat. V této fázi byla také předběžně stanovena tzv. kódovací pravidla.

V další části výzkumu byl uskutečněn vlastní sběr dat, ve kterém byla získána data celkem od 381 českých dětí. Na základě těchto získaných dat vznikla databáze sémantických rysů pro 104 substantivních pojmů z 13 různých sémantických kategorií. Z databáze byly vyvozeny informace o sémantických rysech, pojmech a kategoriích (včetně vztahů mezi nimi) pomocí vybraných statistických metod.

V poslední části výzkumu byl proveden experiment, ve kterém jsme ověřovali využití rysových reprezentací v sémantických úlohách, které vyžadují rychlou reakci při zpracovávání podnětů vztahujících se k obsahům pojmů. I v tomto případě se experimentu zúčastnily děti nižšího školního věku, tj. 8–10letí školáci.

Jednotlivé části práce jsou představeny také ve studiích *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*⁹⁵, *A database of semantic features for chosen concepts (Attested in 8- to 10-year-old Czech pupils)*⁹⁶ a v připravované studii *The influence of featural similarity strength on the priming effect in a speeded task*⁹⁷.

⁹⁵ Konečná, K. a Večeřová, L.: *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*, 2016.

⁹⁶ Konečná, K. a Vaňkátová, K.: *A database of semantic features for chosen concepts (Attested in 8- to 10-year-old Czech pupils)*, 2017.

⁹⁷ Konečná, K. a Smolík, F.: *The influence of featural similarity strength on the priming effect in a speeded task*, 2021. (v recenzním řízení)

3.2 Výběr kategorií a pojmů do databáze sémantických rysů

Představovaná databáze sémantických rysů byla sestavena pro 104 objektových pojmů ze 13 sémantických kategorií.

Naším cílem bylo získat sémantické rysy pojmů/slov, které jsou součástí základní slovní zásoby. Podstatou tohoto výzkumu nebylo ověřit šířku dětského slovníku, tedy to, do jaké míry jsou zkoumané pojmy dětskými participanty osvojené, nebo zda vůbec znají jejich význam. Sémantické rysy mohou být shromážděny a rysová reprezentace mohou být vytvořeny pouze pro takové pojmy, u kterých mají školáci daného věku určitou představu o jejich významu, ať už je správná, nebo ne. Výběr pojmů byl zároveň podmíněn jejich příslušností k vybraným sémantickým kategoriím.

Sémantické kategorie byly vybrány tak, aby polovina z nich byla vztažena k tzv. biologickým kategoriím (*living things, natural kinds*), tj. kategoriím obsahujícím pojmy, jež se vyskytují ve světě přirozeně, a druhá polovina k tzv. artefaktům (*artifacts*), tedy objektům, které jsou dílem člověka a bez přičinění lidské společnosti by neexistovaly. Toto rozdělení vychází z dřívějších vědeckých poznatků, že se tyto dva typy kategorií od sebe strukturně liší (viz například práci F. C. Keila *Concepts, Kinds, and Cognitive Development*⁹⁸).

Pro databázi bylo vybráno celkem šest biologických kategorií a sedm artefaktových kategorií. Výběr konkrétních kategorií byl ovlivněn několika anglosaskými studii. Deset kategorií („savci“, „ptáci“, „ovoce“, „zelenina“, „oblečení“, „nábytek“, „kuchyňské náčiní“, „nářadí“, „dopravní prostředky“ a „zbraně“) bylo převzato ze studie *On the Nature and Scope of Featural Representations of Word Meaning*⁹⁹, přičemž autoři této práce vybírali z kategorií analyzovaných ve studii *Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms*¹⁰⁰. Zbývající tři kategorie („rostliny“, „části těla“, „hudební nástroje“) byly doplněny z dalších studií věnujících se sémantické kategorizaci u dětí (*Does sex influence the age of acquisition of common names? A contrast of different semantic categories*¹⁰¹ a *Category norms for verbal items in 25 categories for children in Grades 2-6*¹⁰²). Pojmy byly pro jednotlivé kategorie vybrány na základě sběru dat od 40 dospělých osob. V rámci dotazníkového šetření tyto osoby dostaly za úkol napsat pro každou z prezentovaných kategorií pět nejtypičtějších zástupců – tedy pět pojmů, které kategorii nejlépe charakterizují či reprezentují.

V prvním kole tohoto sběru byla data sbírána pouze pro deset vybraných kategorií. Byla mezi nimi zařazena také kategorie „zvířata“. V druhém, doplňujícím kole sběru byla tato kategorie rozdělena na dvě samostatné kategorie „savci“ a „ptáci“, dále byly přidány kategorie „zbraně“ a „kuchyňské náčiní“.

Ze získaných odpovědí se zdá, že někteří participanti pravděpodobně nerespektovali zadání a vypisovali spíše ty zástupce, kteří jim přišli na mysl jako první. Nicméně i tyto odpovědi byly akceptovány a zpracovány.

⁹⁸ Keil, F. C.: *Concepts, kinds, and cognitive development*, 1989.

⁹⁹ McRae, K., Sa, V. R. de a Seidenberg, M. S.: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997.

¹⁰⁰ Battig, W. F. a Montague, W. E.: *Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms*, 1969.

¹⁰¹ Barbarotto, R., Laiacona, M. a Capitani, E.: *Does sex influence the age of acquisition of common names? A contrast of different semantic categories*, 2008.

¹⁰² Posnansky, C. J.: *Category norms for verbal items in 25 categories for children in grades 2-6*, 1978.

Pojmy byly seřazeny na základě četnosti jejich produkce participanty od nejméně frekventované. Z různých částí této frekvenční škály byly pro každou kategorii vybrány čtyři pojmy a zbývající čtyři pojmy byly doplněny s ohledem na plánovaný experiment tak, aby pro každou kategorii byly získány čtyři páry pojmů s různě vysokou vzájemnou podobností. Empiricky podložená podobnost rysových reprezentací pojmů může být samozřejmě známa až ve chvíli, kdy je databáze zpracována a pojmová podobnost je přímo vypočítána ze získaných dat. Proto byly pojmy ve fázi jejich výběru do databáze spárovány pouze předběžně, a to na základě našeho předpokladu míry jejich podobnosti. Takovým intuitivním způsobem při párování pojmů postupoval také K. McRae a jeho kolegové: „(...) stupeň podobnosti mezi těmito páry se v určité míře různil (ačkoli byly všechny z podobné skupiny). To znamená, že na základě intuice byly některé v podstatě synonymní (např. *pohovka–gauč*), zatímco jiné si byly podobné jen trochu (*mrkev–celer*)“¹⁰³.

Celkem bylo tedy do databáze vybráno 104 pojmů: 4 páry (tj. 8 pojmů) z 13 různých kategorií. Kompletní přehled vybraných kategorií a pojmových párů lze najít v následující tabulce 2 (pro biologické pojmy) a tabulce 3 (pro artefakty). Páry jsou seřazeny na základě míry podobnosti mezi párovými pojmy. V těchto dvou tabulkách je navíc u každého pojmu uveden počet dospělých participantů (ze 40 celkem), kteří zmínili daný pojem v dotazníku.

Jak lze vidět, některé pojmy byly zmíněny pouze malým počtem dotazovaných osob, přesto se staly součástí databáze. To lze zdůvodnit následovně: za prvé, jejich výběr navzdory jejich nízké frekvenci byl ovlivněn potřebou najít mezi uvedenými pojmy páry s různě vysokou podobností. Za druhé, tyto pojmy byly uváděny jako typičtí zástupci kategorie, i když pouze malým počtem dotazovaných. V neposlední řadě byl výběr pojmů podmíněn – s ohledem na cíl výzkumu – naším osobním zvážením, zda je možné považovat tyto pojmy za dobré zástupce kategorie (proto nebyly vůbec brány v potaz pojmy typu **fénix** v kategorii „**ptáci**“) a zda se jedná o pojmy, u kterých lze předpokládat, že je budou dětští probandi znát, a tudíž budou schopni zapsat jejich význam.

Všechny pojmy byly konfrontovány s frekvencí výskytu ve *Weslalexu* (tj. v databázi slov, která se vyskytují v učebnicích pro děti v západoslovanském prostředí – název *Weslalex* je zkratkou pro *West Slavic lexicon*¹⁰⁴), v českém korpusu *InterCorp verze 12* a ve filmových titulcích obsažených v tomtéž korpusu¹⁰⁵. Tyto údaje jsou uvedeny v přehledové tabulce v příloze 1.

BIOLOGICKÉ KATEGORIE	KATEGORIE	míra podobnosti	POJEM	produkční frekvence (1. pojem)	PÁROVÝ POJEM	produkční frekvence (2. pojem)
	OVOCE	nejvyšší	VIŠEŇ	1/40	TŘEŠEŇ	5/40
		vyšší	MANDARINKA	3/40	POMERANČ	22/40
		nižší	JAHODA	18/40	MALINA	3/40
		nejnižší	JABLKO	40/40	ANANAS	4/40
	ZELENINA	nejvyšší	OKURKA	28/40	CUKETA	2/40

¹⁰³ McRae, K., Sa, V. R. de a Seidenberg, M. S.: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 103. (překlad autorky)

„(...) the degree of similarity among those pairs varied somewhat (even though all were still part of the similar group). That is, based on intuition, some pairs were virtually synonymous (e.g., *sofa–couch*), whereas others were only moderately similar (*carrot–celery*).“

¹⁰⁴ Dostupné z: <http://spell.psychology.wustl.edu/weslalex/> [cit. 15. října 2021]

¹⁰⁵ Rosen, A., Vavřín M. a Zasina, A. J.: *Korpus InterCorp – čeština, verze 12*, 2019. Dostupné z: <https://kontext.korpus.cz>

	vyšší	FAZOLE	1/40	HRÁŠEK	2/40
	nižší	KVĚTÁK	8/40	SALÁT	12/40
	nejnižší	CIBULE	11/40	MRKEV	26/40
ROSTLINY	nejvyšší	SEDMIKRÁSKA	10/40	KOPRETINA	11/40
	vyšší	BOROVICE	3/40	LÍPA	2/40
	nižší	RŮŽE	19/40	TULIPÁN	10/40
	nejnižší	LEKNÍN	1/40	ČTYŘLÍSTEK	1/40
ČÁSTI TĚLA	nejvyšší	NOHA	37/40	RUKA	36/40
	vyšší	PAŽE	3/40	STEHNO	1/40
	nižší	UCHO	6/40	NOS	6/40
	nejnižší	SRDCE	1/40	ZUB	1/40
PTÁCI	nejvyšší	OREL	23/40	SUP	5/40
	vyšší	RACEK	3/40	HOLUB	13/40
	nižší	HUSA	1/40	LABUŤ	3/40
	nejnižší	PAPOUŠEK	9/40	SOVA	6/40
SAVCI	nejvyšší	KRÁLÍK	3/40	ZAJÍC	1/40
	vyšší	KOČKA	33/40	TYGR	5/40
	nižší	JELEN	1/40	KŮŇ	11/40
	nejnižší	DELFIN	10/40	OPICE	6/40

Tab. 2. Kategorie a pojmy vybrané do databáze sémantických rysů (biologické kategorie)

ARTEFAKTY	KATEGORIE	míra podobnosti	POJEM	produkční frekvence (1. pojem)	PÁROVÝ POJEM	produkční frekvence (2. pojem)
		NÁBYTEK	nejvyšší	SEDAČKA	3/40	ŽIDLE
		vyšší	POSTEL	24/40	POHOVKA	12/40
		nižší	SKŘÍŇ	36/40	VĚŠÁK	2/40
		nejnižší	LAMPA	1/40	DVEŘE	2/40
	NÁŘADÍ	nejvyšší	LOPATA	6/40	RÝČ	7/40
		vyšší	KOŠTĚ	1/40	HRÁBĚ	3/40
		nižší	KLADIVO	34/40	SEKERA	7/40
		nejnižší	KLEŠTĚ	17/40	PILA	19/40
	HUDEB. NÁSTROJE	nejvyšší	KLAVÍR	27/40	VARHANY	1/40
		vyšší	HOUSLE	32/40	KYTARA	32/40
		nižší	TRUMPETA	11/40	FLÉTNA	28/40
		nejnižší	TRIANGL	1/40	HARFA	5/40
	OBLEČENÍ	nejvyšší	KLOBOUK	1/40	ČEPICE	11/40
		vyšší	SUKNĚ	21/40	ŠATY	7/40
		nižší	TRIKO	37/40	SVETR	6/40
		nejnižší	BOTA	6/40	PYŽAMO	1/40
	DOPR. PROSTŘEDKY	nejvyšší	TROLEJBUS	5/40	AUTOBUS	32/40
		vyšší	VRTULNÍK	1/40	LETADLO	23/40
		nižší	AUTO	39/40	TRAKTOR	1/40

	nejnižší	VLAK	28/40	LOŽ	11/40
ZBRANĚ	nejvyšší	PISTOLE	25/40	REVOLVER	8/40
	vyšší	MEČ	15/40	DÝKA	5/40
	nižší	PRAK	3/40	LUK	6/40
	nejnižší	OŠTĚP	1/40	DĚLO	11/40
KUCHYŇSKÉ NÁČINÍ	nejvyšší	SKLENICE	1/40	HRNEK	2/40
	vyšší	NABĚRAČKA	18/40	LŽICE	6/40
	nižší	HRNEC	27/40	KONVICE	2/40
	nejnižší	STRUHADLO	5/40	VAŘEČKA	35/40

Tab. 3. Kategorie a pojmy vybrané do databáze sémantických rysů (kategorie artefaktů)

3.3 Pilotní studie

Pilotní sběr dat byl uskutečněn ve spolupráci s L. Večeřovou. Kompletní popis tohoto sběru, analýzy a závěry jsou shrnuty ve studii *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*¹⁰⁶. Zde představujeme pouze tu část pilotní studie, která je relevantní pro naši práci.

3.3.1 Cíl pilotní studie

Pilotní sběr dat byl zaměřen především na výběr nejvhodnější metody pro následný vlastní sběr dat a také na vytvoření tzv. „kódovacích pravidel“, podle kterých jsou odpovědi participantů systematicky rozděleny na jednotlivé sémantické rysy (jsou tzv. zakódovány).

Pro databázi byl nejprve vytvořen seznam kategorií a pojmů. Vedle objektových pojmů, na které je naše práce zaměřena a jejichž výběr popisujeme v předchozí podkapitole (3.2), jsou součástí databáze sémantických rysů také další substantivní, verbální a adjektivní pojmy – tyto pojmy byly předmětem zájmu L. Večeřové a nejsou v rámci naší práce dále analyzovány.

Pro pilotní sběr dat bylo vybráno celkem 20 substantivních pojmů. Metoda sběru dat byla z větší části převzata ze vzorových zahraničních studií, tedy ze studie K. McRae a kol.¹⁰⁷ a D. P. Vinsona a G. Vigliocco¹⁰⁸. S ohledem na náš cíl, tj. získat data od dětských probandů jsme testovali také další způsoby sběru dat, abychom zjistili, který typ dětem vyhovuje nejlépe a zároveň nám přinese dostatek relevantních informací. Originální metoda předepisovala participantům sepisovat vlastnosti daného pojmu do sloupce pod sebe (podle předloženého vzoru). Na základě obavy z toho, že by příliš velké množství pokynů mohlo ovlivnit produktivnost dětí kvantitativně či kvalitativně, jsme v další, námi zvolené metodě nabídli dětem možnost psát odpovědi lineárně, tj. na řádcích, přičemž se mohly individuálně rozhodnout, zda budou jednotlivé informace oddělovat čárkou, nebo budou psát odpovědi v souvislých větách.

¹⁰⁶ Konečná, K. a Večeřová, L.: *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*, 2016.

¹⁰⁷ McRae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005.

¹⁰⁸ Vinson, D. P. a Vigliocco, G.: *Semantic feature production norms for a large set of objects and events*, 2008.

Ve třetí variantě jsme testovali možnosti metody ústní – jednalo se o dialog se sběračem dat. V této metodě jsme sbírali potřebná data zcela jiným způsobem především proto, abychom mohli porovnat kvalitu a kvantitu informací získaných ústně s kvalitou a kvantitou informací získaných písemně. Vycházeli jsme z domněnky, že pro participanty daného věku může být převedení vlastních myšlenek do písemné podoby stále poněkud náročné.

Bylo stanoveno několik kritérií k posouzení jednotlivých metod. Vedle celkové kvality odpovědí jsme se zaměřili na to, kolik rysů bylo v jednotlivých typech sběrů získáno a zda se děti při popisu zaměřily na sdělování různých typů informací. Sledovali jsme také obtížnost charakterizování jednotlivých pojmů a zjišťovali jsme, jak dlouhý časový úsek je potřebný ke splnění zadaného úkolu.

Získané odpovědi byly rozloženy na jednotlivé sémantické rysy na základě tzv. kódovacích pravidel. Tato pravidla byla předběžně připravena podle výše zmiňovaných vzorových studií a při analýze dat z pilotního sběru byla dodatečně doplněna na základě nedostatků objevených v průběhu této analýzy.

3.3.2 Metoda pilotního sběru dat

3.3.2.1 Participanti

Pilotního sběru dat se zúčastnilo celkem 39 žáků třetích tříd běžných českých základních škol, tj. děti ve věku 8–9 let. Více informací je uvedeno v tabulce 4.

Typ sběru dat	Celkový počet dětí	Počet dívek	Počet chlapců	Věk 8 let	Věk 9 let
Písemný (zápis ve sloupcích)	14	10	4	13	1
Písemný (volný styl zápisu)	19	7	12	13	6
Ústní	6	3	3	4	2
Celkem	39	20	19	30	9

Tab. 4. Přehled participantů zúčastněných v jednotlivých typech sběru

3.3.2.2 Materiály

V rámci pilotního sběru jsme otestovali celkem 20 substantivních pojmů z různých sémantických kategorií, přičemž devět z nich (zvýrazněny tučným písmem) je součástí naší databáze sémantických rysů: **borovice, cuketa, racek, rýč, svetr, trolejbus, varhany, višně a zub**, bidýlko, břeh, hranice, chrám, jinovatka, lepidlo, nestvůra, paruka, stádo, turista, lavice. Záměrně byly vybrány takové pojmy, které podle nás zastupují různou obtížnost při jejich popisu v souvislosti s různou mírou znalosti těchto pojmů u dětí mladšího školního věku. U všech testovaných pojmů byl zaznamenán jejich výskyt ve *Weslalexu*, a v korpusu českého jazyka *Syn verze 3*¹⁰⁹. Pro přehled pojmů a frekvenci jejich výskytu v této databázi a tomto korpusu viz následující tabulku 5.

¹⁰⁹ Křen, M. et al.: *Korpus SYN, verze 3*, 2014. Dostupné z: <https://kontext.korpus.cz>

pojmem	ARF SYN V3	Weslalex g3U (frekvence výskytu – 3. třída)	výskyt ve Weslalexu (školní třída)
BIDÝLKO	365,92	0	5
BŘEH	59220,60	141.36	1–5
HRANICE	259631,61	87.84	2–5
CHRÁM	29359,31	6.02	3, 4, 5
JINOVATKA	527,44	33.02	2, 3
LEPIDLO	3488,83	0	1
NESTVŮRA	661,47	0	2, 5
PARUKA	2714,86	5.52	3, 4, 5
STÁDO	10510,82	6.02	3, 5
TURISTA	116446,28	6.02	3
BOROVICE	5494,04	54.68	1–5
CUKETA	885,80	-	-
LAVICE	30982,31	50.91	1–5
RACEK	4412,89	5.52	1, 3, 4
RÝČ	956,72	24.10	1, 3, 4
SVETR	6267,35	5.52	1–5
TROLEJBUS	9044,78	0	4
VARHANY	10357,71	11.05	3, 4, 5
VIŠEŇ	487,20	0	4
ZUB	43896,92	109.70	1–5

Tab. 5. Přehled pojmů testovaných v pilotní studii

Každému participantovi bylo předloženo z výše uvedeného seznamu 10 pojmů, které měl popsat na základě položené otázky „Co to je, co to znamená?“. Participantů byli upozorněni, že cílem není popisovat pouze vizuální stránku pojmu, ale je vhodné zaměřit se také na charakterizaci spjatou s dalšími smyslovými vjemy (např. chuť), je možné uvést funkci, využití dané věci a cokoli dalšího, co jim přijde na mysl. Kompletní instrukce k této úloze (pro všechny tři testované typy sběru) jsou uvedeny v příloze 5. Pro lepší pochopení úkolu bylo uvedeno také několik příkladů.

Pojmy byly namíchány pomocí speciálního doplňku na náhodné promíchání buněk, který byl nainstalován do programu Microsoft Excel. Zkoumané pojmy byly rozděleny do dvou verzí sběrných sešitů. Praktická výhoda dvou variant spočívala v zamezení tomu, aby děti od sebe vzájemně opisovaly odpovědi – proto byl kladen důraz na to, aby byly verze při rozdávání sešitů prostřídány.

Participantům byly zkoumané pojmy představeny ústně nebo ve sběrných sešitech formátu A5. Úvodní strana obsahovala informace o participantovi a na dalších pěti stranách bylo připraveno 10 pojmů k popisu.

3.3.2.3 Postup při sběru dat

Před samotným sběrem dat byl prostřednictvím oslovených škol zaslán rodičům informační dopis s žádostí o účast v našem výzkumu. Pokud rodiče souhlasili, uvedli v příloženém

dotazníku základní informace o svém dítěti (tzn. jméno, věk, mateřský jazyk, znalost dalších jazyků).

Data, která byla sbírána písemnou formou (ve sloupcích i na řádcích), byla získána ve školách v rámci jedné vyučovací hodiny, přičemž všechny děti ze třídy mající souhlas rodičů se mohly zúčastnit souběžně. Sběrač dat přečetl účastníkům pokyny a uvedl písemně vzorový příklad (na tabuli). Ve zbytku vyučovací hodiny participanti individuálně zapisovali popisy pojmů do svých sběrných sešitů.

Sběr dat v rámci ústní metody probíhal odlišně od výše uvedených sběrů písemných. Osoba, která lektorovala sběr (sběrač dat), pracovala s jednotlivými participanty individuálně, pouze ústně. Sběrač dat přečetl pokyny a následně postupně jmenoval jednotlivé pojmy, zatímco participant ústně odpovídal, co tyto pojmy znamenají. Během všech těchto setkání byly pořizovány audiozáznamy a teprve následně byly odpovědi účastníků převedeny sběrači dat do písemné podoby.

3.3.2.4 Metoda skórování odpovědí a kódování získaných dat

Probandi mohli za každý pojem získat tři body, za celý úkol tedy maximálně 30 bodů. V případě, že dítě u pojmu nevedlo žádnou informaci, nebo uvedený popis neodpovídal skutečnému významu pojmu, získalo dítě 0 bodů. Za uznanou odpověď byl dítěti připsán 1 bod a další bod bylo možné získat za uvedení alespoň jednoho významného rysu, který pojem charakterizuje, případně diferencuje od jiných pojmů. Poslední, třetí bod byl přičten v případech, kdy byl popis posouzen autory práce jako velmi kvalitní (např. byly uvedeny další diferenční či významné rysy apod.).

Z individuálních popisů pojmů, které byly shromážděny, byla v dalším kroku vytvořena pilotní databáze sémantických rysů, a to podle stanovených tzv. kódovacích pravidel. Kódovací pravidla umožňují sjednotit jedinečně formulované odpovědi na základě jejich sdíleného významu a vytvořit systematický soubor sémantických dat.

Cílem bylo použít ustanovené kódy napříč celou databází systematicky tak, aby všechny stejné významy byly zahrnuty pod jediným kódem, a naopak významy odlišné měly svůj vlastní kód.

Odpovědi získané od participantů byly rozděleny na nejmenší významové jednotky, na tzv. sémantické rysy. Individuální formulace obsahující stejnou sémantickou informaci byly zaznamenány pomocí jedinečného kódu. Např. formulace „*sedá se na to*“, „*je to sedací věc*“, „*na sezení*“ apod. může být sjednocena a označena pomocí kódu *slouží k sezení* u všech pojmů, u kterých je tento sémantický rys uveden (v tomto případě se nejspíš bude jednat o pojmy označující sedací nábytek, tento rys ale může být zmíněn např. také u pojmů jako je **postel** či u adekvátních částí lidského těla¹¹⁰).

Vedle víceméně mechanického dělení odpovědí na samostatné části zahrnuje proces kódování také komplikovanější případy, kdy musí být ustanovena jednoznačná pravidla, která musí být

¹¹⁰ Přestože takové pojmy nemusí být v aktuální databázi sémantických rysů obsaženy, předpokládá se, že databáze může být kdykoliv rozšířena o nové kategorie a pojmy, přičemž je nezbytné, aby nové informace mohly být kódovány v souladu s daty stávajícími.

následně systematicky aplikována v celé databázi. Např. byla po vzoru K. Mcrae a kol.¹¹¹ rozložena informace typu *má to čtyři kola* na dva samostatné sémantické rysy, tj. *má to kola* a *má to čtyři kola*. Některá ustálená slovní spojení, jako je *je to hudební nástroj*, byla naopak chápána jako jeden rys (a proto v takových případech nedošlo ke sjednocení tohoto rysu s rysem *je to nástroj*, který se mohl objevit u pojmů v kategoriích, jako jsou „nářadí“ či „kuchyňské náčiní“). Vzhledem k tomu, že jsou kódovány sémantické informace, je v některých případech problematické stanovit přesnou hranici, kdy se jedná o stejný a kdy už o jiný rys. Důležité však je zachovat zvolený přístup konzistentní napříč celou databází.

Jednotnost systému kódování se vyznačuje také tím, že všechny kódy mají stejnou strukturu. Sémantické rysy jsou rozděleny na verbální a jmennou část. Například pro rysy týkající se funkce je užíváno verbum *slouží* a jmenná část se skládá z prepozice *k* a podstatného jména ve třetím pádě, tj. *slouží k převozu lidí*.

Při kódování dat byly v rámci možností sjednoceny morfologické formy – v kódech nebyla reflektována mluvnická kategorie rodu apod. Tak například pro vizuální rys označující barvu byl kód uveden vždy s maskulinní koncovkou *-ý* bez ohledu na gramatický rod pojmu, respektive slova (u pojmu **jablko** je proto uváděn rys *je červený*, nikoliv *je červené* atd.).

Při kódování dat dochází k záměrnému zjednodušení odpovědí (popisů), což vede k vyloučení některých výrazů, jako například těch, které vyjadřují frekvenci („někdy“, „obvykle“, „může být“ aj.). Významnost rysu uváděného pro pojem je dána samotnou frekvencí produkce rysu: pokud děti například u jablka téměř vždy uvádějí, že je červené, a jen někdy (někteří) při jeho popisu zmiňují zelenou barvu, odráží se v početnosti těchto rysů i míra jejich typičnosti v jednotlivých rysových reprezentacích.

V databázi jsou zařazeny všechny odpovědi bez ohledu na jejich kvalitu či pravdivost. Případné chybné popisy či neobvyklé charakteristiky se v datech rozptýlí a jejich nízká frekvence svědčí o jejich bezvýznamnosti v reprezentaci vybraných pojmů.

Pravidla byla předběžně ustanovena již před zahájením pilotního sběru dat. V průběhu zpracování dat bylo zaznamenáno několik případů, které vedly k diskusi a k nastavení dodatečných pravidel kódování.

Vzhledem k tomu, že pilotní sběr a zpracování dat bylo uskutečněno pro dva různé výzkumy zároveň a naše cíle byly částečně odlišné, byla kódovací pravidla utvořena tak, aby vyhovovala záměrům obou autorek. Více detailů ke kódování dat v hlavní databázi uvádíme v podkapitole 3.4.2.5 (*Metoda kódování dat*), která je součástí popisu metody hlavního sběru dat.

3.3.3 Statistické zpracování dat a výsledky pilotního výzkumu

Data získaná v pilotním výzkumu byla zpracována vhodnými statistickými metodami. Na základě výsledků těchto zpracování byla vybrána nejvhodnější metoda sběru dat.

Posuzována byla délka plnění a náročnost úkolu, kvalita odpovědí, počet získaných rysů a rovnoměrnost rozložení různých typů rysů. Jako nejvhodnější varianta sběru dat byla pak vybrána ta, která nejlépe splňovala tyto stanovené podmínky – tzn. bylo v ní u jednotlivých

¹¹¹ McRae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 551.

pozorovaných kritérií dosaženo ideálních výsledků. Zároveň by se výsledky hodnocení jednotlivých kritérií neměly příliš odchylovat od průměru.

3.3.3.1 Průměrná doba zpracování úlohy a úspěšnost participantů pro jednotlivé varianty sběru dat

Průměrná doba zpracování úkolu v písemné variantě, ve které participanté zaznamenávali odpovědi do sloupců, byla 26,5 minut. V případě druhé písemné varianty nebyl průměrný čas zaznamenán, nicméně poslední sběrný sešit byl odevzdán po 45 minutách. V ústní metodě trval sběr dat průměrně 3,36 minut. Je však nutné připomenout, že při písemném sběru byla data získávána od většího množství probandů zároveň, zatímco v ústním sběru byla data získávána individuálně, což znamenalo, že pokyny musely být opakovány každému účastníkovi zvlášť. Je zřejmé, že pro participanty je ústní varianta sběru méně náročná (přinejmenším časově). Z hlediska práce sběračů dat se však ústní sběr dat nejeví jako vhodná varianta, především s ohledem na množství dat, které je potřeba získat, aby byl vzorek dostačující, a také s ohledem na množství času potřebného pro následný přepis odpovědí.

Úspěšnost z hlediska uznaných a neuznaných odpovědí (z celkového množství) byla nejvyšší v případě ústní metody. V této metodě bylo dosaženo 86,67% úspěšnosti. V písemných variantách sběru bylo dosaženo lepších výsledků v metodě „řádkové“ (77,37 %), v metodě „sloupcové“ to bylo pouze 62,14 %.

Z hlediska kvality odpovědí byly jednotlivé metody sběru úspěšné ve stejném pořadí: nejlepších výsledků bylo dosaženo v metodě ústní (57,22 %), pak v metodě „řádkové“ (39,65 %) a nakonec v metodě „sloupcové“ (33,81 %).

Na základě tohoto hodnocení je v písemném sběru dosaženo lepších výsledků v metodě s volnějším zadáním, tedy v té, ve které děti mohou odpovědi zaznamenávat na řádcích, jednotlivě, či ve větách. Nejlepších výsledků bylo dosaženo v ústní variantě sběru dat, nicméně při komplexním posuzování je nutné přihlídnout i k výše uvedeným nevýhodám, které tato metoda přináší.

3.3.3.2 Hodnocení kvality jednotlivých typů sběrů dat

Kvalita jednotlivých typů sběru dat byla posouzena za pomoci statistických metod, které popisujeme na následujících řádcích.

Kompatibilita mezi třemi vybranými kritérii (tzn. kompatibilita mezi kvantitou uváděných rysů, kvalitou odpovědí a různorodostí uváděných rysů) byla posouzena pomocí výpočtu korelačního koeficientu. Tento koeficient ukázal, že nejvyšší korelace se objevují u metody, kdy děti zapisovaly informace do sloupců (viz tabulku 6). Významná korelace byla zjištěna i u metody zápisu na řádcích. Významné nedostatky týkající se závislostí jednotlivých kritérií byly zjištěny v případě metody ústní. To je pravděpodobně způsobeno nedostatečnou různorodostí rysů získaných v tomto typu sběru – další výpočty tuto domněnku potvrzují (viz dále).

Metoda zápisu dat ve sloupcích			
	KVANTITA RYSŮ	KVALITA RYSŮ	TYPY RYSŮ
KVANTITA RYSŮ	1	-	-
KVALITA RYSŮ	0,8920	1	-
TYPY RYSŮ	0,8538	0,8279	1
Metoda zápisu dat na řádcích			
	KVANTITA RYSŮ	KVALITA RYSŮ	TYPY RYSŮ
KVANTITA RYSŮ	1	-	-
KVALITA RYSŮ	0,7483	1	-
TYPY RYSŮ	0,8341	0,7334	1
Metoda ústního sběru dat			
	KVANTITA RYSŮ	KVALITA RYSŮ	TYPY RYSŮ
KVANTITA RYSŮ	1	-	-
KVALITA RYSŮ	0,6696	1	-
TYPY RYSŮ	0,7782	0,4396	1

Tab. 6. Kompatibilita sledovaných kritérií v jednotlivých typech sběru

Shoda rozptylů výsledků získaných v odlišných metodách sběru či případné rozdíly mezi středními hodnotami v těchto sběrech byly zjišťovány za pomoci jednoduché analýzy rozptylu (ANOVA). Data byla posuzována podle participantů – posuzování z hlediska pojmů nebylo možné, jelikož pro analýzu rozptylu nebyly v tomto případě splněny podmínky. Pomocí této analýzy bylo potvrzeno, že mezi středními hodnotami existují rozdíly, a to mezi ústní a písemnou sloupcovou variantou a mezi ústní a písemnou řádkovou variantou; mezi písemnými variantami sběru rozdíly nebyly zaznamenány.

Co se týče množství získaných rysů, nebyly při srovnání zvolených metod nalezeny žádné významné rozdíly. Vybraná metoda sběru dat nicméně měla vliv na kvalitu odpovědí – odpovědi, které byly vyhodnoceny jako nejkvalitnější, byly získány při ústním sběru dat. Vyhodnotit vliv výběru typu metody na rozmanitost získaných typů rysů nebylo možné, protože v tomto případě opět nebyly pro analýzu rozptylu splněny podmínky.

Dále byly jednotlivé metody posuzovány celistvě pomocí váženého průměru – byla brána v potaz jednotlivá výše zmíněná kritéria (tj. kvalita, kvantita a typologická různorodost rysů), přičemž pro ně byla nastavena pomocí vah různá důležitost. V jednom případě byla vyšší váha přiřčena kvalitě rysů a jejich typologické různorodosti, ve druhém případě byla pro všechna kritéria nastavena stejná váha. V obou případech dosáhla nejlepších výsledků metoda písemného sběru na řádcích.

Aby nebyla výsledná data zkreslena nadprůměrnými a podprůměrnými výkony jednotlivých žáků v rámci jednotlivých kritérií, byla vedle průměru zohledněna také směrodatná odchylka dat, a to ve výpočtu variačního koeficientu, tedy ve výpočtu podílu směrodatné odchylky a střední hodnoty. Z výsledků vyplývá, že největší odchýlení od průměru přináší metoda zápisu odpovědí do sloupců, nejmenší byla zaznamenána u ústního sběru dat. Pro naše účely je důležité preferovat tu variantu, která se jeví jako nejvíce stabilní a ve které je variabilita dat nižší, tj. nedochází v ní k velkému odchýlení od průměrných hodnot způsobenému nadprůměrnými či podprůměrnými výkony jednotlivců. Z tohoto hlediska je považována za nejvhodnější varianta ústní. Podobně nízká variabilita dat však byla zaznamenána i v případě varianty písemné se zápisem odpovědí na řádcích (variabilitě dat u této písemné metody bylo jen o 1,55 % vyšší než u metody ústní).

Poslední výpočet, který nám pomohl v rozhodování při výběru mezi jednotlivými typy sběru, byl proveden za pomoci tzv. Hurwitzova pravidla, ve kterém se pracuje s maximálními a minimálními možnými hodnotami a s posuzovatelem volenou mírou optimismu a pesimismu. Pomocí tohoto měření je tedy možné vyloučit metodu, v níž je v jednotlivých zvolených kritériích dosaženo nedostatečných výsledků, současně je ale zjišťováno, která metoda přináší výsledky nejlepší. V tomto případě nejlépe jednoznačně nejlépe obstála varianta písemná se zápisem na řádcích.

3.3.3.3 Ke statistickému zpracování databáze sémantických rysů (v kontextu pilotní studie)
V rámci pilotního výzkumu byly rovněž otestovány statistické metody, pomocí kterých jsou zpracovávána data v našich vzorových studiích. Díky těmto statistickým výpočtům lze získat různorodé informace o kategoriích, pojmech, rysech, a především o vztazích mezi nimi.

Byla zjišťována průměrná produkce rysů pro individuální pojmy v jednotlivých metodách sběru. Tato průměrná produkce rysů byla vypočítána pomocí aritmetického průměru: soubor všech realizací pro daný pojem byl vydělen počtem dětí popisujících tento pojem. Zdůrazňujeme, že šlo o realizace odpovědí – to znamená, že byla započítána každá jedna odpověď, nešlo tedy o pouhý součet odlišných sémantických informací. Výsledky tohoto výpočtu napříč všemi metodami sběru dat ukazují, že průměrný počet rysů na participanta byl nejvyšší u pojmu **bidýlko**, **nestvůra** a **zub** (3,8 rysu, 4,5 rysu a 4,4 rysu). Naopak nejméně rysů bylo uváděno u pojmů **hranice** a **varhany** (průměrně 1,8 a 1,4 rysu).

Vzhledem k tomu, že je pilotní výzkum zaměřen především na hledání nejvhodnější metody sběru dat, zaměřili jsme se také na průměrný počet uváděných informací v jednotlivých metodách sběru. Větší množství uváděných rysů u některé z testovaných metod může totiž svědčit o nižší obtížnosti úkolu při plnění touto metodou, případně může znamenat větší ochotu plnit úkol určitým způsobem. Bylo zjištěno, že participanti k pojmům uváděli průměrně tři informace – ve variantě psaní na řádcích a v ústní metodě to konkrétně bylo 3,1 rysů, zatímco v písemném sběru ve sloupcích to bylo jen 2,8 rysu.

Obecně můžeme uvažovat nad tím, že by průměrný počet rysů mohl naznačovat míru osvojení významů jednotlivých pojmů u dané věkové kategorie, případně může souviset se šíří významu pojmu (například **nestvůra** může mít mnoho podob a odpovědi dětí se mohou méně krýt na základě odlišné individuální zkušenosti) či s pocitem nutnosti uvést více rysů k dosažení přesnější definice u méně jasných pojmů.

V pilotním sběru nebylo získáno velké množství dat, proto nepovažujeme za nutné v této části práce data interpretovat a ani uvádět výsledky dalších statistických zpracování. Tyto pilotní výpočty jsou však dostupné v již zmiňované publikaci *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*¹¹², ve které je detailně popsán průběh celé pilotní studie.

Statistiky tohoto typu jsou pak v této práci popsány dále, a to v kapitolách, které popisují statistické zpracování dat z hlavní databáze a výsledky tohoto zpracování.

¹¹² Konečná, K. a Večeřová, L.: *Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat*, 2016.

3.3.4 Vyhodnocení pilotní studie – výběr vhodné metody sběru dat

Hlavním cílem pilotní studie byl výběr nejvhodnější metody sběru dat, tj. metody popisu pojmů pro získání sémantických rysů do databáze. Standardně používaná písemná metoda, ve které participanti sepisují jednotlivé vlastnosti pojmu pod sebe (do sloupce), byla konfrontována s dalšími dvěma metodami, a to s písemným zápisem odpovědí na řádcích a s ústní metodou sběru. Rozhodnutí otestovat více možností, jak data získat, vycházelo z ohledu na věk participantů – bylo potřeba zjistit, která metoda bude dětem nižšího školního věku nejlépe vyhovovat a pomocí které my můžeme získat dostatečně kvalitní popisy pojmů obsahující rozmanité typy informací.

Na základě několika provedených hodnocení jednotlivých typů sběrů pomocí výše představených statistických metod jsme došli k závěru, že nejvhodnější způsob, jak získat potřebná data o pojmech od dětí, je zadat dětem úkol v písemné podobě a zároveň jim dopřát určitou volnost ve způsobu zápisu. Rozhodli jsme se tedy pro variantu písemného sběru se zápisem na řádcích, ve které nejsou děti omezeny předem nastavenými přísnými instrukcemi, jak zapisovat odpovědi, a mohou dle libosti uvádět buď celé věty, nebo vypisovat jednotlivé informace oddělené čárkou apod.

Ze zpracování získaného materiálu vyplývá, že uspokojivé výsledky byly získány také ústní metodou. Už na začátku podkapitoly 3.3.3, věnující se statistickému zpracování dat, jsme však předešleli myšlenku, že sběr dat touto metodou je časově náročný a zpracování takto získaných dat je velice pracné. Vzhledem k předpokládanému množství dat, které je pro databázi potřeba získat a zpracovat, jsme se rozhodli pro zamítnutí této metody, a to ve chvíli, kdy bylo na základě výsledků pilotního sběru zřejmé, že písemná řádková metoda je vhodná pro dětské participanty a zároveň dostačující pro cíl našeho výzkumu.

3.4 Hlavní sběr dat

Podobně jako v případě pilotního sběru dat byl i hlavní sběr dat uskutečněn ve spolupráci s L. Večeřovou. Sběr popisujeme na následujících řádcích tak, jak probíhal ve spolupráci s kolegyní. V samotné databázi jsou však zpracována pouze ta získaná data, která jsou podstatná pro autorku této práce, proto zde uvádíme zpracování a interpretaci pouze tohoto vybraného materiálu. Tato část práce byla publikována ve studii *A database of semantic features for chosen concepts (Attested in 8- to 10-year-old Czech pupils)*¹¹³.

3.4.1 Cíl hlavního sběru dat

V hlavním sběru dat byla získána data pro databázi sémantických rysů pro 104 objektových pojmů z 13 různých sémantických kategorií od 381 dětí, jejichž mateřský jazyk je čeština. Podstata zadání úkolu zněla „Co to znamená? Popiš dané slovo/pojem.“ Úkolem dětí bylo tedy sepisovat charakteristické znaky daného pojmu. Pomocí metody kódování odpovědí pak byla z těchto odpovědí vytvořena databáze sémantických rysů.

¹¹³ Konečná, K. a Vaňkátová, K.: *A database of semantic features for chosen concepts (Attested in 8- to 10-year-old Czech pupils)*, 2017.

Ze získaných dat si lze utvořit představu o tzv. rysových reprezentacích vybraných pojmů. Pomocí statistických metod byla data z databáze analyzována a následně byla interpretována. Databáze a základní statistické výpočty jsou dostupné online na adrese www.childrenfeatures.webnode.cz.

Tento materiál lze využít v dalších psycholingvistických experimentech. V našem případě byla v návazném sémantickém experimentu využita rysová podobnost pojmů vypočtená ze získaných dat. V tomto experimentu jsme se pokusili ověřit roli sémantických rysů v rychlé sémantické úloze (viz kapitolu 3.5).

3.4.2 Metoda hlavního sběru dat

3.4.2.1 Participanti

Pro tvorbu databáze byla získána data celkem od 381 českých dětí navštěvujících české základní školy (data byla sbírána konkrétně na Moravě a ve Slezsku, tj. v Olomouci, v Opavě, v Ostravě, v Krnově, v Novém Jičíně a ve Velké Bystřici). Jednalo se o děti třetích ročníků nižšího stupně, tedy o věkovou kategorii 8–9 (výjimečně 10) let.

V centru našeho zájmu byla věková skupina osmi a devítiletých dětí, během sběru dat jsme však získali data také od několika desetiletých dětí, které chodí do třetí třídy společně s mladšími dětmi. Data od těchto starších dětí jsou rovněž součástí databáze – předpokládáme, že děti navštěvující jednu třídu jsou mentálně a znalostně přibližně na stejné úrovni, bez ohledu na jejich věk. Navíc databáze obsahuje pouze malé množství dat získaných od starších dětí a u všech dat je uvedena informace o věku participanta.

Zdůrazněme, že data byla sbírána pouze v moravskoslezské části České republiky. Domníváme se, že na sémantickou stránku databáze nebude mít tato skutečnost významný vliv a rozdíly se mohou objevit spíše pouze při zkoumání konkrétních reálných odpovědí participantů, protože děti nemusejí umět posoudit spisovnost a správnost výrazů a mohou při popisu pojmů užívat dialektické výrazy typické pouze pro danou oblast, ve které žijí. V každém případě se zde nabízí možnost rozšířit databázi o data z dalších oblastí České republiky.

V databázi rovněž uvádíme všechny mateřské jazyky dítěte, které rodiče uvedli ve vstupním dotazníku – lze zde tedy snadno dohledat například informaci o tom, zda je participant bilingvní.

V databázi neuvádíme jména a příjmení jednotlivých dětí. Databáze je anonymní – každý participant má svůj kód, jehož součástí jsou údaje o pohlaví, věku, verzi sběrného sešitu a číslo participanta (jedno z 381 možných).

3.4.2.2 Výběr pojmů do databáze – shrnutí

Na následujících řádcích shrnujeme zásadní informace o výběru pojmů do databáze. Zevrubný popis obsahuje samostatná kapitola *Výběr kategorií a pojmů do databáze sémantických rysů*, kterou uvádíme na začátku praktické části práce.

Pro databázi bylo vybráno celkem 104 substantivních pojmů ze 13 sémantických kategorií. Jedná se především o pojmy/slova spadající do základní slovní zásoby. Kategorie byly rozděleny na dva typy podle toho, zda se jedná o věci (části věcí), které se vyskytují volně

v přírodě nezávisle na člověku (tyto kategorie nazýváme biologické kategorie), nebo se jedná o věci, které vznikly zásluhou člověka (tzv. artefakty).

Celá struktura sémantického rozboru pak vypadá následovně: biologická kategorie/artefakt → sémantická kategorie → pojem/slovo¹¹⁴ → sémantické rysy → typy sémantických rysů¹¹⁵.

Pro každou kategorii bylo zvoleno 8 zastupujících pojmů s různou mírou (proto)typičnosti. Tyto pojmy nebyly zvoleny náhodně, nýbrž na základě dotazníkového průzkumu, kterého se zúčastnilo 40 dospělých respondentů.

Ze získaného seznamu typických zástupců pro každou kategorii jsme dále vybírali pojmy s ohledem na další výzkumný cíl tak, aby bylo možné je spárovat na základě podobnosti – v rámci každé kategorie byly vybrány čtyři páry pojmů s různou mírou podobnosti, od nejvyšší po nejmenší nebo žádnou podobnost. Tato míra podobnosti byla v této fázi práce pouze předpokládána a skutečná rysová podobnost byla vypočítána posléze ze získaných dat.

Pro názornější představu uvedme připravené pojmové páry z kategorie savců: pojmy **zajíc** a **králík** byly vybrány jako vzájemně nejvíc sémanticky podobné; pojmy **kočka** a **tygr** tvořily pár s vysokou mírou podobnosti; pojmy **jelen** a **kůň** tvořily pár vykazující nějakou podobnost, avšak poměrně nízkou; nakonec pár s nejnižší (případně žádnou) podobností tvořily pojmy **delfín** a **opice**.

Vybrané pojmy byly rovněž označeny informací o jejich frekvenci výskytu v korpusu SYN verze 3 a také četnosti výskytu ve *Weslalexu*.

Slova byla vybírána (intuitivně a také na základě jejich četnosti výskytu ve zmíněném korpusu a databázi) tak, abychom redukovali situace, kdy participanti nebudou znát význam daného slova a my proto nebudeme moci od nich získat žádná data. Přesto je nutné dodat, že u některých pojmů nakonec vyplynulo, že jsou zřejmě na deskripci významu náročnější než jiné – například uvést význam pro pojem **revolver** zvládlo pouze 16 dětí ze 40, kterým byl pojem k popisu předložen. V následující tabulce 7 uvádíme také několik dalších příkladů pojmů a úspěšnosti participantů v jejich popisu.

	PRAK	REVOLVER	KŮŇ	CIBULE	NOS	LŽÍCE
celkový počet dětí, kterým byl pojem předložen k popisu	24	40	20	21	23	25
uznané popisy	20	16	20	20	20	24
neuznané popisy	3 ¹¹⁶	0	0	0	0	1 ¹¹⁷
nevyplněno	1	24	0	1	3	0

Tab. 7. Celkový počet získaných popisů pojmů a počet vyřazených položek (pro vybrané pojmy)

Přestože jsme se při výběru pojmů snažili respektovat pravidla „bazálnosti“ a jednoduchosti, vyskytuje se v seznamu několik slov, u kterých nelze jednoznačně mluvit o nevhodném výběru, ale považujeme za nutné dodat k nim další komentář vycházející z našich pozdějších poznatků.

¹¹⁴ K problematice „pojem versus slovo“ se stručně vyjadřujeme v teoretické části práce.

¹¹⁵ Rysy jsou rozděleny podle toho, zda popisují nějakou funkci, vzhled, chuť apod. Více viz dále.

¹¹⁶ Pojem **prak** byl několika dětmi zaměněn s pojmem **park**.

¹¹⁷ Pojem **lžíce** byl jedním dítětem zaměněn s pojmem **lyže**.

Součástí kategorie „oblečení“ je pojem **boty**, který uvádíme v množném čísle na základě párového charakteru tohoto objektu. Vzhledem k existenci jednotného čísla by bylo možné uvést jej, podobně jako tomu bylo například u pojmů/slov **ruka** či **noha**, v jednotném čísle – u dvou posledně zmiňovaných pojmů participantů uváděli (i když v malé míře) *rys je párový*, přičemž tento rys nebyl zmíněn u pojmu boty ani jednou – možná právě kvůli tomu, že jsme jej uváděli v plurálu.

V databázi se rovněž vyskytuje několik pojmů/slov, které se řadí mezi jména pomnožná (například **hrábě** či **varhany**), nezdá se však, že by tato jejich vlastnost měla jakýkoliv vliv na jejich popis.

Z tohoto hlediska se však jako komplikovaný jeví pojem **šaty** – předpokládali jsme, že participantů budou popisovat dámské oblečení, které toto slovo označuje, a nepřipustili jsme možnost, že by se některé děti mohly zaměřit na popis pojmu ve smyslu oblečení jako takového. V případě dámského oblečení se jedná o slovo pomnožné, zatímco v obecnějším smyslu tohoto slova existuje také jednotné číslo (*šat*), byť je dnes chápáno jako knižní.

Další nejednoznačnost vyvstala u pojmů **třešeň** a **višeň**. Někteří participantů uváděli kromě popisu plodu, kterýžto jsme předpokládali, také popis stromu, na kterém tyto plody rostou. Teprve při zpracování databáze jsme se v jazykové příručce¹¹⁸ seznámili se skutečností, že pojem **třešeň/višeň** je zpravidla užíván ve smyslu celé rostliny a pro plod se i v jednotném čísle primárně používá výraz **třešně/višně**. Avšak vzhledem k nejednoznačnosti vyplývající z definice uvedené v příručce, možné regionální podmíněnosti použití těchto výrazů a také vzhledem k tomu, že část participantů skutečně popisovala pojem tak, jak jsme očekávali, zůstaly tyto pojmy součástí databáze. Seznam sémantických rysů patřících danému pojmu se však rozšířil o další význam.

3.4.2.3 Materiály

Pro každého participanta byl připraven speciální osmistránkový sešit formátu A5, který obsahoval 10 pojmů. Na úvodní straně byl štítek s osobními údaji a údaji o sběru, včetně místa sběru, tj. školy. Byly zde také uvedeny příklady jednotlivých typů pojmů a možných popisů jejich významů (**citrón** jako zástupce biologických kategorií, **nůž** jako zástupce artefaktů, **klepat** jako zástupce verb a **průhledný** jako zástupce adjektiv). Příklady sémantických rysů byly různého typu (vzhled, chuť, vůně, funkce aj.), abychom participanty motivovali k uvádění rozmanitých informací při popisu. Na každé následující straně byly uvedeny dva pojmy a pro každý pojem bylo k dispozici několik volných řádků, na které participantů mohli zapisovat své odpovědi. Na základě pilotního sběru dat bylo rozhodnuto ponechat dětem volnost při rozhodování způsobu zápisu informací. V příkladech popisů pojmů předkládaných na úvodní straně byly sémantické rysy zapisovány na řádcích za sebou a byly oddělovány čárkami. Jakýkoliv jiný způsob zápisu rysů – např. pod sebou ve sloupcích či formou celých vět a souvětí – byl plně tolerován.

Výběr pojmů do jednotlivých verzí sešitů byl náhodný, proběhl pomocí speciálního nástroje na míchání buněk v Microsoftu Excelu. Při míchání bylo respektováno pravidlo, že se v každém sešitě smí vyskytovat maximálně jeden pojem z každé kategorie. Vzniklo tak několik verzí

¹¹⁸ *Internetová jazyková příručka* [online], 2008–2021. Dostupné z: <https://prirucka.ujc.cas.cz/>

sešitu, které byly dále kopírovány k dosažení celkového potřebného množství. Naším cílem bylo získat sémantické rysy pro každý pojem přibližně od 20 dětí. Po zpracování části získaných dat bylo proto ověřeno, které pojmy už byly do té chvíle popsány dostatečným množstvím probandů. Pokud participanti pojem nedokázali nebo nestihli v časovém limitu popsat, vzali jsme tuto informaci v potaz – pojmy, které nebyly vyplněny, byly opětovně namíchány a rozdány mezi participanty v dalším kole sběru dat. Proces jsme opakovali, dokud nebylo získáno potřebné množství popisů. Naším cílem však nebylo ověřování znalostí pojmů. Součástí databáze jsou proto i takové sémantické rysy, které nemusíme považovat za správné, ale jsou součástí reprezentace daného pojmu minimálně u jednoho z participantů (například v kategorii „**ptáci**“ byl u pojmů **holub**, **racek** a **orel** uveden rys *je savec*). Do databáze nebyly zahrnuty pouze takové popisy, u kterých bylo zjevné, že došlo při čtení pojmu/slova k záměně s jiným slovem.

3.4.2.4 Postup při sběru dat

Vzhledem k tomu, že data byla sbírána na základních školách, byli nejprve osloveni a požádáni o spolupráci ředitelé těchto škol – byl jim doručen informační dopis pro ředitele (viz přílohu 2). Pokud vedení školy účast ve výzkumu schválilo, byly dětem ve vybraných třídách předány informační dopisy pro rodiče, ve kterých kromě základních informací o svém dítěti uvedli rodiče rovněž písemný souhlas s účastí ve výzkumu (viz přílohu 3).

Děti, které přinesly souhlas rodičů s účastí, vyplnily s pomocí sběračů dat na začátku vyučovací hodiny první stranu svého sběrného sešitu s osobními údaji (tj. jméno, příjmení a věk). Zbývající údaje – týkající se místa, času a formy sběru – vyplnili sběrači a autoři výzkumu dodatečně. Dětem byly přečteny pokyny (viz přílohu 6) a byly seznámeny s příklady popisů uvedenými na první straně sešitu. Následně byly v rámci přípravy požádány o spolupráci při ústním popisu dalšího pojmu, a to pojmu **slunce**.

Ve zbývajícím čase téže vyučovací hodiny, ve které byla představena pravidla úkolu, popisovali dětští participanti samostatně deset vybraných pojmů, které obsahoval jejich sešit. Vzhledem k tomu, že bylo připraveno větší množství variant sběrných sešitů, byly dětem, které seděly v jedné lavici, předloženy dvě různé verze, což snížilo riziko, že budou participanti odpovědi navzájem opisovat.

Děti byly k této práci motivovány odměnou v podobě výtvarných potřeb (pastelek), kterou dostaly v závěru vyučovací hodiny, tj. po splnění úkolu.

V další fázi sběrači dat přepsali získaná data do tabulky v programu Excel a předběžně rozdělili odpovědi dětí na jednotlivé významové jednotky.

3.4.2.5 Metoda kódování dat

Základní informace o způsobu kódování získaných dat jsou uvedeny také v podkapitole *Metoda skórování odpovědí a kódování získaných dat*, která je součástí kapitoly o pilotní studii. Na tomto místě shrneme nejdůležitější pravidla a věnujeme se několika komplikovanějším případům.

Stěžejní pravidla zpracování získaných dat vycházejí z našich vzorových studií K. McRaeho a kol.¹¹⁹ a D. P. Vinsona a G. Vigliocco¹²⁰. Postupně jsme v průběhu pilotního a hlavního sběru dat ustanovili také několik dalších vlastních pravidel. Naším cílem bylo především vytvořit takový systém, který bude přehledný a bude jednotný napříč celou databází tak, aby stejné jednotky významu byly zakódovány pod jediným kódem, a naopak každý rozdílný význam měl svůj vlastní kód. Je zřejmé, že dělení získaných informací na nejmenší významové jednotky i jejich kódování může být v určité míře subjektivní, a proto v této části práce popisujeme jednotlivá pravidla, podle kterých jsme data rozdělovali (tzv. kodovali), ale také všechna možná rizika a nevýhody zpracovávání dat takovýmto způsobem.

Data získaná od dětí – tzn. popisy významů pojmů, které dětští participanti zapisovali – byla za pomoci sběračů dat přepsána do počítače. Popis každého pojmu v přesné podobě tak, jak jej uvedli jednotliví participanti¹²¹, byl zaznamenán do tabulek v programu MS Excel a následně byl rozložen na jednotlivé sémantické informace, tedy na sémantické rysy.

Každá varianta odpovědi se stejným významem pak byla zapsána pod jediným kódem. Jednotliví participanti například mohli uvést, že nějaké zvíře (tj. k rozboru předkládaný pojem může v takovém případě být např. **kůň**) „*můžeme najít na farmě*“, „*je na farmě*“, „*bývá na farmě*“, „*žije na farmě*“ apod., přičemž význam je přes rozdílnou formu zápisu stále stejný, a proto jsou všechny tyto odpovědi zahrnuty pod jediný kód *je na farmě*. Tvoření těchto kódů však nebylo zcela nahodilé. Zpravidla bylo vybráno jednoduché, jednoznačné slovní spojení, které nese daný význam.

Pro přehlednost jsme navíc v našem případě dělili jednotlivé kódy na tzv. jmennou a verbální část: oddělili jsme vybrané verbum od zbytku formulace, takže v tzv. verbální části odpovědi bylo například jen verbum *je* nebo *má* a v druhé, jmenné části odpovědi bylo uvedeno *na farmě*, *nohy* apod. Toto rozdělení se podílí na udržení jednotnosti a přehlednosti celé databáze.

Ze stejných důvodů nejsou rozlišovány jednotlivé gramatické formy. V kódech tedy nebyla reflektována mluvnická kategorie rodu apod. Tak například kód pro vizuální rys označující barvu je zpravidla popisován pomocí adjektiva – to bylo uvedeno vždy s maskulinní koncovkou -ý bez ohledu na rod pojmu, respektive slova (například u pojmu **mrkev** byly uváděny rysy jako *je oranžový* apod.). V rámci verbální části byla v rámci možností udržována forma třetí osoby singuláru, přítomný čas a oznamovací způsob.

Ani kvantifikátory typu „někdy“, „často“, „může být“ apod. nejsou součástí kódů. Výskyt a typičnost daného rysu pro pojem je dána produkční frekvencí rysu. Tato frekvence byla vypočítána z informace, kolik dětí z celkového počtu dětí popisujících pojem uvedlo tento rys. Pokud tedy uvede 18 participantů z 20 celkem, že **holub má šedou barvu**, pouze dvě děti zmíní, že *může být bílý*, a jen jedno dítě napíše, že *je někdy černý*, je zřejmé, že v případě **holuba** lze vnímat jako nejtypičtější barvu právě šedou.

¹¹⁹ McRae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005.

¹²⁰ Vinson, D. P. a Vigliocco, G.: *Semantic feature production norms for a large set of objects and events*, 2008.

¹²¹ Příklady takových kompletních popisů jsou uvedeny v příloze 8.

V databázi se však vyskytují také komplikovanější případy, pro které musí být ustanovena pravidla a ta musí být následně systematicky aplikována v celé databázi. Již jsme uváděli příklad, že např. informace typu „*má to čtyři kola*“ je rozdělena na dva sémantické rysy, tedy na *má to kola* a *má to čtyři kola*. Inspiraci pro tento způsob dělení jsme našli u K. Mcrae a kol.¹²². také jsme již zmiňovali, že ustálená slovní spojení typu *je to hudební nástroj* byla chápána jako jeden rys.

Je zřejmé, že je často obtížné objektivně rozhodnout, kde začíná a končí hranice „nejmenší sémantické jednotky“, a to i přes ustanovená pravidla. Problémy nastávají především tehdy, když se musíme rozhodnout, zda jsou dvě informace uváděné dětmi skutečně shodné a lze je zapsat pod jeden kód, nebo zda je nutné chápat je jako dvě informace samostatné. Například z výše uvedené realizované odpovědi probanda „*žije na farmě*“ jsme se rozhodli vyčlenit informaci o živém stavu pojmu pod samostatný kód. Z informace, že „*žije na farmě*“ jsme tedy získali dvě sémantické jednotky: *je na farmě* a *je živý*. K tomuto rozhodnutí nás vedl fakt, že participanti mohou uvádět informaci „*je živý*“ nebo „*žije*“ jako samostatnou skutečnost. Proto se domníváme, že ve spojení *žije na farmě* je minimálně podvědomě začleněna informace o životnosti daného pojmu, ačkoliv je pravděpodobné, že dítě se prostě jen rozhodlo upřednostnit verbum „*žít*“ před verbem „*být*“.

Problematické je také stanovení hranice při určování míry podobnosti dvou informací. Lze chápat sémantický rys *je velký* jednotně, přestože je participanty uváděn nejen u pojmu **trolejbus**, ale také u pojmu **cuketa**? Ačkoliv je zřejmá relativita významu tohoto rysu, zdá se, že není možné určit taková kritéria, která by nám umožnila tento rys rozdělit na více typů. Musíme se tedy spokojit s jedním kódem pro oba tyto případy.

Podobně komplikovaná je situace např. s barvami. Například červená barva je dětmi uváděna u pojmu **višeň**, ale také u pojmu **trolejbus**, přestože se ve skutečnosti nemusí jednat o totožnou barvu (může to být dáno tím, že v daném jazyce pro takový rozdíl mezi barvami vůbec neexistují dva různé výrazy, nebo participanti tento rozdíl při vypisování atributů ignorují, přestože o něm vědí).

Naopak informaci *je dobrý* jsme se rozhodli rozdělit na dva různé rysy. V případě, kdy děti uváděly „*je dobrý*“ u jedlých objektů, kódovali jsme tuto informaci jako *je chutný*. Vycházeli jsme přitom z vlastního přesvědčení, že dítě v těchto případech mluvilo o chuti. V ostatních případech jsme ponechali pro tuto informaci kód *je dobrý*.

K podobným komplikacím se členěním odpovědí docházelo zákonitě poměrně často. Zpracování rozmanitého významu do zjednodušených systematických kódů, které lze zaznamenat do přehledných tabulek, s sebou přináší určitou redukci. Je nutné počítat také s určitou subjektivitou ze strany zpracovatelů databáze. Z těchto důvodů byly nejasné a komplikované případy dále diskutovány s lingvisty. Část dat byla rovněž zpracována více zpracovateli, aby byla zajištěna alespoň částečná objektivita (došlo ke srovnání kódování dvou zpracovatelů – autorky a kontrolora – a hledání shod a rozdílů, což vedlo k následné diskusi). O tom více pojednává samostatná podkapitola 3.4.5 (*Revize databáze sémantických rysů*).

¹²² McRae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 551.

3.4.3 Statistické zpracování dat¹²³

Na začátku této podkapitoly shrňme ještě jednou všechny podstatné informace o databázi sémantických rysů. Sémantické rysy byly získány pro 104 substantivních pojmů z 13 různých sémantických kategorií – z každé kategorie bylo vybráno 8 zástupných pojmů. Každý pojem byl popsán přibližně 20 participanty. Celkem se sběru zúčastnilo 381 probandů ve věku 8–9 (výjimečně i 10) let. Databáze obsahuje celkem 2146 různých informací přiřazených k pojmům, tzn. 2146 odlišných sémantických rysů.

Aby mohla být získaná data dále interpretována, byla zpracována pomocí několika statistických metod, přičemž jsme se inspirovali výpočty, pomocí kterých svou databázi vyhodnotil K. McRae a kol.¹²⁴

V této podkapitole popisujeme jednotlivé statistické metody, které byly pro zpracování databáze využity. V další podkapitole jsou pak prezentovány dílčí výsledky a k nim přiloženy jednotlivé interpretace. Kompletní výsledky statistických zpracování lze najít na internetové adrese www.childrenfeatures.webnode.cz – jedná se zpravidla o velmi rozsáhlé datové tabulky, které není možné představovat na jednotlivých stránkách této práce.

Výchozím materiálem pro následující statistické výpočty byly dvě sady dat – v prvním případě jde o veškeré sémantické rysy, které byly od probandů získány, ve druhé sadě jsou odstraněny taxonomické rysy – ty totiž mají specifický charakter, jak uvádí K. McRae a kol.: „(...) proměnné byly propočítány také bez taxonomických rysů, protože popisy týkající se nadřazené kategorie (nebo kategorií), ke kterým pojmy náležejí, považujeme za odlišný typ informací než ty, které přinášejí jiné typy rysů (části, funkce atd.)“¹²⁵.

Prezentované statické analýzy jsou zaměřeny na zpracování rysových reprezentací pojmů, na popisy kategorií a pojmů, ale také na vztahy, které se mezi kategoriemi, pojmy i rysy objevují, a to i napříč celou databází. K. McRae podotýká, že „Distribuční statistiky, jako jsou statistické zákonitosti mezi rysy, se osvědčily jako obzvláště prospěšné pro pochopení sémantických komputací.“¹²⁶

K. McRae dále uvádí, že v jejich případě byly do výpočtů zahrnuty pouze ty rysy, které byly pro daný pojem zmíněny minimálně pětkrát¹²⁷. V našem případě byly z výpočtů vyjmuty rysy, které byly pro daný pojem zmíněny pouze jedním probandem. Rozhodli jsme se pro tuto mírnější variantu z toho důvodu, že naše databáze není tak obsáhlá jako ty, které jsou prezentovány v našich vzorových studiích. Na druhou stranu považujeme za vhodné vyjmut

¹²³ Za pomoc se statistickým zpracováním dat srdečně děkuji Mgr. Kristýně Vaňkátové.

¹²⁴ **McRae, K. et al.:** *Semantic feature production norms for a large set of living and non-living things*, 2005.

¹²⁵ Tamtéž, s. 552. (překlad autorky)

„(...) variables were also calculated with taxonomic features excluded because we consider statements regarding the superordinate category (or categories) to which a concept belongs as being somewhat different from information regarding other types of features (parts, function, etc.).“

¹²⁶ Tamtéž, s. 548. (překlad autorky)

„Distributional statistics, such as statistical regularities among features, have proven particularly insightful with regard to understanding semantic computation.“

¹²⁷ Tamtéž, s. 552.

alespoň tyto rysy s nevýznamnou produkční frekvencí. Lze totiž předpokládat, že se nejedná o rysy pro pojem informativně přínosné – jsou pouze vyjádřením specifické představy/znalosti (či neznalosti) daného probanda, nepopisují tedy obecnou představu pojmu u dané věkové kategorie, nýbrž individuální reprezentaci. Nicméně v databázi tyto sémantické rysy v našem případě zahrnuty jsou a je proto možné s nimi případně dále pracovat a nějakým způsobem je v případě potřeby interpretovat.

Odstranění takového ojedinělého sémantického rysu u jednoho konkrétního pojmu nemusí však nutně znamenat jeho celkové vyloučení z výpočtů – uveďme příklad, kdy pouze jeden proband napíše, že jahoda je zelenina. V případě jahody bude tento rys odebrán, ve výpočtech ale zůstane, protože je podstatný pro jiné pojmy, které skutečně patří mezi zeleninu.

Zjednodušeně pak můžeme říci, že v matici, která je výchozí pro následující statistické výpočty, jsou všechny číslice 1 změněny na číslici 0. Pokud pak celý řádek (tj. sémantický rys) obsahuje pouze nuly, je z matice úplně odebrán.

Základem pro statistické výpočty je tedy $(n \times p)$ matice M . Každý sloupec v matici náleží jednomu pojmu a každý řádek jednomu sémantickému rysu. Jednotlivé buňky s číselnou hodnotou pak ukazují tzv. produkční frekvenci – tedy údaje o počtu participantů, kteří uvedli rys i pro pojem j . S ohledem na okolnost, že každý pojem nebyl popsán stejným počtem participantů, preferujeme relativní četnost namísto absolutní četnosti. Tato relativní informace je uložena v matici R a je rovna podmíněné pravděpodobnosti výskytu rysu v daném pojmu: $P(F_i|C_j)$, kde F_i vyjadřuje, že je popisován rys i , a C_j vyjadřuje, že je popisován pojem j . Tento vzorec může být rovněž zapsán jako

$$r_{ij} = \frac{m_{ij}}{\sum_{h=1}^n m_{hj}}$$

kde m jsou prvky matice M .

V případě, kdy nás zajímá informativnost analýzy s ohledem na jednotlivé pojmy, je důraz kladen na otestování různorodosti a podobnosti mezi pojmy. Na začátku je tedy provedeno rozlišení mezi podobnými pojmy. Základním výpočtem (s výchozí maticí R) je tzv. validita vodítka (anglicky „cue validity“). Jde v podstatě o výpočet podmíněné pravděpodobnosti, která nám říká, jaká je pravděpodobnost, že daný pojem patří do dané kategorie, pokud mu náleží určitý rys (v našem případě se jedná o pojem j a rys i): „Například validita vodítka, že je okřídlený objekt pták, je $P(\text{pták}|\text{křídla})$, což je pravděpodobnost, že něco je pták, když to má křídla.“¹²⁸ Podle Bayesovy první věty může být pravděpodobnost $P(F_i|C_k)$ vyjádřena také jako

$$P(C_j|F_i) = \frac{P(F_i|C_j)P(C_j)}{\sum_{k=1}^p P(F_i|C_k)P(C_k)} = \frac{P(F_i|C_j)}{\sum_{k=1}^p P(F_i|C_k)}$$

¹²⁸ **Murphy, G. L.:** *The big book of concepts*, 2002. S. 213. (překlad autorky)

„For example, the cue validity for a winged thing that being a bird is $P(\text{bird}|\text{wings})$, which is the probability that something is a bird, given that it has wings.“

– za předpokladu, že je pravděpodobnost pro libovolný pojem popsána vždy stejně. Pokud je dána matice V , která zaznamenává validitu vodítka a je utvořena stejným způsobem jako předešlé dvě matice, vzorec pro výpočet bude následující:

$$v_{ij} = \frac{r_{ij}}{\sum_{k=1}^p r_{ik}}$$

To znamená, že validita vodítka je vyjádřena podílem produkční frekvence rysu i patřícího pojmu j a součtu všech produkčních frekvencí rysu i napříč pojmy. E. Roschová a C. B. Mervisová definovaly validitu vodítka jako „procesní model klasifikace, ve kterém je pravděpodobnost podnětu (rysu) definována ve smyslu jeho celkové frekvence v kategorii a jeho proporční frekvence v této kategorii ve vztahu k jiným kategoriím.“¹²⁹

Tzv. distinktivní rysy (angl. distinguishing features) jsou rysy, které jsou specifické pro daný pojem a tím jej odlišují od všech ostatních pojmů. Rys je považován za distinktivní, pokud se objevuje pouze u jednoho, případně u dvou pojmů z celé databáze. Tento jev byl zaznamenán rovněž procentuálně: pro každý jednotlivý pojem byl vypočítán podíl počtu distinktivních ze všech rysů u něj zmíněných. S tím souvisí další jev, tzv. distinktivnost, která zachycuje celou škálu – od rysů, které jsou významně distinktivní, až po rysy, které sdílí velké množství pojmů, a tudíž distinktivní nejsou. Podíváme-li se pak na tyto rysy v rámci jednotlivých pojmů, je možné vysledovat, které pojmy se na základě těchto distinktivních rysů významově vzdalují od ostatních pojmů v databázi. Tato distinktivnost byla vypočtena pro každý rys pomocí inverzní hodnoty počtu pojmů, ve kterých se rys objevil. V takovém případě má rys, který je maximálně distinktivní (tj. neobjevuje se u žádného jiného pojmu) hodnotu 1,0. A naopak – v čím více pojmech se rys objeví, tím nižší je jeho distinktivnost a validita vodítka.

Dále jsme se zaměřili na rysy, které jsou nějakým způsobem typické pro celou sémantickou kategorii. V tomto případě jsme zjišťovali, které rysy byly probandy velmi často zmiňovány a zároveň se vztahují k jedné kategorii. Pro každou kategorii byly vygenerovány tři takovéto nejtypičtější rysy.

Podobnost mezi pojmy lze zjistit pomocí výpočtu, který udává míru podobnosti mezi dvěma proměnnými. V našem případě pracujeme s výše zmíněnou maticí R , kde je každý sloupec chápán jako vektor s n pozorováními náhodné proměnné C_j , $j = 1, \dots, p$. Standardně se pro takový výpočet podobnosti používá Pearsonův korelační koeficient, nicméně P. Ahlgren, B. Jarneving a R. Rousseau¹³⁰ uvádějí, že tento korelační koeficient je citlivý na nuly. V souboru dat, jako je ten náš, mohou nuly přispět ke zvýšení podobnosti, což nakonec vede k oslabení korelačního koeficientu mezi proměnnými. Proto bylo vhodné najít alternativní metodu výpočtu, která výsledky nezkruskuje. Preferujeme zde tedy – podle G. Saltona a M. J. McGilla¹³¹

¹²⁹ Rosch, E. a Mervis C. B.: *Family resemblances: Studies in the internal structure of categories*, 1975. S. 575. (překlad autorky)

„(...) processing model of classification in which the validity of a cue is defined in terms of its total frequency within a category and its proportional frequency in that category relative to contrasting categories.“

¹³⁰ Ahlgren, P., Jarneving, B. a Rousseau, R.: *Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient*, 2003.

¹³¹ Salton, G. a McGill, M. J.: *Introduction to modern information retrieval*, 1983.

– kosinus, který není citlivý na nuly. Kosinus pro dva vektory x a y , oba z dimenze n , je možné v rámci našich výpočtů formulovat následovně:

$$\cos(C_j, C_l) = \frac{\sum_{h=1}^n r_{hj}r_{hl}}{\sqrt{\sum_{h=1}^n r_{hj}^2} \sqrt{\sum_{h=1}^n r_{hl}^2}}$$

Kosinus se tedy rovná skalárnímu součinu dvou pojmových vektorů matice R , která je složena z produkčních frekvencí, a je vydělen násobkem jejich délek. Vzhledem k tomu, že Pearsonův kolerační koeficient lze vyjádřit vzorcem

$$r(\mathbf{x}, \mathbf{y}) = \frac{\sum_{h=1}^n (x_h - \bar{x})(y_h - \bar{y})}{\sqrt{\sum_{h=1}^n (x_h - \bar{x})^2} \sqrt{\sum_{h=1}^n (y_h - \bar{y})^2}}$$

kde $\bar{x}(\bar{y})$ je průměr výběrového souboru vektoru $x(y)$, je zřejmé, že vzorec kosinu je stejný jako Pearsonova korelace, ale nedochází zde k centrování vektorů k průměru. Z geometrického hlediska to znamená, že u Pearsonovy korelace je počátek vektorového prostoru situován do středu souboru, zatímco v případě kosinu se vytváří vektorový prostor na počátku, kde mají všechny vektory hodnotu 0^{132} . To nám pomáhá vypořádat se s takovými vektory, jejichž součástí je velké množství nulových hodnot. Preferujeme tedy v tomto případě kosinus. Kosinus se pohybuje mezi hodnotami -1 (opačné vektory) až 1 (identické vektory), přičemž 0 zde odpovídá nezávislým vektorům. Pojmová podobnost byla vypočítána pro každou párovou kombinaci pojmových vektorů a je zanesena v matici (pxp).

Podobnost mezi sémantickými rysy byla naopak zjišťována právě pomocí Pearsonova korelačního koeficientu, jako je tomu i v případě studie K. McRaeho a kol.¹³³. Takový postup je totiž vhodnější pro následnou interpretaci dat a je výhodnější pro případné další výpočty.

Vzhledem k charakteru dat, která jsou zpracovávána, je důležité vyhnout se nepravým korelacím, které indikují vztahy mezi jednotlivými prvky tam, kde ve skutečnosti žádné neexistují. Abychom se těmto nežádoucím podobnostem vyhnuli, byly do korelační analýzy zahrnuty pouze ty rysy, které byly participanty uvedeny minimálně u tří pojmu.

Výsledky tohoto statistického zpracování dat a jejich interpretace jsou uvedeny v následující podkapitole. Údaje o pojmové podobnosti byly navíc využity v experimentu, který představujeme v kapitole 3.5.

3.4.4 Výsledky zpracování databáze a diskuse

Statistickým zpracováním dat, které podrobně popisujeme v předchozí podkapitole, jsme získali velké množství informací. Výsledky těchto statistických výpočtů jsou v kompletní podobě publikovány na webové adrese www.childrenfeatures.webnode.cz.

¹³² Egghe, L. a Leydesdorff, L.: *The relation between Pearson's correlation coefficient r and Salton's cosine measure*, 2009.

¹³³ McRae, K. et al.: *Semantic feature production norms for a large set of living and non-living things*, 2005. S. 553.

Celkem bylo získáno 2146 různých sémantických rysů pro 104 pojmů z 13 sémantických kategorií. Cílem bylo získat pro každý pojem popis minimálně od 20 participantů. Bylo zjištěno, že každý účastník uvedl v průměru čtyři sémantické rysy pro jeden pojem. Následující tabulka 8 znázorňuje, kolik rysů pro pojem participantů průměrně uváděli v rámci jednotlivých sémantických kategorií.

kategorie	SAVCI	OVOCE	PTÁCI	ZELENINA	ROSTLINY	ČÁSTI TĚLA	
průměrný počet rysů uváděný pro jeden pojem	5,8	5,2	4,8	4	4	3,5	
kategorie	DOPRAVNÍ PROSTŘEDKY	NÁŘADÍ	ZBRANĚ	NÁBYTEK	KUCHYŇSKÉ NÁČINÍ	OBLEČENÍ	HUDEBNÍ NÁSTROJE
průměrný počet rysů uváděný pro jeden pojem	4,2	4	3,9	3,9	3,8	3,7	3,5

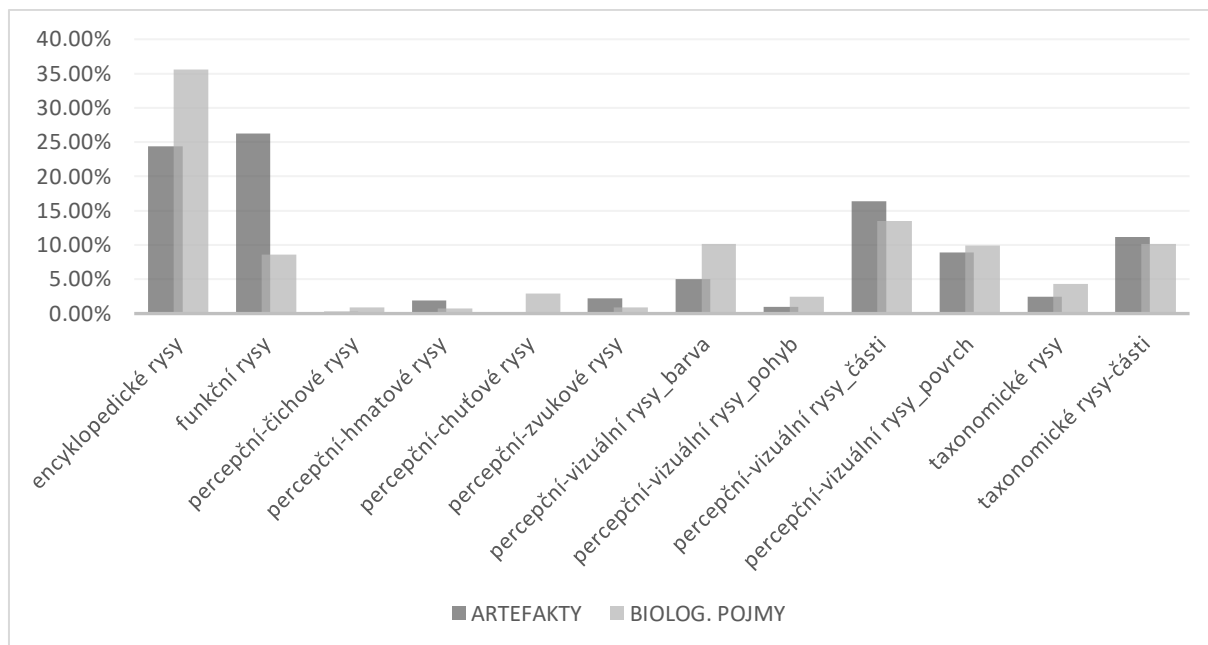
Tab. 8. Průměrný počet rysů uváděných pro pojem v jednotlivých sémantických kategoriích

Rysy byly podle typu informace rozděleny do několika kategorií. Do kategorie taxonomických rysů byly zařazeny především rysy korespondující se sémantickými kategoriemi, které jsme si na počátku zvolili („savci“, „ptáci“, „ovoce“, „zelenina“, „zbraně“, „nábytek“, „oblečení“ atd.), ale zařadili jsme zde také jiné rysy označující základní taxonomické kategorie – například u „savců“ a „ptáků“ se jednalo o rys *je to zvíře* apod.; do taxonomických rysů–částí byla pak zařazena méně specifická taxonomická určení, jako například rys *je to věc*, který se u dětských participantů vyskytoval poměrně běžně, a to v podstatě u všech popisovaných pojmů. Percepční rysy byly rozděleny na takové, které popisují zrakové, zvukové, chuťové, hmatové a čichové vjemy. V případě percepčních rysů popisujících zrakové vjemy (nazýváme je „percepční-vizuální rysy“) byly tyto rysy dále rozřazeny podle toho, zda se jednalo o popis povrchového vzhledu (například *je dlouhý, je krásný*), zda byly uváděny části, ze kterých se objekt skládá (například *má kola, má oči*), zda šlo o popis nějakého pohybu (*běhá, teče*), nebo zda byly pro daný objekt uvedeny informace o barevnosti (*je bílý, je barevný*). Mezi rysy popisující zvukové vjemy jsme zařadili například *je hlasitý, píská, cinká, mňouká*; mezi ty, které popisují chuťové vjemy, jsme zařadili například *je chutný, je sladký*; mezi rysy popisující hmatové vjemy jsme zařadili například *pálí, píchá, je ostrý, je teplý, je měkký*; jako rysy popisující čichové vjemy jsme označili rysy *smrdí/(ne)voní* a rysy, které tuto vůni / tento zápach specifikují. Další typ rysu byl takový, který popisuje nějakou funkci daného pojmu (objektu), zde patří rysy jako *hrabe se tím, hraje se na to, je k jídlu* apod. Ostatní rysy, které jsme nedokázali zařadit do žádné z výše uvedených kategorií, byly zařazeny mezi rysy encyklopedické – většinou se jednalo o nějaké specifické znalosti a rozšiřující informace o daném pojmu: například u pojmu **zub** participantů uváděli, že se zuby čistí zubní pastou; dále zde patří například informace o tom, kde se daný objekt nachází, o tom, že je nebezpečný apod.

V příloze 8 uvádíme dva příklady pojmů (jednoho zástupce biologických kategorií a jednoho zástupce artefaktů) se všemi sémantickými rysy, které byly u těchto pojmů uvedeny více než

jedním participantem. Dále je u jednotlivých rysů uvedena jejich frekvence a typ (dle výše uvedeného rozdělení).

V následujícím grafu 1 lze vidět srovnání distribuce jednotlivých typů rysů u artefaktů a u biologických kategorií. V tomto případě nebyla brána v potaz u jednotlivých sémantických rysů jejich produkční frekvence: například pro pojem **orel** děti uvedly čtyřikrát *rys je černý*, dvakrát *rys je bílý* a dvakrát *rys je hnědý*. Podstatné tedy je, že se jedná o tři různé rysy týkající se zrakového vjemu (popisující barvu) a není podstatné, kolikrát byly tyto rysy dětmi zmíněny.



Graf 1. Distribuce jednotlivých typů rysů u artefaktů a biologických kategorií

Jak lze z tohoto grafu vyčíst, dětské participanty inklinovali k uvádění rysů týkajících se vzhledu, především k takovým, které popisují jednotlivé části daného pojmu a vlastnosti týkající se jeho povrchu. Participanty uváděli také velké množství encyklopedických rysů, a to především pro biologické kategorie. To pravděpodobně souvisí nejen s osobní zkušeností, ale také s mírou znalostí nabytých vzděláním – dá se předpokládat, že se dítě ve výuce setká spíše s výkladem týkajícím se například **ptáků** či **rostlin** než **nábytku** či **nádobí**, a zdá se, že děti rády při popisování pojmů tyto nabyté znalosti prezentovaly. Rysy popisující chuť byly uváděny především u biologických kategorií, což není překvapivé, vzhledem k tomu, že součástí této skupiny jsou mimo jiné také kategorie „**ovoce**“ a „**zelenina**“. Přesto byl jednou tento typ rysu uveden i u artefaktů, konkrétně u pojmu **pistole**, kdy jedno z dětí uvedlo rys *nemá chuť*. Taková „negativní informace“ není v databázi ojedinělá – domníváme se, že uvádění takovýchto rysů bylo dáno vlivem příkladů, které byly představeny na začátku sběračem dat, případně mohl být participant ovlivněn popisem předchozího pojmu a docházelo zde ke srovnání. Dětské participanty často uváděli taxonomické rysy – to může souviset se skutečností, že děti v tomto věku začínají o pojmech přemýšlet více objektivně a méně je popisují pomocí nějaké konkrétní osobní zkušenosti (viz k tomu také podkapitolu *Osvojování pojmů u dětí*). Základní taxonomické zařazení bylo častěji uváděno u pojmů z biologických kategorií, což nejspíše koresponduje s každodenní realitou a běžnou komunikací – nadřazené kategorie jako je „**ovoce**“, „**zelenina**“, „**květina**“ či „**strom**“ a každodenní zařazování jednotlivých předmětů

do těchto kategorií je pravděpodobně v běžném životě častější než poukazování například na to, že **stůl** je „nábytek“ a **autobus** je „dopravní prostředek“. Na druhou stranu, taxonomické rysy-části se objevovaly častěji v případě kategorií artefaktů. To může být dáno i tím, že dětské probandi velmi často uváděli *rys je to věc*, a to právě především u artefaktů, ačkoliv ne výlučně u nich – v několika případech byla tato vlastnost také zaznamenána například také u pojmů z kategorie „části těla“.

Další statistické výpočty nám přináší informace o tzv. distinktivnosti a typičnosti rysů. Přibližně u 40 % pojmů byla zaznamenána průměrná distinktivnost s vyšší hodnotou než 0,5, což poukazuje na to, že účastníci uváděli velké množství rysů odlišujících pojem od pojmu ostatních. Pojem, u kterého bylo uvedeno nejméně takovýchto distinktivních rysů, byl **orel** (s hodnotou 0,168). To znamená, že rysy, které účastníci u tohoto pojmu uváděli, nebyly specifické pouze pro tento pojem, ale naopak byly sdílené (například s ostatními pojmy z kategorie „ptáci“). Může to být dáno tím, že děti nemají konkrétnější představu o tomto zástupci kategorie a dokáží ho popsat spíše jen na obecné úrovni jako jiné zástupce ptáčích říše. Obecně ale můžeme říci, že přítomnost nediferencujících, sdílených rysů mezi pojmy jednotlivých biologických kategorií byla poměrně typická, což může být vysvětleno homogenitou uvnitř těchto kategorií, potažmo uvnitř celé skupiny biologických kategorií. Nejvíce distinktivních rysů bylo zaznamenáno v případě pojmu **noha** (0,810). Obecně lze říci, že pro celou kategorii „části těla“ byla přítomnost distinktivních rysů výrazně typická.

Ostatně připouštíme, že celá tato biologická kategorie je specifická – jejím základem jsou pojmy označující pouze části celistvého organismu, a tudíž se liší od ostatních kategorií z této skupiny. Přestože je zřejmé, že děti mají hlubší představu o pojmech z této kategorie, jedná se o jednu ze dvou kategorií, ve kterých bylo zmiňováno průměrně nejméně rysů pro pojem, tj. 3,5 rysů. Druhou kategorií s takto nízkým průměrným počtem rysů pro pojem byly „**hudební nástroje**“. V této kategorii však mohla být tato nízká produkce dána menším povědomím o významu některých pojmů (v této kategorii účastníci popisovali pojmy mimo jiné pojmy, jako je **harfa** či **varhany**).

Následující tabulka 9 představuje typické sémantické rysy pro vybrané sémantické kategorie (do tohoto přehledu nejsou zahrnuty taxonomické rysy).

biologické kategorie příklady	typické rysy	kategorie artefaktů příklady	typické rysy
PTÁCI	<i>létá, má peří, je živý</i>	NÁŘADÍ	<i>je ze železa, je ze dřeva, má rukojeť</i>
SAVCI	<i>je živý, má ocas, skáče</i>	ZBRANĚ	<i>střílí se tím, je nebezpečný, je ostrý</i>
OVOCE	<i>je sladký, roste, je červený</i>	HUDEBNÍ NÁSTROJE	<i>hraje se na to, zvuk/y vydává, je ze dřeva</i>
ZELENINA	<i>je k jídlu, je zelený, roste</i>	OBLEČENÍ	<i>nosí se, má různou barvu, je z látky</i>

Tab. 9. *Účastníci nejčastěji uváděné sémantické rysy (vybrané kategorie)*

Pomocí Pearsonova korelačního koeficientu bylo z databáze vygenerováno 15 významných párů korelačních rysů. Mezi tyto páry patří například dvojice *je kyselý* a *má vitamíny*, *je šťavnatý* a *je sladký*, *má křídlo/a* a *létá, má pecku/y* a *je na stromě*, *pluje* a *je na vodě*, *voní* a *je*

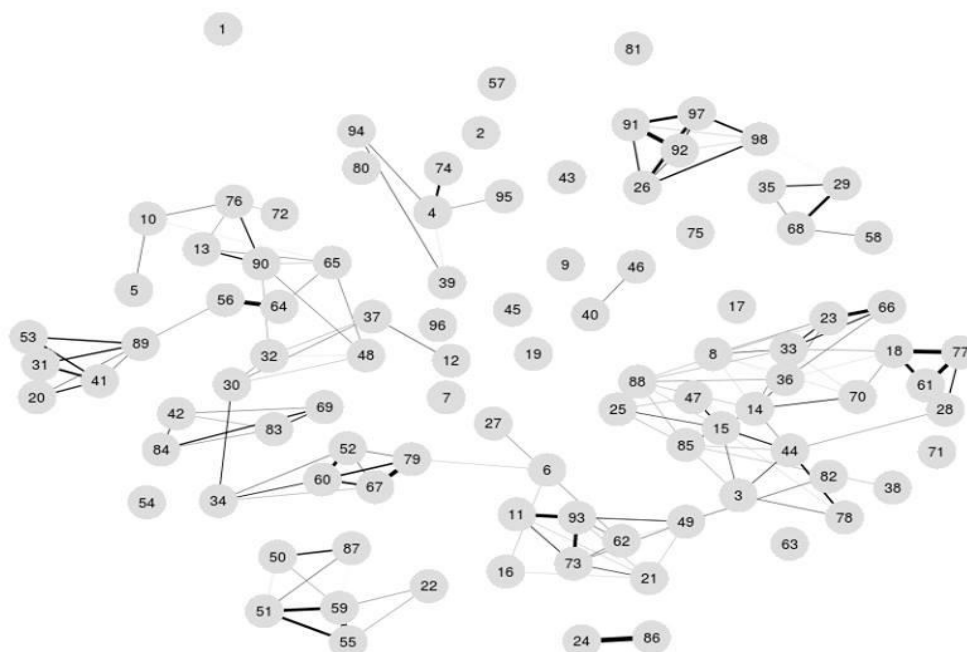
hezký. Ze získaných dat vyplývá, že vztahy se typicky objevují mezi rysy různých typů (například mezi rysem popisujícím část pojmu a rysem popisujícím pohyb, který tato část umožňuje apod.).

Graf 2 znázorňuje vztahy a závislosti mezi rysy. Čím tlustší a tmavší je čára spojující dva rysy, tím silnější je vztah mezi těmito dvěma rysy. Vzhledem k množství dat není možné uvádět u každého čísla rys, který toto číslo zastupuje, graf má tedy pouze ilustrativní charakter. Pro představu alespoň uvádíme, pod kterými čísly se skrývají výše uvedené páry: *je kyselý a má vitamíny* jsou označeny čísly 18 a 77; *je šťavnatý a je sladký* mají čísla 36 a 33; *má křídlo/a a létá* mají čísla 60 a 52; *rysy má pecku/y a je na stromě* mají čísla 66 a 23; *pluje a je na vodě* jsou čísla 86 a 24; a nakonec, *voní a je hezký* jsou čísla 93 a 11.

Nejvyšší rysová podobnost (bez zahrnutí taxonomických rysů) byla zaznamenána v kategorii biologických pojmů mezi pojmy **pomeranč** a **mandarinka** (0,860) a **třešeň** a **višeň** (0,841), u artefaktů pak mezi pojmy **čepice** a **klobouk** (0,827) a **kytara** a **housle** (0,814). Tyto informace o rysové podobnosti jsou dále využity v návazném experimentu, ve kterém ověřujeme významnost sémantických rysů při rychlém rozhodování týkajícím se pojmů a jejich významů.

Databáze, kterou představujeme, obsahuje data získaná od dětí mladšího školního věku, konkrétně od 8–10letých dětí. Dětské participanti popisují pojmy způsobem, který se může významně lišit od popisu těchto pojmů dospělými jedinci. To je viditelné především na konkrétních odpovědích (které jsou pak pro naše účely dále kódovány do zjednodušené a jednotné podoby). Děti ve věku 8–10 let už dokáží popsat mnohé důležité vlastnosti představovaných pojmů, zároveň je ale při popisu zřejmá jazyková i obsahová kreativita, která je daná nejspíše prozatímni neznalostí různých jazykových výrazů a po obsahované stránce nedostatečnou životní zkušeností. Uvedme si několik případů z kategorie „**dopravní prostředky**“. U **traktoru** jeden z participantů uvádí, že „*má v předu velkou naběračku*“, přičemž se, pokud je nám známo, běžně používá spíše termín „*lžíce*“. Další uvádí u **letadla**, že „*vejde se tam 1000 lidí*“, což podle nás svědčí o nedostatečně vyvinuté schopnosti odhadu a zároveň o potřebě uvádět informace dostatečně konkrétně (předpokládáme, že dospělý participant by uvedl spíše informaci, že **letadlo může** mít velkou kapacitu a **může** přepravovat velké množství pasažérů současně, ovšem zdůrazněná modalita je zde důležitá, protože existují i malá letadla). V případě **trolejbusu** bylo dvěma probandy uvedeno, že jezdí po kolejích: „*jezdí po kolejích tam kde jezdí auta*“ a „*jezdí to na koleje*“ – to může být způsobeno nedostatečnou zkušeností s tímto „**dopravním prostředkem**“, který je používán pro veřejnou dopravu jen v několika městech České republiky; podobně jako v případě lžíce u traktoru i zde nacházíme nepřesná označení, kdy například jeden z participantů uvádí, že **trolejbus** „*je na draty*“, protože nejspíše nezná slovo „*troleje*“.

Je také zřejmé, že se v tomto věku u některých participantů stále uplatňuje tendence upřednostňovat při popisování pojmů osobní zkušenost před obecnější charakteristikou.



Graf 2. Ilustrace vztahů a závislostí mezi rysy

Jedno z dětí například pojem **okurka** popisuje takto: „*To je jídlo je dobré a dnes sem ho měl na sváču*“, další k pojmu **auto** mimo jiné uvádí, že „*k babičce když nechceš jít pěšky tak dež do auta a jsi tam z 2 minut*“, jiné dítě k popisu **flétny** dodává, že „*jeto těžké naní těžké hraje nato můj brácha*“. Zdá se, že takové typy subjektivních odpovědí korespondují s vývojovou fází poznávacích procesů u školáků, tj. s vývojem od konkrétnější podoby obsahu pojmu k obecnější podobě, jak to uvádí Vágnerová¹³⁴, kterou jsme citovali v podkapitole *Osvojování pojmů u dětí*.

U dětských participantů byl také zaznamenán vliv struktury úlohy na způsob popisu pojmů. Výše jsme zmiňovali, že jedno z dětí uvedlo u pojmu **pistole** rys *nemá žádnou chuť*. Podobně další děti uvádějí u pojmu **věšák** rys *nemá obličej*, u pojmu **ananas** rys *nehýbe se*, u pojmu **židle** rys *není zvíře* apod. Tento jev se v získaných datech vyskytuje poměrně často.

V příloze 10 předkládáme porovnání encyklopedické definice dvou vybraných pojmů se sémantickými rysy získanými od dětí. Takové srovnání sice považujeme vzhledem k získaným zjištěním za zajímavé, není ale cílem naší práce, proto uvádíme pro ilustraci pouze dva pojmy. Pojem **papoušek** z kategorie „**ptáci**“ zde zastupuje skupinu biologických pojmů a pojem **triangl** z kategorie „**hudební nástroje**“ zastupuje artefakty. U uvedených příkladů vidíme, že dětští participanté zmínili všechny rysy, které jsou uváděny ve *Slovníku spisovné češtiny pro školu a veřejnost*¹³⁵. Rysy, které najdeme ve slovníku, lze považovat za ty nejdůležitější – definice by

¹³⁴ Vágnerová, M.: *Kognitivní a sociální psychologie žáka základní školy*, 2001. S. 110.

¹³⁵ Červená V. et al.: *Slovník spisovné češtiny pro školu a veřejnost*, 2005.

měla vystihovat daný pojem, případně jej odlišit od ostatních, především podobných pojmů. Rysy odpovídající těm ve slovníku byly při sběru dat uvedeny zpravidla více než jedním dítětem (tj. mají vyšší frekvenci než 1), je zde tedy patrná korespondence mezi obecně platnou důležitostí těchto rysů a jejich důležitostí pro skupinu dětí ve věku 8–10 let. Problematické může být přiřazení slovníkových charakteristik k sémantickým rysům uváděnými dětmi, protože ty se mohou rozdělovat na více – různě přesných či nepřesných – rysů (tyto rysy se nemusejí vždy shodovat s realitou). U pojmu **triangl** je ve slovníku uvedeno, že *je kovový*. U dětí však vedle sémantického rysu *je z kovu*¹³⁶, který uvedli tři participanty a ke kterému charakteristickou vlastnost uvedenou ve slovníku přiřazujeme, byl pětkrát zaznamenán také výskyt rysu *je ze železa*, tedy rys specifikující druh kovu. Nelze ani vyloučit možnost, že děti kov a železo chápou jako synonyma. Tento hudební nástroj navíc bývá zpravidla vyráběn z oceli, slitiny vyráběné zpracováním surového železa, přičemž tento materiál není uveden ani ve slovníku, ani mezi popisy získanými od dětí. U **papouška** pak přiřazujeme charakteristiku *pestře zbarvený* k rysu *je barevný* a *s pronikavým hlasem* k rysu *vydává zvuk/y*, přestože se popis těchto vlastností mírně liší.

Podobným způsobem v příloze 11 porovnáváme sémantické rysy dvou vybraných pojmů z naší databáze se sémantickými rysy dvou pojmů z databáze, kterou vytvořil K. McRae se svými kolegy¹³⁷ pro dospělé anglické mluvčí. Vzhledem k tomu, že naše databáze neobsahuje stejné pojmy jako ta databáze, která byla vytvořena K. McRaem a kol., pokusili jsme se pro ilustraci porovnat alespoň sémantické rysy pojmů, které jsou si nějakým způsobem podobné – srovnali jsme tedy dva pojmy, **jelena** a **dýku**, které je součástí naší databáze, s **losem** a **nožem**.

V obou přílohách (10 a 11), ve kterých porovnáváme sémantické rysy produkované dětskými participanty se slovníkovou definicí a sémantické rysy získané námi a K. McRaem a kol., jsme v tabulkách pomocí barevného podbarvení zdůraznili ty rysy, které byly zmíněny více než jedním dítětem.

Zdá se, že v celkové rysové reprezentaci daných pojmů u dětí patří mezi nejdůležitější právě ty rysy, které se objevují ve slovníkové definici, i u dospělých jedinců. Dětské probandy se nezaměřovali pouze na popis vzhledu prezentovaných objektů, ale také například na jejich funkci či způsob pohybu apod. Přestože se jednotlivé reálné odpovědi dětí zdají být občas nepřesné či chaotické (což není u takového typu úkolu typické pro dospělé participanty), zdá se, že v rysové reprezentaci pojmů, které jsou vygenerovány ze všech získaných popisů, různé věkové skupiny vyzdvihují stejné či podobné sémantické rysy (případně typy rysů).

Domníváme se, že je zde možné pozorovat několik skutečností. Je zřejmé, že děti ještě nemají dostatečné zkušenosti a znalosti, aby jejich odpovědi mohly kvalitativně odpovídat odpovědím dospělých. To koresponduje se způsobem vyjadřování při popisu pojmů, ale také s některými nepřesnostmi při definování pojmů. Na druhou stranu, děti se často učí poznávat a chápat okolní svět za pomoci dospělých, což může mít vliv na jejich perspektivu, s kterou vše okolo sebe vnímají. To může být příčinou toho, proč se při popisu pojmů obě věkové skupiny zaměřují na podobné typy informací.

¹³⁶ Pod kódem *je z kovu* jsou zahrnuty i odpovědi typu „*je to kovové*“ či „*je to kovový nástroj*“.

¹³⁷ **McRae, K. et al.:** *Semantic feature production norms for a large set of living and non-living things*, 2005.

3.4.5 Revize databáze sémantických rysů

První verze databáze byla publikována v roce 2017, v letech 2019–2020 byla provedena kontrola a revize dat. Aktualizovaná data, stejně jako ta původní, jsou k dispozici na internetové adrese www.childrenfeatures.webnode.cz. Databáze nebyla rozšiřována o nová data, pouze došlo k drobným úpravám a opravám. Po této revizi se celkový počet získaných sémantických rysů snížil z 2146 na 2124.

Níže pro ilustraci změn uvádíme tabulku 10 s několika páry pojmů a hodnotou určující míru jejich vzájemné podobnosti, přičemž srovnáváme výsledky statistických výpočtů provedených na původních a na aktualizovaných datech. Změny ilustrujeme právě na pojmové podobnosti, protože se jedná o výpočet, který byl zásadní pro náš návazný výzkum. Z tabulky je zřejmé, že se nové hodnoty od těch původních výrazně neliší.

pojmový pár	hodnota korelace (2016)	hodnota korelace (2020)
SUP-OREL	0,879	0.877
POMERANČ-MANDARINKA	0,897	0.898
SEDMIKRÁSKA-KOPRETINA	0,878	0.877
ČEPICE-KLOBOUK	0,780	0.806
KYTARA-HOUSLE	0,876	0.880
TROLEJBUS-VLAK	0,699	0.702

Tab. 10. Srovnání vzájemné podobnosti dvou pojmů ve starší a novější verzi databáze

Při tvorbě databází sémantických rysů bývá zpravidla prováděna kontrola dalšími subjekty, kteří mají za úkol nezávisle posoudit vhodnost a konzistenci zvoleného kódování. V našem případě byla v průběhu tvorby databáze konzultována pravidla kódování v rámci výzkumného týmu. S dostatečným časovým odstupem byla databáze revidována samotnou autorkou práce. Dále byl požádán student z Katedry obecné lingvistiky a teorie komunikace Univerzity Palackého o vlastní zpracování části databáze na základě představených kódovacích pravidel.

Pro kontrolní rozbor byla vybrána data od 60 probandů následujícím způsobem. 60 sešitů bylo vybráno tak, aby z každé verze byly vzaty maximálně 3 sešity, tedy tak, aby byl materiál určený ke kontrole rozmanitější a pojmy se příliš neopakovaly. Student (kontrolor) dostal k dispozici originální sešity vyplněné probandy a také elektronický přepis těchto sešitů v textovém dokumentu. Dále mu byl předložen základní kódovací manuál a odkaz na naše vzorové zahraniční studie.

V dokumentu byl rovněž uveden seznam typů rysů (vizuální, funkce apod.) a kontrolorovi byl předán pokyn určit u každého rysu typ, a to na základě vlastního zvážení, kam se rys řadí. Rozdělení na taxonomické rysy a taxonomické-části bylo upraveno dodatečně autorkou výzkumu, na základě pevně daných pravidel, se kterými nebyl kontrolor seznámen. Podrobně jsou tato pravidla popsána na předchozích stranách této práce.

Dále byl v manuálu uveden přehledový seznam pojmů a jejich zařazení do jednotlivých sémantických kategorií. Poté, co kontrolor data zpracoval, odevzdal materiál autorce výzkumu. Následně byl požádán o dodatečné úpravy – v případě chybějících kódů, místo kterých byly uvedeny pouze poznámky svědčící o pochybnostech kontrolora, jak tato data zpracovat, byl požádán o doplnění těchto kódů pouze na základě svých vlastních rozhodnutí. Pro jednodušší rozhodování mu byla připomenuta některá pravidla z manuálu.

Před samotným vyhodnocením srovnání a byla celá databáze upravena na základě inspirace rysy uvedenými kontrolorem. Například informace *živí se* byla vyčleněna jako samostatný rys a také byly sjednoceny případy, kdy ze strany probandů docházelo k personifikaci – například u pojmu **ryč** byly sjednoceny rysy *kope* a *kope se tím*. Avšak v případě pojmů, jakou jsou **autobus** či **letadlo**, byly zachovány oba rysy *jezdí* a *jezdí se tím* jako dvě samostatné informace, protože v tomto případě není první varianta chápána jako pouhá personifikace, nýbrž jde o vyjádření pohybu, zatímco druhý rys vyjadřuje funkci.

Po těchto částečných úpravách v celé databázi bylo kontrolorovo zpracování části databáze srovnáno se zpracováním stejného úseky autorkou, přičemž byl pro získání objektivnějších a přehlednějších výsledků srovnání vytvořen skórovací systém tak, jak jej uvádíme v následující tabulce 11.

typ případu	bodování	vysvětlení
A	1 bod = max.	> absolutní shoda (případně jiná formulace, ale se shodným významem) Poznámky: 1. U kontrolora zde v některých případech nedošlo ke sjednocení kódů, ale ve většině případů je zřejmé, že jde o stejný význam u obou kódů. 2. Patří zde i případy, kdy dochází k preciznějšímu dělení informací ze strany autorky než ze strany kontrolora (např. u autorky je jako univerzální kód užíván <i>je v lese</i> , zatímco kontrolor uvádí různé kódy, např. <i>žije v lese</i> nebo <i>roste v lese</i> apod., nicméně význam výskytu v lese je zde shodný.
B	-0,5 b.	> informace se shoduje jen částečně , nejde jen o jinou formulaci, ale částečně také o jiný význam
C	-0,8 b.	> rys byl uveden oběma zpracovateli, ale značně se liší (tento případ je málo častý)
E	-0,2 b.	> uvedení odlišného typu rysu Poznámky: 1. V případě taxonomických rysů a taxonomických rysů-částí zavedla autorka v celé databázi jednotný systém. Pro více informací o rozřazení rysů do těchto dvou kategorií viz předchozí podkapitulu 3.4.4. U kontrolora nebyly autorkou provedeny tyto sjednocující úpravy, pokud uváděl jiný typ rysu. 2. Autorka se v některých případech inspirovala ve vzorových studiích, kontrolor mohl uvádět typy dle těchto studií či dle vlastního uvážení.
F	0 b. = min.	> kód uvádí jen autorka, nebo jen kontrolor, u druhého zpracovatele neexistuje odpovídající varianta

Tab. 11. Skórovací systém uplatněný při srovnání zpracování části databáze autorkou a kontrolorem

Na základě těchto pravidel skórování byla mezi autorkou práce a kontrolorem při kódování a určování typů rysů zaznamenána celková shoda 84,1 %.

Dodejme, že pojmy, které probandi nevyplnili či vyplnili špatně, nebyly zpracovávány. V tomto duchu byl také v případě jednoho participanta vyřazen pojem „**sup**“, jelikož zde popisoval „**šuplík**“, tedy zasouvací přihrádku, která bývá součástí některého nábytku. Protože byl původně popis tohoto pojmu chápán jako chybný, byl v případě materiálu připraveného pro kontrolora označen jako vyřazený. Později však došlo v tomto případě k přehodnocení a popis byl uznán: jedná se totiž o případ, kdy proband opsal v rámci popisu jméno pojmu, když uvedl, že „*sup je suplik na věci*“. Z toho vyplývá, že záměna nebyla způsobena špatným přečtením pojmu. V takových případech považujeme za nutné pojem do databáze zařadit: participant jednoznačně popisuje zadaný pojem, byť jeho představa o jeho významu není správná.

3.5 Rychlá sémantická úloha: ověření vlivu míry rysově podobnosti na sílu primingového efektu

3.5.1 Cíl experimentu

V této části práce čerpáme inspiraci ze studie *On the Nature and Scope of Featural Representations of Word Meaning*¹³⁸. Konkrétně nás zajímá druhý experiment z této studie (2A)¹³⁹.

Cílem experimentu je ověřit hypotézu, že rychlé sémantické úsudky (jako je například rychlé posouzení podobnosti dvou pojmů) jsou založeny především na prostých rysových reprezentacích pojmů¹⁴⁰. Zároveň se předpokládá, že v případě úloh, které nejsou časově omezené, jsou využívány rovněž jiné, hlubší schopnosti a znalosti, kterými jedinec disponuje.

Sémantická úloha založená na rychlém rozhodování, pomocí které testujeme naši hypotézu, využívá efekt automatického sémantického primingu. Jak uvádí K. McRae a jeho kolegové (na základě studie M. I. Posnera a C. R. R. Snydera¹⁴¹), rozhodující je zde délka tzv. SOA (tj. *stimulus onset asynchrony*, tedy čas, který uplyne mezi počátkem prezentace prvního a počátkem prezentace druhého podnětu¹⁴², v našem případě tedy mezi počátkem prezentace primového a cílového slova), který ovlivňuje typ sémantického primingu: „(...) pokud je SOA krátký, sémantický priming je automatický, zatímco s přibývajícím časem participantů využívají prime k vygenerování množiny očekávaných blížících se targetů.“¹⁴³

Pokud jsou tyto předpoklady správné, měla by být v představované sémantické úloze při rychlém rozhodování využita především rysová reprezentace pojmů a síla primingového efektu by mohla být přímo úměrná míře podobnosti pojmů. To znamená, že čím více jsou si dva pojmy (prime a target) rysově podobné, tím silnější bude primingový efekt a odpověď na otázku týkající se targetu bude rychlejší (měřený reakční čas RT bude kratší).

V našem případě byl experiment aplikován v českém prostředí a byl do jisté míry přizpůsoben věku participantů. Rysová podobnost pojmů byla vygenerována z databáze sémantických rysů, která byla sestavena z dat získaných od dětí mladšího školního věku, samotný experiment pak byl uskutečněn s jinými participanty, ovšem ze stejné věkové kategorie.

¹³⁸ **McRae, K., de Sa, V. R. a Seidenberg, M. S.:** *On the Nature and Scope of Featural Representations of Word Meaning*, 1997.

¹³⁹ Tamtéž, s. 106–110.

¹⁴⁰ V našem případě jsme pracovali pouze s rysovými reprezentacemi a podobnostmi vystavenými na individuálních rysech, zatímco McRae a kol. se v této studii zabývají také vztahy (korelacemi) mezi rysy a rysovými reprezentacemi a podobnostmi pojmů reflektujícími tyto korelace.

¹⁴¹ **Posner, M. I. a Snyder, C. R. R.:** *Facilitation and Inhibition in the Processing of Signals*, 1975.

¹⁴² *APA Dictionary of psychology* [online], heslo: *Stimulus onset asynchrony*. Dostupné z: <https://dictionary.apa.org/stimulus-onset-asynchrony> [cit. 20. 7. 2021]

¹⁴³ **McRae, K., Sa, V. R. de a Seidenberg, M. S.:** *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 106. (překlad autorky)

„(...) when SOA is short, semantic priming is automatic, whereas with more time, subjects use the prime to generate an expectancy set for the impending target.“

3.5.2 Metoda experimentu

3.5.2.1 Participanti

Experimentu se zúčastnilo celkem 39 dětí ve věku 8–10 let. Všechny zúčastněné děti v době experimentu měly ukončenou docházku do třetí třídy na běžné české základní škole v Olomouckém či Moravskoslezském kraji.

Experiment se uskutečnil během letních prázdnin a podmínkou pro účast dítěte byl poprázdňinový nástup do čtvrté třídy. Toto kritérium bylo stanoveno v souladu s věkovou kategorií účastníků se sběru dat do databáze. Vzhledem k tomu, že databáze byla sestavována z dat získaných v průběhu školního roku a experiment byl prováděn až po ukončení školního roku, jsou data v databázi získána především od dětí nižšího věku a experimentu se zúčastnily děti spíše vyššího věku. V obou případech však byla dodržena zvolená věková škála 8–10 let. Cílem bylo získat data od dětí se stejnou úrovní vzdělání, a tedy i s podobnými zkušenostmi a znalostmi, aby byla data z databáze kompatibilní s výsledky experimentu.

První tři participanti se zúčastnili pilotního testování, proto jejich data nejsou součástí níže uvedených výsledků. Výsledky tedy zahrnují data od 36 účastníků experimentu.

Sémantického experimentu se rovněž nemohly zúčastnit děti, které mají některou z poruch učení – zkušebně byly otestovány dvě děti s dyslexií a jedno dítě s dysfázií, tyto děti však nedokázaly včas přečíst všechna zobrazená slova: problém jim činila především primová slova, která se zobrazovala na obrazovce pouze po dobu 500 ms. Často však nestíhaly reagovat včas ani na cílová slova, u kterých byl maximální časový limit pro zodpovězení otázky desetinásobný. Výsledná chybovost odpovědí na otázku byla u těchto dětí mnohem vyšší než u ostatních účastníků. Děti samy po skončení úkolu přiznaly, že primová slova nestíhaly číst. V těchto případech tudíž nemohl vůbec nastat primingový efekt, který má v předkládaném experimentu stěžejní význam.

3.5.2.2 Materiály

Sémantická úloha, kterou v této části práce představujeme, byla časově limitovaná a vyžadovala rychlou reakci participanta. Úkol byl postaven na jednoduchém rozhodování, kdy participanti odpovídali na snadnou otázku zaměřenou na cílové slovo (respektive pojem a jeho význam). Jednalo se o otázku „Roste to?“ a „Vyrábějí to lidé?“. Z jazykovědného hlediska jde o tzv. otázky zjišťovací, kdy jediné možné odpovědi jsou pouze „ano“, nebo „ne“. Vzhledem k tomu, že otázka byla směřována na cílové slovo, kterému předcházelo krátké zobrazení primingového slova, mohlo dojít k vyvolání primingového efektu – to je pro experiment zásadní, jelikož předpokládáme, že tento efekt může mít vliv na rychlost participantovy reakce.

Typ sémantických otázek byl zvolen záměrně a liší se od standardního postupu, kdy se v primingových experimentech k targetům vztahují zpravidla otázky lexikálního charakteru („Jedná se o skutečné slovo?“). Sémantický typ úlohy, kdy probandi odpovídají na otázky typu „Roste to?“ a „Vyrábějí to lidé?“, více odpovídá tomu, proč byl experiment vytvořen – zajímá nás sémantická rovina jazyka a myšlení. Navíc je důležitý výběr velice obecné otázky, protože pokud bude vybraná kategorie dostatečně obsáhlá, nebudou mít participanti tendenci si předem vygenerovat v mysli zástupce kategorie. V případě otázek typu „Je to květina?“ si může proband, i když nevědomě, předem ve své mysli připravit několik zástupných pojmů pro danou kategorii a narušit tak průběh experimentu; takovéto generování zástupců je v případě otázky typu „Je to živé?“ mnohem náročnější, protože kategorie živých věcí je mnohem

obsáhlejší než kategorie květin. K. McRae a kol. vysvětlují důvody výběru takové otázky¹⁴⁴: „Byla upřednostněna úloha sémantického rozhodování jako „Je to živé?“ před úlohou detekce slova, která je standardní v primingových experimentech. (...) zatímco sémantické rozhodování je jasně založeno na sémantické znalosti, lexikální úloha nabízí možnost využít jiné možnosti při tvorbě odpovědí (...).“¹⁴⁵

Podle K. McRae a kol.¹⁴⁶ tento typ úlohy navíc umožňuje vytvořit takové výplňové položky, které zabraňují tomu, aby si participant na základě již zobrazených primových a cílových slov odvodil pravidla párování těchto slov. V úloze se proto objevují dva typy výplní, které byly vytvořeny za pomoci náhodně vybraných slov, jež nejsou součástí databáze. Jako příklad prvního typu výplně uveďme pár pojmů **čítanka** (prime) a **houska** (target) pro otázku „Roste to?“ – odpověď na danou otázku je pro primové i cílové slovo „ne“. Druhý typ výplně se skládal z páru, ve kterém sice byla odpověď na výše uvedenou otázku pro cílové slovo „ano“ (například pojem **kaštan**), ale nemohlo zde dojít k primingovému efektu, protože odpověď na otázku pro primový pojem byla „ne“ (například pojem **cihla**). Tyto výplňové položky tedy slouží k tomu, aby příbuznost pojmů automaticky neevokovala kladnou odpověď na otázku.

V případě sledovaných párů byla odpověď na danou otázku „ano“ pro prime (například **sup**) i target (například **orel**), mohlo tedy dojít k primingovému efektu, který urychlí odpověď pro cílové slovo.

Přestože pojmy byly párovány na základě jejich podobnosti v rámci užší sémantické kategorie (např. kategorie „**ptáci**“, „**hudební nástroje**“ apod.), využití velice obecných otázek, kdy je příbuznost dvou pojmů dána pouze jejich životností (schopností růst) nebo tím, že se jedná o předměty, jejichž existence ve světě je podmíněna lidskou tvořivostí, znemožňuje participantovi přemýšlet s předstihem nad možnými zástupci pro jednotlivé kategorie (otázky): „(...) je nepravděpodobné, že by participant trávili čas generováním příkladů *živých věcí a věcí vytvořených lidmi*, protože tyto kategorie obsahují velké množství členů.“¹⁴⁷

Participantům bylo navíc oznámeno, že se jedná o testování jejich paměťových schopností, a proto je nezbytné, aby četli i slova zobrazená na krátký čas, na která neodpovídají (primová slova). Po skončení vlastního experimentu byly děti dotazovány na přítomnost několika pojmů/slov v úloze, například „Viděl jsi v průběhu plnění úkolu slovo **orel**?“, přičemž část těchto slov se na obrazovce skutečně během experimentu objevila a část ne. Tato strategie přiměla číst participanty primová slova, ale také pomohla zakrýt skutečný cíl úkolu, čímž mohla být podpořena efektivita automatického primingu. Tento „paměťový test“ byl uskutečněn

¹⁴⁴ Odkazují přitom na tyto studie: **Seidenberg, M. S. a McClelland, J. L.**: *A distributed, developmental model of word recognition and naming*, 1989, a **Balota, D. A. a Chumbley, J.**: *Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage*, 1984.

¹⁴⁵ **McRae, K., Sa, V. R. de a Seidenberg, M. S.**: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 106. (překlad autorky)

„Rather than the lexical decision task that is standard in priming experiments, semantic decision tasks such as „is it animate?“ were used (...) whereas semantic decisions are clearly based on semantic knowledge, the lexical task affords other bases for making responses (...).“

¹⁴⁶ Tamtéž, s. 107.

¹⁴⁷ Tamtéž. (překlad autorky)

„(...) it is unlikely that subjects would spend time generating exemplars of *living things* or *things made by humans* because these categories have so many members.“

dvakrát – poprvé byli participanti na přítomnost slov zkušebně dotazováni po dokončení instruktážní části, podruhé po dokončení obou otázek ve vlastní části experimentu.

Výchozím materiálem pro primingový experiment je výše popsána databáze sémantických rysů, která byla vytvořena z dat získaných od dětských probandů z věkové kategorie 8–10 let. Z této databáze byla vygenerována rysová podobnost pojmů.

Z dat obsažených v aktualizované databázi byla připravena matice, ve které každý sloupec odpovídal jednomu pojmu (z celkem 104 pojmů) a každý řádek jednomu probandy uvedenému rysu (z celkem 2124 rysů). Číselné hodnoty uvedené v jednotlivých buňkách této matice vyjadřují produkční frekvenci rysu pro daný pojem (tj. kolik probandů uvedený rys zmínilo). Do tohoto výpočtu nebyly zařazeny rysy, které byly pro daný pojem zmíněny pouze jedním probandem (to znamená, že z matice byly vyřazeny všechny hodnoty 1, tj. byly přepsány na 0). Taxonomické rysy zůstaly v matici zahrnuté, a jsou tedy součástí rysově podobnosti. Z této matice pak byla vypočítána rysová podobnost mezi pojmy pomocí kosinu (v případě, že by kosinus dosáhl hodnoty 1, jednalo by se o pojmy identické; čím více se hodnota blížila k 0, tím méně se rysová reprezentace dvou pojmů vzájemně podobala)¹⁴⁸, a to zvláště pro jednotlivé kategorie.

Celkem bylo připraveno k pozorování 72 párů z různých sémantických kategorií, přičemž cílové slovo se zde objevilo vždy dvakrát: v jednom případě byl target spárován s primem, který byl cílovému slovu velmi podobný, v druhém případě byl target spárován s málo podobným primovým pojmem, aby mohlo dojít k porovnání vlivu míry podobnosti na sílu primingového efektu. Byly vytvořeny dvě sady – 36 párů pojmů obsahovala jedna sada párů s různou mírou podobnosti a 36 doplňujících párů s opačnou mírou podobnosti bylo zařazeno do druhé sady. Polovině účastníků pak byla prezentována slova z jedné sady a další polovině ze sady druhé.

Každá sada byla doplněna o další, výplňové páry pojmů tak, aby na polovinu všech párů (včetně pozorovaných dvojic) byla odpověď na otázku kladná (tj. 25 % otázek s odpověďmi „ano“ a „ano“, 25 % otázek s odpověďmi „ne“ a „ano“) a na druhou polovinu párů byla odpověď záporná (tj. 50 % otázek s odpověďmi „ne“ a „ne“). U obou hlavních otázek bylo navíc vždy na počátku prezentováno nejprve pět přípravných párů (s dalšími výplňovými slovy), během kterých děti získaly čas zorientovat se v úkolu. Hlavnímu experimentu předcházely dvě zkušební otázky „Má to oči?“ a „Je to sladké?“, pro každou z nich bylo připraveno šest párů pojmů. Během první části zkoušky (otázka „Má to oči?“) pracovalo dítě s pomocí examinátora, při druhé otázce („Je to sladké?“) šlo o samostatnou práci. V této zkušební fázi byl s dětmi procvičen také maskovací „paměťový experiment“.

¹⁴⁸ Způsob výpočtu podobností pojmů pomocí kosinu je detailně popsán v předchozí kapitole věnující se vzniku hlavní databáze a statistickým metodám použitým při jejím zpracování. Pro experiment byla použita data z revidované databáze. Pojmové podobnosti vypočítané z původních i aktualizovaných dat lze najít na naší webové stránce www.childrenfeatures.webnode.cz.

přítom odkazují na studie¹⁵⁰, ve kterých byla publikována zjištění, že priming je automatický a zdrojem informací jsou pouze interní lexikální procesy, pokud je interval mezi počátečním zobrazením primu a targetu maximálně 250 ms. V našem případě byl prime zobrazen po delší časový úsek s ohledem na věk participantů, protože jsme se obávali, že 200 ms není dostačující pro začínající čtenáře.

Domníváme se, že prodloužení intervalu neohrožuje správný průběh experimentu přinejmenším ze dvou důvodů. Za prvé, vědomé přemýšlení a generování potenciálních targetů bylo do značné míry znemožněno typem sémantické otázky, jak jsme uvedli již výše. Za druhé, C. S. Holderbaumová a J. F. de Sallesová¹⁵¹ zjistily, že sémantický primingový efekt se sice objevuje u dětí i v případě, kdy je SOA jen 250 ms, avšak v případě 500 ms byl tento efekt silnější.

Celý experiment byl prováděn na laptopu značky Dell s monitorem velikosti 13 palců a jako ovládací zařízení byla použita standardní drátová optická myš se dvěma tlačítky (vzhledem k tomu, že participanty myš drželi v ruce, optický senzor byl zakryt, aby nenarušoval průběh experimentu).

3.5.2.3 Postup

Rodiče potenciálních účastníků byli osloveni prostřednictvím přátel a také v adekvátních zájmových skupinách na sociální síti Facebook. Pokud rodiče projeví zájem o účast svého dítěte v jazykovém experimentu, byl jim zaslán informační dopis společně s krátkým dotazníkem a žádostí o písemný souhlas s účastí (viz přílohu 4). Zájemcům byly s ohledem na prázdninové období nabídnuty dvě možnosti, co se týkalo místa uskutečnění experimentu – rodiče mohli se svými dětmi navštívit Filozofickou fakultu Univerzity Palackého, případně byl úkol splněn v jejich vlastní domácnosti za účasti experimentátorky (tj. autorky této práce). Práce však vždy probíhala na stejném přenosném zařízení (laptopu). Každý participant byl testován individuálně a průměrná doba plnění úkolu byla 30 minut.

Děti byly za splněné dílčí úkoly oceňovány a byly motivovány k další práci drobnými odměnami v podobě samolepek. Po splnění celého úkolu dostalo každé dítě odměnu v podobě balíčku školních potřeb.

Sémantická úloha se skládala z několika částí, přičemž všechny tyto části byly založeny na stejném principu. Před samotným úkolem examinátor seznámil participanty s průběhem experimentu a s instrukcemi, jak správně postupovat při plnění úkolu. Tyto instrukce jsou v písemné podobě uvedeny v příloze 7.

Participanty byli instruováni, aby odpovídali kliknutím na „ano“ nebo „ne“ tlačítko myši pouze pro každé druhé slovo (neodpovídají tedy na slovo, které se zobrazí na velmi krátký čas, ale až na slovo následující, které se zobrazuje na delší čas). Zároveň byli nabádáni, aby si přečetli každé zobrazené slovo – to bylo zdůvodněno výše popsanou záminkou „paměťového testu“.

¹⁵⁰ Neely, J. H.: *Semantic priming effects in visual word recognition: A selective view of current findings and theories*, 1991; Den Heyer, K., Briand, K. a Dannenbring, G. L.: *Strategic factors in a lexical-decision task: Evidence for automatic and attention-driven processes*, 1983 a De Groot, A. M. B.: *Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target*, 1984.

¹⁵¹ Holderbaum, C. S. a Salles, J. F. de: *Semantic Priming Effects in a Lexical Decision Task: Comparing third Graders and College Students in two Different Stimulus Onset Asynchronies*, 2011.

Přestože byly pro experiment vybrány velice obecné a dle našeho mínění jednoduché otázky, zjistili jsme, že některým dětským participantům činí určité problémy otázka „Roste to?“ – zdá se, že „růst“ může být asociován především s rostlinami, ne vždy však s živými tvory. Proto experimentátorka před započítáním každé ze dvou hlavních fází experimentu předběžně seznámila participanty se zněním otázek a bylo uvedeno několik příkladů.

Po instruktážní části následovaly postupně jednotlivé fáze experimentu – v rámci prvních dvou otázek s šesti páry slov pro každou z nich bylo ověřeno porozumění úkolu. Třetí v pořadí byla otázka „Roste to?“ s 69 páry pojmů. Po ní následovala poslední, čtvrtá otázka „Vyrábějí to lidé?“ s 85 páry pojmů.

3.5.3 Výsledky

V následující tabulce 13 uvádíme průměrnou rychlost odpovědi participantů s ohledem na míru podobnosti mezi primovým a cílovým slovem.

Průměrný reakční čas (bez chybných odpovědí a vypršené lhůty)			
	celkový průměrný RT	RT – otázka 1	RT – otázka 2
vysoká podobnost	1728.88	1466.25	1937.6
směrodatná odchylka	803.96	615.69	871.99
nízká podobnost	1720.80	1480.6	1923.27
směrodatná odchylka	806.29	697.54	835.66

Tab. 13. Rychlost reakce (odpovědi) na danou otázku u párů s vysokou a nízkou mírou sémantické podobnosti

V tabulce 14 pak uvádíme kolik procent odpovědí z celkového množství bylo chybných, případně nebyla otázka zodpovězena včas, tj. v daném časovém limitu 5000 ms.

	celkem	otázka 1	otázka 2
vysoká podobnost	5.6 %	5.9 %	5.8 %
nízká podobnost	5.9 %	3.1 %	8.1 %

Tab. 14. Chybovost při odpovídání participanty na danou otázku

Vyhodnocení se týkalo pouze těch párů pojmů, které jsou součástí databáze sémantických rysů, díky čemuž u nich mohla být vypočítána jejich vzájemná sémantická podobnost založená na sdílených sémantických rysech. Jinými slovy, do výpočtů nebyly zařazeny výplňové páry slov. Vyřazeny byly také ty páry, u nichž účastník experimentu uvedl špatnou odpověď na položenou otázku.

Vliv míry sémantické podobnosti mezi primovým a cílovým slovem na sílu primingového efektu (tj. na rychlost reakce na cílové slovo) byl zjišťován analýzou dat pomocí smíšených regresních modelů. Závisle proměnnou byl tedy stanoven reakční čas. Byla také pozorována interakce efektu míry podobnosti a typu otázky. V této analýze však nebyly naměřeny žádné významné hodnoty. Následující tabulka 15 shrnuje výsledky této analýzy pomocí smíšeného regresního modelu.

Vzhledem k tomu, že předpokládaný vliv sémantické podobnosti nebyl zaznamenán, byly dále z dat odstraněny pro každou z otázek čtyři páry pojmů, u kterých nebyl signifikantní rozdíl v míře podobnosti mezi párem, který měl zastupovat vysokou rysovou podobnost, a párem zastupujícím vzájemnou nízkou podobnost (o omezeních týkajících se výběru pojmů z databáze s různou mírou podobnosti viz dále, v závěrečné diskusi k experimentu). Další

tabulka 16 shrnuje výsledky analýzy poté, co byly pro obě části vlastního experimentu odebrány tyto čtyři páry pojmů (tzn. čtyři páry pro otázku „Roste to?“ a čtyři páry pro otázku „Vyrábějí to lidé?“), tedy těch, u kterých nebyl rozdíl mezi nízkou a vysokou podobností výrazný. Do analýzy tak byly zahrnuty pouze ty páry pojmů, které dobře splňovaly potřebné podmínky: byl zde významnější rozdíl mezi nízkou a vysokou hodnotou podobnosti u párů vztahených ke stejnému cílovému slovu. Ani v tomto případě však nebyly zaznamenány žádné signifikantní výsledky.

Proto byly dále pozorovány vztahy mezi rychlostí reakce a dalšími proměnnými, které mohly potenciálně tuto rychlost ovlivnit. Poslední tabulka 17 zobrazuje výsledky výpočtů, ve kterých byla zohledněna délka cílových slov (počet písmen, ze kterých se cílové slovo skládá). I zde byl brán v potaz také typ otázky. Vliv vysoké podobnosti na rychlost reakce na cílové slovo byl v tomto specifickém případě zaznamenán: byla-li ve výpočtech zohledněna délka slova, efekt se objevil u cílových slov, která se skládala minimálně ze sedmi písmen, avšak pouze v případě biologických kategorií. Výsledky analýz, které jsou uvedeny ve výše uvedené tabulce 17, dokládají tento jev. Grafické znázornění rozdílných reakčních časů s ohledem na délku slova a míru podobnosti je prezentováno v následujících dvou grafech – graf 3 zobrazuje výsledky pro otázku „Vyrábějí to lidé?“ (kategorie artefaktů) a graf 4 pro otázku „Roste to?“ (pro biologické kategorie).

prediktory	log(RT+1)			log(RT)		
	odhad	střední chyba průměru	hodnota p	odhad	střední chyba průměru	hodnota p
(regresní konstanta)	7.36	0.10	<0.001	7.36	0.10	<0.001
nízká podobnost	0.00	0.02	0.809	-0.01	0.03	0.784
otázka „Vyrábějí to lidé?“	0.30	0.03	<0.001	0.29	0.04	<0.001
cílové slovo – počet písmen	0.01	0.01	0.324	0.01	0.01	0.328
cílové slovo (frekvence ve <i>Weslalexu</i>) + 1 [log]	-0.02	0.01	0.011	-0.02	0.01	0.011
cílové slovo (frekvence v <i>Intercorp12-czech</i> -titulky) [log]	-0.03	0.01	0.002	-0.03	0.01	0.002
nízká podobnost: faktor (otázka „Vyrábějí to lidé?“)				0.02	0.04	0.555
N	36 participantů			36 participantů		
	36 cílových slov			36 cílových slov		
počet pozorování	1222			1222		

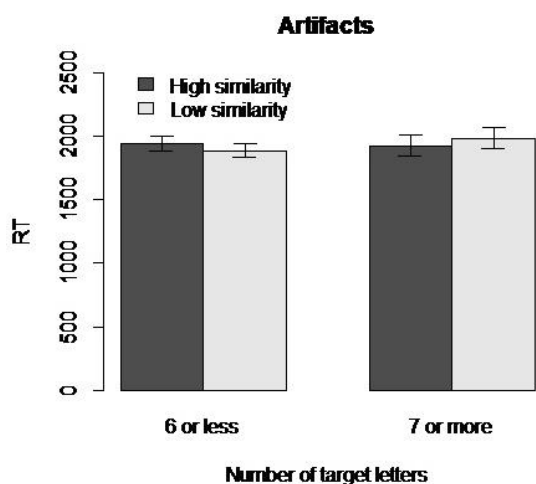
Tab. 15. Analýza dat pomocí smíšeného regresního modelu (model vlevo pro artefakty a model vpravo pro biologické kategorie)

prediktory	log(RT)			log(RT)		
	odhad	střední chyba průměru	hodnota p	odhad	střední chyba průměru	hodnota p
(regresní konstanta)	7.33	0.11	<0.001	7.33	0.11	<0.001
nízká podobnost	0.02	0.02	0.385	0.02	0.03	0.537
otázka „Vyrábějí to lidé?“	0.29	0.04	<0.001	0.29	0.05	<0.001
cílové slovo – počet písmen	0.01	0.01	0.514	0.01	0.01	0.514
cílové slovo (frekvence ve <i>Weslalexu</i>) + 1 [log]	-0.02	0.01	0.082	-0.02	0.01	0.082
cílové slovo (frekvence v <i>Intercorp12-czech</i> -titulky) [log]	-0.03	0.01	0.046	-0.03	0.01	0.046
nízká podobnost: faktor (otázka „Vyrábějí to lidé?“) podobnost x kategorie				-0.00	0.04	0.954
N	36 participantů			36 participantů		
	28 cílových slov			28 cílových slov		
počet pozorování	952			952		

Tab. 16. Analýza dat pomocí smíšeného regresního modelu po vyřazení pojmových párů s málo diferencovanou mírou podobnosti

prediktory	log(RT)			log(RT)		
	odhad	střední chyba průměru	hodnota p	odhad	střední chyba průměru	hodnota p
(regresní konstanta)	7.47	0.08	<0.001	7.46	0.07	<0.001
podobnost	-0.06	0.03	0.047	-0.03	0.02	0.149
zaokrouhlení dolů (cílové slovo 7+ písmen)	-0.18	0.10	0.066	-0.13	0.06	0.044
otázka „Roste to?“	0.24	0.07	0.001	0.30	0.03	<0.001
cílové slovo (frekvence v <i>Intercorp12-czech</i> -titulky) [log]	-0.03	0.01	0.009	-0.04	0.01	0.001
cílové slovo (frekvence ve <i>Weslalexu</i>) + 1 [log]	-0.03	0.01	0.001	-0.03	0.01	0.006
podobnost * zaokrouhlení dolů (cílové slovo 7+ písmen)	0.18	0.06	0.002	0.11	0.04	0.005
podobnost * „Roste to?“	0.06	0.04	0.168			
zaokrouhlení dolů (cílové slovo 7+ písmen) * otázka „Roste to?“	0.08	0.13	0.517			
(podobnost * zaokrouhlení dolů (cílové slovo 7+ písmen)) * otázka „Roste to?“	-0.13	0.07	0.090			
N	36 participantů			36 participantů		
	36 cílových slov			36 cílových slov		
počet pozorování	1222			1222		

Tab. 17. Analýza dat pomocí smíšeného regresního modelu – interakce míry podobnosti a délky cílového slova



Graf 3. Srovnání reakčního času u párů s nízkou a vysokou podobností mezi pojmy – s ohledem na délku slova (pro kategorie artefaktů)



Graf 4. Srovnání reakčního času u párů s nízkou a vysokou podobností mezi pojmy – s ohledem na délku slova (pro biologické kategorie)

3.5.4 Diskuse

Výsledek sémantického experimentu byl do jisté míry překvapivý, protože nebyl zjištěn předpokládaný vliv rysové podobnosti pojmů na sílu primingového efektu. Jinými slovy řečeno, rychlost reakce na cílové slovo v případě, že mu předcházelo sémanticky podobné primové slovo ze stejné sémantické kategorie, nebyla mírou rysové podobnosti ovlivněna. Nicméně byl zaznamenán vliv na rychlost reakce v případě jedné skupiny slov – tento efekt byl nalezen u pojmů z tzv. biologických kategorií, jestliže tomuto targetu předcházelo primové slovo s vysokou podobností, avšak pouze v takovém případě, kdy se slova zastupující tyto pojmy skládala ze sedmi a více písmen.

Z toho vyvozujeme, že primingový efekt jako takový se v experimentu projevil a nelze říci, že by experiment selhal na základě absence výskytu primingového jevu. Zdá se, že participanté četli primová slova, což ovlivnilo rychlost reakce na cílové slovo, i když byl tento vliv zaznamenán pouze pro specifickou skupinu slov.

Je možné, že se primingový efekt projevil především u delších slov proto, že v tomto případě došlo díky primovému slovu k otevření mentálního přístupu k sémanticky příbuzným slovům, a tudíž bylo následující slovo rozpoznáno bez nutnosti detailního čtení tohoto slova. V případě delších slov musejí dětští participanté vynaložit větší úsilí při čtení a následné identifikaci významu daného slova. Pokud slovu s velkým množstvím písmen předchází významově příbuzné slovo, může dojít k usnadnění čtení a pochopení slova následujícího. V tomto případě lze očekávat, že čím více bude rysová reprezentace primového slova sémanticky příbuzná s rysovou reprezentací cílového slova (v našem případě je míra příbuznosti vyvozena z rysové podobnosti pojmů), tím rychleji bude target zpřístupněn a projeví se to na rychlosti reakce. Vzhledem k tomu, že v případě kratších slov není proces čtení tak náročný, nemusel se zde tento efekt uplatnit.

Výskyt tohoto efektu pouze v případě pojmů z biologických sémantických kategorií může být dán větší sémantickou soudržností pojmů v rámci těchto kategorií. Navíc je možné předpokládat, že děti jsou lépe obeznámeny právě s pojmy z těchto více přirozených kategorií. Výsledky našeho experimentu se liší od výsledků získaných K. McRaem a jeho kolegy – v případě dospělých participantů byl vliv rysové podobnosti na sílu primingového efektu (respektive na rychlost reakce) zaznamenán u pojmů z kategorií artefaktů, pokud byla podobnost odvozena z rysových reprezentací založených na individuálních rysech: „Podobnost

založená na individuálních rysech předpověděla primingový efekt pro artefakty, (...) ale ne pro kategorie živých věcí (...).¹⁵²

K. McRae a kol. testovali stejnou hypotézu také pro případ, kdy byly podobnosti pojmů vypočítány z reprezentací založených na korelacích mezi rysy. V tomto případě byl sledován opačný efekt: vyšší míra podobnosti předpověděla rychlejší reakční čas pro kategorie živých věcí, ne však pro kategorie artefaktů.¹⁵³

V našem případě sice nebylo cílem srovnávat výsledky reprezentací utvořených z individuálních rysů a reprezentací, které reflektovaly vztahy mezi jednotlivými rysy, nicméně následující myšlenka K. McRae a jeho kolegů objasňuje výsledky pro oba typy rysových reprezentací (individuálních i korelačních) v obou typech kategorií (biologických i artefaktech): protože při měření korelací mezi rysy byly zaznamenány významnější vztahy v kategoriích živých objektů, v reprezentacích, které byly utvořeny z individuálních rysů, nebyla podobnost dostatečně zachycena, zatímco v případě artefaktů byly vztahy mezi rysy méně významné, a proto mohly být pro výpočet podobností mezi pojmy dostačující reprezentace vystavěné na individuálních rysech.¹⁵⁴

V našem experimentu založeném na individuálních rysech bylo tedy možné očekávat pozitivní výsledky především v kategorii artefaktů, nicméně tento předpoklad nebyl potvrzen.

V následujících odstavcích uvádíme několik možných příčin (i praktických) toho, proč se v experimentu výrazněji neprojeví předpokládané jevy.

Jak bylo zmíněno výše, významnější korelace mezi rysy byly nalezeny v případě kategorií živých objektů, které se vyskytují ve světě přirozeně, bez vlivu člověka (kategorie zvířat, rostlin apod.) – zdá se, že tyto vztahy mezi jednotlivými rysy přímo souvisejí s evolucí, zatímco v případě artefaktů, tedy objektů vyprodukovaných člověkem, bývá nejdůležitější jejich funkce a ostatní vlastnosti týkající se vzhledu apod. mohou být zcela náhodné: „Struktura živých věcí je určena geneticko-evolučními principy; rostliny a zvířata se v průběhu času vyvinuly až do jejich současné podoby. Korelované skupiny rysů se vyvinuly souběžně a staly se součástí mnoha rostlin a živočichů. (...) v případě živých věcí mohou být vztahy mezi rysy pevnější, protože pro ně platí výraznější omezení než pro artefakty, v jejichž případě bývají některé vlastnosti více nahodilé. (...) Je možné, že mentální reprezentace artefaktů obsahují méně korelačních rysů, protože lidé s nimi zacházejí jako s předměty, které jsou vyrobeny s cílem plnit jednu konkrétní funkci.“¹⁵⁵

Tento argument může být považován za vhodné vysvětlení toho, proč v našem případě nebyl nalezen vliv rysové podobnosti na primingový efekt v případě biologických kategorií – je možné, že efekt by se projevil v případě, kdy by pojmová podobnost byla vystavěna na korelačních rysech. Je však nutné hledat odpovědi na otázku, proč nebyl zaznamenán žádný

¹⁵² **McRae, K., Sa, V. R. de a Seidenberg, M. S.:** *On the Nature and Scope of Featural Representations of Word Meaning*, 1997. S. 108. (překlad autorky)

„Similarity in terms of individual features predicted priming effects for artifacts, (...) but not for living things (...).“

¹⁵³ Tamtéž.

¹⁵⁴ Tamtéž, s. 110.

¹⁵⁵ Tamtéž, s. 111–112. (překlad autorky)

„The structure of living things is determined by genetic-evolutionary principles; plants and animals have evolved over time into their present form. Correlated sets of features have evolved in parallel and become instantiated in a number of plants and animals. (...) the features of living things may be more densely intercorrelated because the constraints tend to be more arbitrary for artifacts than for living things. (...) It may be that mental representations of artifacts contain fewer correlated features because people treat them as items that are designed to perform a single function.“

efekt v případě artefaktových kategorií, u kterých lze podobnost dostatečně zachytit pomocí individuálních rysů.

Pokud porovnáme naši studii a studii K. McRae a jeho kolegů¹⁵⁶, domníváme se, že odlišný výsledek našich experimentů může být dán rozdílným věkem participantů, respektive odlišnou kompetencí na lexikálně-sémantické jazykové úrovni, spíše než samotnou skutečností, že byl experiment aplikován v jiném jazykovém prostředí (anglickém a českém).

Na závěr zmiňme další možná vysvětlení výsledků experimentu, která jsou spíše pragmatického charakteru.

Za prvé je nutné připomenout, že rysově podobné dvojice byly vygenerovány z databáze sémantických rysů, což vedlo ke značnému omezení možností – pro každou sémantickou kategorii jsme měli k dispozici pouze osm pojmů, které bylo nutné vhodně spárovat tak, abychom dostali čtyři dvojice s různou mírou podobnosti, přičemž předběžné intuitivní spárování podle podobnosti nekorespondovalo s rysovou podobností, která byla vygenerována z databáze. V některých případech navíc nebyly hodnoty kosinu dostatečně rozptýlené na celé škále 1–0 (v kategorii se například objevilo několik párů s vysokou podobností, ale nedostatečné množství párů s nízkou podobností).

Při párování pojmů muselo být dodrženo několik zásadních pravidel. Vybrané cílové slovo muselo být vždy spárováno se slovem, se kterým sdílelo vysokou podobnost, a zároveň se slovem, se kterým sdílelo podobnost nízkou. Jeden pár byl součástí sady 1 a pár s opačnou mírou podobnosti byl součástí sady 2 – to znamená, že jednomu participantovi byl daný target představen v průběhu experimentu pouze jednou. Zároveň se v rámci jedné sady mohl vyskytnout každý pojem pouze jednou, aby na základě opakovaného zobrazení jednoho pojmu nedocházelo k narušení podmínek nutných pro správný průběh experimentu. Jestliže byly tyto podmínky splněny, nezbývalo mnoho kombinačních možností při párování pojmů. V rámci některých kategorií dokonce nebylo možné vytvořit tři páry s různou mírou podobnosti, ale pouze dva. Pouze pro kategorii „zelenina“ byly vytvořeny páry čtyři. Pro ilustraci uvádíme v tabulce 18 příklady párů ze dvou různých sémantických kategorií, které byly pro experiment vybrány, společně s mírou jejich sémantické podobnosti (hodnota kosinu se nachází mezi 1–0). Kompletní seznam pojmových párů s číselnou hodnotou míry podobnosti lze najít v příloze 12.

	sada 1			sada 2		
	prime	target	míra sémantické podobnosti	prime	target	míra sémantické podobnosti
ZELENINA	CIBULE	CUKETA	0.459 (nízká)	KVĚTÁK	CUKETA	0.801 (vysoká)
	MRKEV	HRÁŠEK	0.340 (nízká)	SALÁT	HRÁŠEK	0.701 (vysoká)
	KVĚTÁK	FAZOLE	0.784 (vysoká)	CIBULE	FAZOLE	0.414 (nízká)
	SALÁT	OKURKA	0.742 (vysoká)	MRKEV	OKURKA	0.411 (nízká)
NÁBYTEK	SEDAČKA	ŽIDLE	0.672 (vysoká)	LAMPA	ŽIDLE	0.101 (nízká)
	VĚŠÁK	POHOVKA	0.089 (nízká)	SEDAČKA	POHOVKA	0.750 (vysoká)
	LAMPA	SKŘÍŇ	0.077 (nízká)	DVEŘE	SKŘÍŇ	0.249 (vysoká)

Tab. 18. Příklady párů pojmů a míra jejich sémantické podobnosti

¹⁵⁶ McRae, K., Sa, V. R. de a Seidenberg, M. S.: *On the Nature and Scope of Featural Representations of Word Meaning*, 1997.

Druhým možným vysvětlením výsledku je prostá skutečnost, že metoda experimentu není dostatečně citlivá pro případ, kdy je úkol plněn dětskými účastníky. Přestože byl prodloužen časový úsek, po který se zobrazovalo primové slovo (ve srovnání s délkou zobrazení u dospělých účastníků experimentu v původní studii), nemůže být vyloučena možnost, že úloha byla pro děti příliš náročná. Děti musely být schopné rychle číst zobrazovaná slova a zároveň přemýšlet o jejich významu – to mohlo způsobit, že děti nestíhaly číst primová slova, a tudíž primingový efekt nemohl vůbec nastat. Některé z dětí po dokončení úlohy ostatně samy přiznávaly, že čtení primů jim činilo potíže. Na druhou stranu nelze považovat nedostatečnou čtenářskou kompetenci za zásadní problém – pokud by čtení dětem skutečně činilo významné takové potíže, promítlo by se to nejspíše také v rozdílných reakčních časech u cílových slov s různým počtem slabik – žádný rozdíl v tomto případě však zaznamenán nebyl. Přestože v experimentu nedošlo k potvrzení předkládané hypotézy (tj. míra rysově podobnosti primového a cílového slova bude mít v rychlé sémantické úloze vliv na rychlost reakce v rámci primingového efektu), zdá se, že na základě námi získaných výsledků nelze jednoznačně tvrdit, že tato hypotéza není správná. Výše uvedené argumenty zvažují (respektive přímo zpochybňují) citlivost metody v případě dětských účastníků, stejně jako dostatečnost množství materiálu, který databáze obsahuje. Považujeme tedy za vhodné dále rozšiřovat obsah databáze sémantických rysů získaných od českých dětí a ověřit předloženou hypotézu na základě většího množství dat, aby mohla být výsledkům tohoto experimentu přisouzena větší váha. Naše poznatky, které uvádíme výše, mohou posloužit jako podklad pro další studie tohoto charakteru.

4. Závěr

V této práci představujeme náš psycholingvistický výzkum, jehož cílem bylo vytvořit databázi sémantických rysů z dat získaných od dětí věku 8–10 let a v následujícím kroku provést experiment, který ověřuje využití rysových reprezentací v rychlých sémantických úlohách.

V první části práce předkládáme teoretický přehled, přičemž se – s ohledem na mezioborový přesah tématu – pokoušíme představit všechny relevantní poznatky o pojmech, jejich kategorizaci a jejich mentální reprezentaci skládající se mimo jiné z tzv. sémantických rysů. Uvádíme proto jak filozofické, tak také lingvistické (sémantické) a psychologické přístupy.

V úvodu teoretické části se pokoušíme nastínit důležitost rozšiřování psycholingvistického výzkumu v českém prostředí a význam předkládané práce. Nejobsáhlejší podkapitola se pak věnuje už výše zmíněnému tématu, tj. pojmům a jejich mentální kategorizaci, která je podle mnohých teorií založena na podobnosti a odlišnosti na úrovni sémantických rysů, z nichž se tyto pojmy skládají. Dále se věnujeme významu zkoumání rysových reprezentací a využití modelů těchto reprezentací utvořených na základě empirického materiálu získaného zpracováním dat sesbíraných od velkého počtu jednotlivců. Poukazujeme také na to, že pojmová struktura je podle posledních vědeckých poznatků mnohem sofistikovanější a nelze ji popsat pomocí pouhého souboru sémantických rysů, a to ani ve chvíli, kdy neopomeneme respektovat vztahy mezi těmito rysy. Existuje však předpoklad, že v každodenním rychlém rozhodování o objektech, které nás obklopují, jejich významech a vztazích mezi nimi příliš nevyužíváme komplikovanější znalosti – zdá se, že tyto základní úsudky jsou naopak postaveny právě na rysových reprezentacích pojmů a na jejich vzájemném porovnávání. Doklady této hypotézy představujeme na konkrétních pracích – poukazujeme na rozdílné mentální postupy u úloh, které vyžadují rychlé rozhodování o pojmech a jejich významech, a u těch, ve kterých není prostor k úvahám časově limitován. Vzhledem k tomu, že se v rámci našeho výzkumu pokoušíme ověřit využití rysových reprezentací pojmů v rychlé sémantické úloze, představujeme v další podkapitole také (sémantický) priming, protože právě na tomto jevu je předkládaný experiment založen a hraje v něm tedy významnou roli. V závěru teoretické části se stručně zmiňujeme o osvojování pojmů u dětí – ve většině případů se psycholingvistický výzkum na lexikální a sémantické úrovni zaměřuje na rozšiřování dětského slovníku (rozdávání slovní zásoby o nová slova), méně často je však zkoumán dětský slovník do hloubky; i z toho důvodu obvykle stojí v centru zájmu práce s dětmi předškolního věku, kdy dochází k největším a nejrychlejším změnám ve vývoji řeči. Předkládaná práce tedy přináší nové poznatky a nová data z méně tradiční oblasti tohoto oboru.

V empirické části představujeme námi vytvořenou databázi sémantických rysů, která byla sestavena z dat získaných od 381 českých dětí ve věku 8–10 let. Před oficiálním sběrem dat byl v rámci pilotní studie nasbírán menší objem dat, na kterém jsme mimo jiné ověřili několik možných metod sběru a vybrali jsme tu metodu, která se nám jevila jako nejvhodnější. Celá pilotní studie je představena v praktické části této práce; dále pak popisujeme postup při hlavním sběru dat, celý proces tvorby databáze sémantických rysů, její zpracování pomocí vybraných statistických metod a její interpretaci. Celá databáze, včetně statistické analýzy dat, je publikována online, čímž ji nabízíme zájemcům z řad vědecké komunity k využití v dalších (nejen) psycholingvistických výzkumech. Zároveň věříme, že publikovaný materiál může být přínosný také pro pedagogy či rodiče. Data z této databáze byla využita při zkoumání významnosti rysových reprezentací a ověřování jejich využití v sémantické úloze vyžadující rychlou reakci. V této úloze byl využíván tzv. sémantický primingový efekt: sémantická blízkost dvou pojmů/slov ovlivňuje rychlost reakce na cílový pojem. Předpokládali jsme, že rychlost reakce může ovlivnit také míra rysové podobnosti mezi dvěma pojmy – čím si budou rysové

reprezentace dvou pojmů podobnější, tím bude naměřen kratší reakční čas. Překvapivě nebyl žádný významný vliv rysové podobnosti na rychlost reakce zaznamenán, což bylo v rozporu s výsledky získanými autory naší vzorové studie. V případě biologických kategorií byla sice zaznamenána rychlejší reakce na cílová slova, která byla spárována s pojmy s vyšší podobností, avšak pouze v případech, kdy tato cílová slova obsahovala minimálně sedm písmen. V závěru části práce věnující se tomuto experimentu předkládáme diskusi, ve které uvádíme možná vysvětlení toho, proč docházelo k urychlení reakce pouze v tomto jediném specifickém případě. Pokoušíme se také objasnit, proč se nám nepodařilo ověřit předkládanou hypotézu a proč nebyly naše výsledky shodné s těmi ve vzorové studii. Přestože nebyl zaznamenán předpokládaný efekt, věříme, že uskutečněný experiment i jeho výsledek má svůj význam a bude přínosem pro českou psycholingvistiku – například jako inspirace a motivace pro další výzkum tohoto charakteru.

5. Seznam použité literatury

AHLGREN, P., JARNEVING, B. a ROUSSEAU, R.: Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American society for information science and technology*. 2003; **54**(6), s. 550–560.

ALTMANN, Gerry T. M. *Výstup na babylonskou věž: otázky jazyka, mysli a porozumění*. Praha: Triáda, 2005.

APA Dictionary of psychology. Dostupné z: www.dictionary.apa.org/stimulus-onset-asynchrony

APA Peeps. Dostupné z: www.apa.org/pubs/highlights/peeps/issue-33

BALOTA, David A. a CHUMBLEY, James I. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of experimental psychology: human perception and performance*. 1984; **10**(3), s. 340–357.

BARBAROTTO, Riccardo, LAIACONA, Marcella a CAPITANI, Erminio. Does sex influence the age of acquisition of common names? A contrast of different semantic categories. *Cortex*. 2008; **44**(9), s. 1161–1170.

BARSALOU, Lawrence W. Deriving categories to achieve goals. In **BOWER, Gordon H. (ed.)**. *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press. 1991; 27, s. 1–64.

BATTIG, William F. a MONTAGUE, William E.: Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of experimental psychology*. 1969; **80**(3/2), s. 1–46.

BLOOM, Paul. *Jak se děti učí významu slov*. 1. čes. vyd. Praha: Karolinum, 2015.

ČERNÝ, Jiří. *Dějiny lingvistiky*. 1. soubor. vyd. Olomouc: Votobia, 1996.

ČERNÝ, Jiří. *Úvod do studia jazyka*. Olomouc: Rubico, 1998.

ČERNÝ, Jiří a HOLEŠ, Jan. *Sémiotika*. Praha: Portál, 2004.

ČERVENÁ Vlasta et al.: *Slovník spisovné češtiny pro školu a veřejnost*. 4. vyd. Praha: Academia, 2005.

DeCASPER, Anthony a SPENCE, Melanie J.: Prenatal maternal speech influences newborns' perception of speech sounds. *Infant behavior and development*. 1986; **9**(2), s. 133–50.

EGGHE, Leo a LEYDESDORFF, Loet: The relation between Pearson's correlation coefficient r and Salton's cosine measure. *Journal of the American society for information science and technology*. 2009; **60**(5), s. 1027–1036.

ERHART, Adolf. *Úvod do obecné a srovnávací jazykovědy*. Praha: SPN, 1973.

EYSENCK, Michael W. a KEANE, Mark T. *Kognitivní psychologie*. Praha: Academia, 2008.

GROOT, Anette M. B. de. Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *The quarterly journal of experimental psychology*. 1984; **36A**(2), s. 253–280.

HEYER, Ken den, BRIAND, Kevin a DANNENBRING, Garry L. Strategic factors in a lexical-decision task: Evidence for automatic and attention-driven processes. *Memory & Cognition*. 1983; **11**(4), s. 374–381.

HOLDERBAUM, Candice S. a SALLES, Jerusa F. de. Semantic priming effects in a lexical decision task: Comparing third graders and college students in two different stimulus onset asynchronies. *The Spanish journal of psychology*. 2011; **14**(2), s. 589–599.

Internetová jazyková příručka [online] (2008–2021). Praha: Ústav pro jazyk český AV ČR, v. v. i. Dostupné z: <https://prirucka.ujc.cas.cz/>

KARLÍK, Petr, Marek NEKULA a Jana PLESKALOVÁ, eds. *Nový encyklopedický slovník češtiny* [online]. Praha: NLN, Nakladatelství Lidové noviny, 2016. Dostupné z: <https://www.czechency.org/>

KEIL, Frank C. *Concepts, Kinds, and Cognitive Development*. Massachusetts: The MIT Press, 1989.

KONEČNÁ, Kristýna a VAŇKÁTOVÁ, Kristýna. A database of semantic features for chosen concepts (Attested in 8- to 10-year-old Czech pupils). *Topics in linguistics*. 2017; **18**(1), s. 71–97.

KONEČNÁ, Kristýna a VEČEŘOVÁ, Lucie. Baterie sémantických rysů: data získaná od dětí z třetích tříd základních škol – zpráva z pilotního sběru dat. *Studie z aplikované lingvistiky*. 2016; **7**(2), s. 48–71.

KONEČNÁ, Kristýna a SMOLÍK, Filip. *The influence of featural similarity strength on the priming effect in a speeded task*. 2021 (v recenzním řízení)

KŘEN, Michal et al.: *Korpus SYN, verze 3*, 2014. Dostupné z: <https://kontext.korpus.cz>

Labels. Dostupné z: www.labels.ff.cuni.cz

LAKOFF, George. *Ženy, oheň a nebezpečné věci: co kategorie vypovídají o naší mysli*. Praha: Triáda, 2006.

LOTKO, Edvard. *Slovník lingvistických termínů pro filology*. 3. vyd. (dotisk). Olomouc: Univerzita Palackého, 2005.

McRAE, Ken, CREE, George S., SEIDENBERG, Mark S. a McNORGAN, Chris. Semantic Feature Production Norms For a Large Set of Living and Non-living Things. *Behavior Research Methods, Instruments & Computers*. 2005; **37**(4), s. 547–559.

McRAE, Ken, SA, Virginia R. de a SEIDENBERG, Mark S. On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*. 1997; **126**(2), s. 99–130.

MEDIN, Douglas L. Concepts and conceptual structure. *American psychologist*. 1989; **44**(12), s. 1469–1481.

MEHLER, Jacques et al. A precursor of language acquisition in young infants. *Cognition*. 1988; **29**, s. 143–178.

MURPHY, Gregory L.: *The big book of concepts*. Massachusetts: The MIT Press, 2002.

MURPHY, Gregory L. a MEDIN, Douglas L. The Role of Theories in Conceptual Coherence. *Psychological review*. 1985; **92**(3), s. 289–316.

NAGY, Marek. Produkce sémantických rysů u dospělých mluvčích. In **KESEELOVÁ, Jana, IMRICHOVÁ, Mária a OLOŠTIK, Martin (eds.)**. *Registre jazyka a jazykovedy (I)*. Prešov: Filozofická fakulta Prešovskej university v Prešove. 2014; s. 237–249.

NEELY, James H. Semantic priming effects in visual word recognition: A selective view of current findings and theories. In **BESNER Derek a HUMPHREYS, Glyn W. (eds.)**. *Basic processes in reading*. Hillsdale, NJ: Erlbaum. 1991; s. 264–336.

PAIVIO, Alan. Dual coding theory and the mental lexikon. *Words and Their Meaning: A Deep Delve from Surface Distribution Into Underlying Neural Representation*. **5**(2), s. 205–230.

POSNANSKY, Carla J. Category norms for verbal items in 25 categories for children in grades 2–6. *Behavior research methods & Instrumentation*. 1978; **10**(6), s. 819–832.

POSNER, Michael I. a SNYDER, Charles R. R. Facilitation and Inhibition in the Processing of Signals. In **RABBITT P. M. A. a DORNIC Stanislav (eds.)**. *Attention and Performance 5*. London: Academic Press. 1975; s. 669–682.

POTTIER, Bernard. Vers une sémantique moderne. In **HANKS, P. (ed.)**. *Lexicology II*. 2009; s. 108–139.

PRŮCHA, Jan. *Dětská řeč a komunikace: poznatky vývojové psycholingvistiky*. Praha: Grada, 2011.

RAMSCAR, Michael. Wittgenstein and the representation of psychological categories. In **RAMSCAR, M., HAHN, U., CAMBOUROPOLOS, E. a PAIN H. (Eds.)**. *Proceedings of simcat 1997: An interdisciplinary workshop on similarity and categorisation*. Edinburgh: Department of Artificial Intelligence, University of Edinburgh. 1997, s. 205–211.

ROSEN, Alexandr, VAVŘÍN Martin a ZASINA, Adrian J.: *Korpus Intercorp – čeština, verze 12*; 2019. Dostupné z: <https://kontext.korpus.cz>

ROSCH, Eleanor a MERVIS, Carolyn B. Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*. 1975; **7**(4), s. 573–605.

SALTON, Gerard a MCGILL, Michael J.: *Introduction to modern information retrieval*. New York: McGraw-Hill, 1983.

SAUSSURE, Ferdinand de. *Kurs obecné lingvistiky*. 2. vyd. Praha: Academia, 1996.

SEDLÁKOVÁ, Miluše. *Vybrané kapitoly z kognitivní psychologie: mentální reprezentace a mentální modely.* Praha: Grada Publishing, 2004.

SEIDENBERG, Mark S. a McCLELLAND, James L. A distributed, developmental model of word recognition and naming. *Psychological review.* 1989; **96**(4), s. 523–568.

SMITH, Edward E. a MEDIN, Douglas L.: *Categories and concepts.* Harvard University Press, 1981.

SMOLÍK, Filip. Psycholingvistika a čeština: některá slibná témata. *Naše řeč.* 2009; **92**(5), s. 240–251.

STERNBERG, Robert J. *Kognitivní psychologie.* Vyd. 2. Praha: Portál, 2009.

VÁGNEROVÁ, Marie. *Kognitivní a sociální psychologie žáka základní školy.* Praha: Karolinum, 2001.

VÁGNEROVÁ, Marie. *Základy psychologie.* Praha: Karolinum, 2004.

VINSON, David, P. a VIGLIOCCO, Gabriella. Semantic feature production norms for a large set of objects and events. *Behavior Research Methods.* 2008; **40**(1), Psychology Module, s. 183–190.

Weslalex. Dostupné z: <http://spell.psychology.wustl.edu/weslalex/>

WITTGENSTEIN, Ludwig. *Filosofická zkoumání.* Praha: Filozofický ústav AV ČR, 1993.

6. Přílohy

6.1 Příloha 1 (Přehled kategorií a pojmů a frekvence jejich výskytu ve vybraných korpusech)

KATEGORIE	POJEM	Weslalex (g3u)	Inter corp12 czech (ARF)	Inter corp12 czech subtitl. (ARF)	PÁROVÝ POJEM	Weslalex (g3u)	Inter corp12 czech (ARF)	Inter corp12 czech subtitl. (ARF)	
BIOLOGICKÉ KATEGORIE	OVOCE	VIŠEŇ	0	91.44	10.93	TŘEŠEŇ	30.12	174.74	23.87
		MANDARINKA	--	86.57	13.87	POMERANČ	16.58	726.62	83.5
		JAHODA	150.34	486.29	86.41	MALINA	6.03	327.94	26.5
		JABLKO	94.04	2003.76	259.29	ANANAS	5.53	214.52	38.96
	ZELENINA	OKURKA	0	514.94	74.37	CUKETA	--	27.94	3.86
		FAZOLE	5.53	649.86	161.52	HRÁŠEK	20.57	264.22	52.95
		KVĚTÁK	--	104.38	10.06	SALÁT	11.06	1263.25	290.88
		CIBULE	0	837.63	115.96	MRKEV	29.39	438.39	72.41
	ROSTLINY	SEDMIKRÁSKA	6.03	127.24	15.22	KOPRETINA	0	97.36	6.96
		BOROVICE	29.39	643.17	37.99	LÍPA	82.1	529.57	18.2
		RŮŽE	46.19	2312.04	340	TULIPÁN	18.08	236.29	32.39
		LEKNÍN	12.05	107.44	1.99	ČTYŘLÍSTEK	6.03	73.14	11.67
	ČÁSTI TĚLA	NOHA	524.80	36369.23	3661.33	RUKA	1129.85	106839.59	8214.78
		PAŽE	84.75	6441.72	199.45	STEHNO	6.03	1938.31	115.06
		UCHO	320.03	12911.1	1038.44	NOS	464.68	11103.76	1160.03
		SRDCE	319.40	27011.11	3967.37	ZUB	109.71	11103.16	1212.19
PTÁCI	OREL	16.58	954.77	226.06	SUP	0	442.15	77.77	
	RACEK	5.53	522.86	43.91	HOLUB	50.41	997.19	138.06	
	HUSA	44.14	1241.19	150.71	LABUŤ	6.03	479.71	63.88	
	PAPOUŠEK	16.58	543.30	80.48	SOVA	6.03	525.15	73.25	
SAVCI	KRÁLÍK	6.03	1561.69	316.48	ZAJÍC	27.64	781.1	89.58	
	KOČKA	218.92	5231.70	1237.27	TYGR	71.86	1040.9	261.32	
	JELEN	64.18	651.72	115.83	KŮŇ	160.78	10601.27	1590.71	
	DELFIN	0	332.64	90.52	OPICE	94.7	1736.86	503.32	
ARTEFAKTY	NÁBYTEK	SEDAČKA	0	403.62	85.13	ŽIDLE	96.07	7657.75	393.73
		POSTEL	112.48	19038.56	2505.3	POHOVKA	6.03	1913.83	99.8
		SKŘÍŇ	61.71	3332.46	275.3	VĚŠÁK	5.53	460.36	17.98
		LAMPA	5.53	3177.8	206.09	DVEŘE	266.44	44355.07	4052.07
	NÁŘADÍ	LOPATA	30.12	896.99	105.75	RÝČ	24.1	220.12	11.49
		KOŠTĚ	22.11	640.61	67.41	HRÁBĚ	12.05	153.99	13.06
		KLADIVO	5.53	1143.67	156.63	SEKERA	103.7	833.08	105.85
		KLEŠTĚ	0	584.64	57	PILA	61.71	1514.35	200.95
	HUDEB. NÁSTROJE	KLAVÍR	0	908.71	127.32	VARHANY	11.06	348.98	18.29
		HOUSLE	5.53	694.92	98.65	KYTARA	0	830.87	230.08
		TRUMPETA	0	243.12	41.27	FLÉTNÁ	6.03	402.46	64.38
		TRIANGL	0	26.9	5.69	HARFA	0	216.61	28.54
	OBLEČENÍ	KLOBOUK	79.91	4694.49	603.5	ČEPICE	11.06	2575.96	234.68

	SUKNĚ	30.12	2653.96	225.61	ŠATY	--	11293.83	1545.68
	TRIKO	0	424.06	129.19	SVETR	5.53	1352.65	178.66
	BOTA	191.32	9416.3	1529.89	PYŽAMO	22.11	820.45	113.12
DOPR. PROSTŘEDKY	TROLEJBUS	0	60.06	3.55	AUTOBUS	16.58	3486.56	896.11
	VRTULNÍK	--	716.82	273.67	LETADLO	250.43	6002.21	1662.82
	AUTO	143.11	21919.64	6501.91	TRAKTOR	0	521.63	64.2
	VLAK	46.19	6108.7	1093.68	LOĎ	185.32	12263.9	2475.09
ZBRANĚ	PISTOLE	0	3420.22	828.01	REVOLVER	12.05	853.27	93.63
	MEČ	106.54	2494.02	619.7	DÝKA	0	816.66	114.83
	PRAK	0	197.24	30.64	LUK	6.03	738.16	157.04
	OŠTĚP	12.05	335.46	51.11	DĚLO	6.03	3218.63	406.3
KUCHYŇSKÉ NÁČINÍ	SKLENICE	0	4226.72	368.11	HRNEK	47.22	989.93	99.11
	NABĚRAČKA	0	139.95	13.42	LŽÍCE	37.04	994.01	83.81
	HRNEC	113.81	1377.3	105.31	KONVICE	0	667.86	49.4
	STRUHADLO	0	42.66	3.97	VAŘEČKA	22.11	80.78	3.19

6.2 Příloha 2 (Informační dopis ředitelům vybraných základních škol – hlavní sběr dat)

Informace o studii a prosba o spolupráci

Osvojování jazyka: produkce sémantických rysů se zaměřením na vybrané aspekty osvojování významu u dětí školního věku

Vážená paní ředitelko / Vážený pane řediteli,

rádi bychom Vás touto cestou požádali o spolupráci na výzkumu *Osvojování jazyka: produkce sémantických rysů se zaměřením na vybrané aspekty osvojování významu u dětí školního věku*. V první fázi tohoto projektu, v níž bychom chtěli spolupracovat s Vaší školou, se zabýváme sběrem tzv. sémantických rysů. Jde o nejmenší významové jednotky slov, např. pojem *auto* lze rozdělit na rysy: "je dopravní prostředek", "má čtyři kola", "slouží k přepravě osob" atd. Data získaná tímto experimentem (děti písemně generují rysy k vybraným pojmům) budou dále využita a zpracována: budou zjišťovány rozdíly mezi dívkami a chlapci, dojde ke srovnání dětí s dospělými, data budou také součástí návazného výzkumu věnovaného čtenářské gramotnosti u dětí.

Děti na začátku výzkumu dostanou jednoduché instrukce, jak mají při popisování slov postupovat, a následně budou vyzvány, aby samostatně vyplnily připravené archy, které budou obsahovat kolem deseti slov. O experimentu budou informováni rodiče testovaných dětí a budou také požádáni o souhlas účasti jejich dítěte ve výzkumu. Jedno testování v jedné školní třídě je připraveno tak, aby trvalo maximálně jednu vyučovací hodinu. Důležité je, že nechceme testovat individuální schopnosti dětí, naopak hledáme to, co by mohlo být charakteristické či naopak kontrastní v rámci sledované dětské skupiny.

Do výzkumu potřebujeme zapojit děti ve věku 8–9 let, tedy žáky třetích tříd, a to alespoň 20 dětí (jedna třída). Administrace experimentu by měla proběhnout **v dubnu 2015**. Naše přítomnost na Vaší škole by zabrala jen jednu vyučovací hodinu, v níž by sběr probíhal. Experiment může probíhat v běžném vyučovacím dni přímo ve třídě.

Výzkum je součástí disertačních prací aktuálně zpracovávaných na katedře Obecné lingvistiky Filozofické fakulty Univerzity Palackého v Olomouci, konkrétně disertací Mgr. Kristýny Konečné a Mgr. Lucie Večeřové. Výzkum je dále podpořen grantem IGA téže univerzity. Supervizorem projektu je Mgr. Marek Nagy, Ph.D.

Budeme vděční za pomoc a umožnění administrace experimentu.

/podpisy autorů, kontakty/

6.3 Příloha 3 (Informační dopis a dotazník pro rodiče – hlavní sběr dat)

Informace o studii a prosba o spolupráci

Osvojování jazyka: produkce sémantických rysů se zaměřením na vybrané aspekty osvojování významu u dětí školního věku

Vážení rodiče,

dovolujeme si Vás požádat o souhlas se zařazením Vašeho dítěte do výzkumu *Osvojování jazyka: produkce sémantických rysů se zaměřením na vybrané aspekty osvojování významu u dětí školního věku*, realizovaného na Filozofické fakultě Univerzity Palackého v Olomouci. V následujícím dopise bychom Vás rádi informovali o tom, co bude výzkum obnášet a čemu se věnuje. Připojením svého podpisu na přiložený formulář vyjádříte souhlas s účastí Vašeho dítěte.

Výzkum je zaměřen na produkci sémantických rysů u dětí. Data, která experimentem získáme, budou předmětem rozboru, v němž dojde například ke srovnání rozdílů mezi děvčaty a chlapci, srovnání dětí a dospělých; budou také součástí návazného výzkumu věnovaného čtenářské gramotnosti u dětí a experimentů zaměřujících se na využití sémantických rysů ve vybraných sémantických úlohách.

V připraveném experimentu sledujeme, jakým způsobem a jak zevrubně děti chápou význam jednotlivých slov. Zúčastněné děti budou popisovat vybraná slova a určovat tak jejich charakteristické významové rysy. Např. *auto* "je stroj" a "má čtyři kola". Patří ale podle dětí mezi auta například také „dodávka“? Je při popisu „psa“ důležité zmínit například to, že má čtyři tlapy, ocas a čumák, nebo se děti spokojí s tím, že je to zvíře, které štěká? Po získání dat budeme sestavovat „katalog“ (tzv. baterii), jenž bude přehledně ukazovat, které z významových rysů daná slova charakterizují více a které naopak méně (v dětském pojetí), v jakých vzájemných vztazích sledované pojmy a jejich rysy jsou apod. Naším záměrem není sledování individuálních schopností a odlišností dětí ve školní třídě. Získaná data by měla ukázat, co je obecné pro všechny děti daného věku.

Délka jednoho experimentu odpovídá jedné školní vyučovací hodině, ve které děti samostatně vyplní připravený arch. Testování proběhne na půdě školy, kterou navštěvuje Vaše dítě, po dohodě s personálem a vedením školy. Účast dítěte je dobrovolná – nebude-li se chtít experimentu zúčastnit, nebude samozřejmě muset.

Rádi bychom Vás požádali o vyplnění krátkého dotazníku a také o souhlas, aby se Vaše dítě mohlo zúčastnit tohoto výzkumu. Budeme moc vděční za pomoc a umožnění administrace experimentu.

Vyplněním a podepsáním přiloženého formuláře dáváte souhlas k účasti Vašeho dítěte ve výzkumu a následné publikaci dat. Data budou zpracována a zveřejněna zcela anonymně, v souladu s platnou zákonnou úpravou. Data nebudou poskytována žádné třetí osobě.

Vyplněný dotazník prosím předejte třídní učitelce svého dítěte. Tento dopis uchovejte pro svou potřebu.

Projevíte-li zájem, rádi Vás budeme informovat o výsledcích výzkumu. Děkujeme Vám za spolupráci!

/podpisy autorů, kontakty/

Dotazník

1. Je Vaše dítě **bilingvní** (tj. ovládá **aktivně alespoň dva jazyky, rozumí jim, mluví jimi** – zpravidla proto, že každý rodič dítěte je jiné národnosti/z jiného jazykového prostředí, např. dítě přijímá od matky češtinu a od otce němčinu)?

Prosím zakroužkujte: **ANO/NE**

Uveďte prosím rodný jazyk/rodné jazyky dítěte:

2. Jaké jazyky se Vaše dítě **učí**? Ve škole, v kroužku, doma atd. **Prosím, vypište je:**

Místo pro případnou poznámku:

Prohlášení rodiče o souhlasu s účastí dítěte ve výzkumu

Potvrzuji, že jsem se seznámil/a s obsahem dopisu Informace o studii a prosba o spolupráci, který poskytuje informace o výzkumu, prováděného katedrou Obecné lingvistiky a teorie komunikace FF UP. Souhlasím s účastí svého dítěte ve výzkumu, který je v dopisu popsán a který proběhne v průběhu **března 2015** na půdě školy.

Beru na vědomí, že účast dítěte je dobrovolná a svůj souhlas mohu kdykoli před, ale i v průběhu výzkumu odvolat.

/podpis rodičů/

datum vyplnění tohoto formuláře a souhlasu:

jméno dítěte:

datum narození dítěte:

6.4 Příloha 4 (Informační dopis a dotazník pro rodiče – sémantická úloha)

Vážení rodiče,

velice si vážíme toho, že jste se rozhodli zúčastnit se se svými dětmi našeho experimentu. Vaše účast pomůže rozšířit obsah našeho psycholingvistického výzkumu, který se věnuje jazyku, dětské řeči a významu slov. Výsledky naší práce – obsahující rovněž data získaná od Vašeho dítěte – by měly být v následujícím roce publikovány v zahraničním odborném časopise.

Pro jazykový experiment hledáme děti, které ve školním roce 2019/2020 navštěvovaly 3. třídu ZŠ a je jim tedy 8–9 (případně 10) let. Pandemie COVID-19 bohužel zkomplikovala naše plány nasbírat data přímo ve školách. O to víc si vážíme Vaší ochoty a věříme, že děti potěší odměna za účast v podobě balíčku školních potřeb.

A co vlastně budou děti muset udělat?

Jedná se o jednoduchý jazykový experiment, ve kterém děti budou číst z obrazovky česká slova a odpovídat na námi položené, velice snadné otázky.

Tyto výše zmíněné informace jsou však pouze pro rodiče, tedy pro Vás. Pokud Vás můžeme poprosit, dětem zatím nic neprozrazujte – můžete se jen zmínit, že se zúčastní „jazykové hry“. Přesné pokyny jim pak zadáme na začátku experimentu osobně.

Předpokládáme, že samotný experiment bude trvat asi 30–45 minut, celkově Vás tedy zdržíme nejdéle na jednu hodinu.

Je důležité zdůraznit, že získaná data budou zpracována anonymně a výsledky nebudou interpretovány pro jednotlivce, nýbrž pro celou skupinu účastníků – v práci se neobjeví jména Vašich dětí, pouze jejich věk, pohlaví a oblast, ve které žijí, případně informace o jejich mateřském jazyce.

Na závěr bychom Vás poprosili o vyplnění velmi krátkého dotazníku, o souhlas s účastí a Váš podpis.

Srdečně děkujeme a v případě dalších dotazů nás neváhejte kontaktovat!

/podpis autora, kontakt/

Dotazník:

Dnešní datum: _____

Jméno dítěte: _____ Datum narození: _____

Je Vaše dítě bilingvní, tzn. ovládá plynule více než jeden jazyk? **ANO – NE**

Jeho mateřský jazyk je: _____ a _____

Souhlasím s účastí mého dítěte v tomto jazykovém experimentu. **ANO – NE**

/podpis rodičů/

6.5 Příloha 5 (Zadání úkolu – pilotní sběr dat)

Písemná varianta č. 1 (psaní odpovědí na řádcích)

Mám pro vás teď úkol. Jde spíše **o hru**, není to na známky, nemusíte se bát.

Poprosím vás, abyste ke každému slovu, které najdete v sešitku, který jste dostali a máte ho před sebou, **napsali krátce a stručně popis**, jako byste vysvětlovali mladšímu sourozenci nebo mladšímu kamarádovi, **co to slovo znamená**.

Například, když se řekne **kiwi**, co vás napadne?

Mohli byste napsat, že:

/Přečíst + zapisovat na tabuli/

Kiwi je to ovoce, je šťavnaté, má hnědou slupku a uvnitř je zelené, je sladké nebo kyselé, hezky voní, je měkké a je chlupaté na povrchu, je k jídlu, jí se lžičkou, může se oškrábat, někdo ho jí také se slupkou, roste na stromě, neroste u nás a tak podobně, že?

Zkuste to teď vy: co je to **slunce**? *.../společně/...* Výborně.

Rozumíte tomu, co po vás teď chci?

Není to těžké, že?

Teď máte celou vyučovací hodinu na to, abyste mi psali ke každému slovu, **co znamená**.

Nezapomeňte, že mě zajímá, co o tom víš a co si myslíš **zrovna a jen ty**. Prosím, jste velké děti, jistě **zvládnete neopisovat**. Otočte se ke kamarádovi zády a pište každý sám. Není to písemka, zapojte fantazii a **pište to, co máte v hlavě**. Neraďte se mezi sebou, prostě napiš, co víš zrovna ty. Pokud nevíte k některému slovu **nic, nevíte, co napsat**, vůbec to **nevadí**. Můžete se k němu **později vrátit** a vzpomenout si ještě nebo ho **vynechat úplně**, nechat místo u slova **prázdné**.

Nemusíte spěchat, ale snažte se psát **čitelně**.

Prosím, teď začněte.

Písemná varianta č. 2 (psaní odpovědí ve sloupcích)

Mám pro vás teď úkol. Jde spíše **o hru**, není to na známky, nemusíte se bát.

Poprosím vás, abyste ke každému slovu, které najdete v sešitku, který jste dostali a máte ho před sebou, **napsali krátce a stručně popis**, jako byste vysvětlovali mladšímu sourozenci nebo mladšímu kamarádovi, **co to slovo znamená**. **Na každý řádek napíšete jednu informaci o té věci**.

Například, když se řekne **kiwi**, co vás napadne?

Na jeden řádek byste mohli napsat, že **je ovoce**, na další řádek **je šťavnaté**, na další, že **má hnědou slupku** a na další, že **uvnitř je zelené**, potom zase na další, že **je sladké** a pod to můžete napsat, že **je někdy taky kyselé**, další řádek bude, že **hezky voní**, další, že **je měkké**, na další, že **je chlupaté na povrchu**, další, že **je k jídlu**, další, že **se jí lžičkou**, pod to zase, že **se může oškrábat**, pod to, že **někdo ho jí také se slupkou**, pod to, že **roste na stromě**, a nakonec třeba to, že **neroste u nás** a tak podobně, že?

/Přečíst + zapisovat na tabuli/

Kiwi

je ovoce
je šťavnaté
má hnědou slupku
uvnitř je zelené

je sladké
... (všechny příklady)

Zkuste to teď vy: co je to **slunce**? .../společně/... Výborně.

Rozumíte tomu, co po vás teď chci?

Není to těžké, že?

Teď máte celou vyučovací hodinu na to, abyste mi psali ke každému slovu, **co to znamená**.

Nezapomeňte, že mě zajímá, co o tom víš a co si myslíš **zrovna a jen ty**. Prosím, jste velké děti, jistě **zvládnete neopisovat**. Otočte se ke kamarádovi zády a pište každý sám. Není to písemma, zapojte fantazii a **pište to, co máte v hlavě**. Neradte se mezi sebou, prostě napiš, co víš zrovna ty. Pokud nevíte k některému slovu **nic, nevíte, co napsat**, vůbec to **nevadí**. Můžete se k němu **později vrátit** a vzpomenout si ještě nebo ho **vynechat úplně**, nechat místo u slova **prázdné**.

Nemusíte spěchat, ale snažte se psát **čitelně**.

Prosím, teď začněte.

Ústní varianta

Mám pro tebe teď úkol. Jde spíše **o hru**, není to na známky, nemusíš se bát.

Poprosím tě, abys ke každému slovu, které ti řeknu, **řekl krátce a stručně popis**, jako bys to vysvětloval mladšímu sourozenci nebo mladšímu kamarádovi, **co to slovo znamená**.

Například, když se řekne **kiwi**, co tě napadne?

Že je to **ovoce, je šťavnaté, má hnědou slupku a uvnitř je zelené, je sladké nebo kyselé, hezky voní, je měkké a je chlupaté na povrchu, je k jídlu, jí se lžičkou, může se oškrábat, někdo ho jí také se slupkou, roste na stromě, neroste u nás a tak podobně, že?**

Zkuste to teď ty: co je to **slunce**? ...(*společně*)... Výborně.

Rozumíš tomu, co po tobě teď chci?

Není to těžké, že?

Teď máš celou vyučovací hodinu na to, abys mi ke každému slovu řekl, **co to znamená**.

Není to zkoušení, zapoj fantazii a **říkej mi to, co máš v hlavě**. Pokud nebudeš vědět k některému slovu **nic, nevíš co říct**, vůbec to **nevadí**.

Nemusíš spěchat, ale snaž se mluvit **zřetelně**.

Prosím, teď začneme.

6.6 Příloha 6 (Zadání úkolu – hlavní sběr dat)

Mám pro vás teď úkol. Jde spíše **o hru**, není to na známky, nemusíte se bát.

Poprosím vás, abyste ke každému slovu, které najdete v sešítku, který jste dostali a máte ho před sebou, **napsali krátce a stručně popis**, jako byste vysvětlovali mladšímu sourozenci nebo mladšímu kamarádovi, **co to slovo znamená**.

Například, když se řekne **citrón**, co vás napadne? Podívejte se se mnou na příklady na přední straně sešitu.

Mohli byste napsat, že:

Citrón: je ovoce, je šťavnatý, má žlutou kůru a uvnitř je také žlutý, je kyselý, hezky voní, má pecky, je k jídlu, může se přidat do čaje, někdo ho jí jen tak, roste na stromě, neroste u nás a tak podobně, že?

A třeba **nůž: je to věc, nástroj, pomocí kterého můžeme řezat věci a krájet jídlo, je hodně ostrý, je dlouhý, má dřevěnou rukojeť a má kovové ostří, je nebezpečný, používáme ho v kuchyni nebo třeba v dílně, může to být i zbraň, že?**

Klepat: znamená vydávat zvuk nebo ťukat, zvuk může vydávat člověk prsty, a může ho také vydávat nástroj nebo zvíře, člověk klepe na dveře a datel klepe do stromu, nejčastěji klepeme, když chceme vstoupit k někomu domů.

Průhledný: když je něco průhledné, znamená to, že je skrz to vidět a je to průsvitné, také tím prochází světlo, může také být sklo, ale i voda, často je průhledná výplň oken.

Zkuste to teď vy: co je to **slunce**? *.../společně/...* Výborně.
Rozumíte tomu, co po vás teď chci? Není to těžké, že?

Důležité je vědět, že můžete psát, **k jakému druhu nebo skupině ta věc patří, jak vypadá, třeba to, jak voní, jak chutná, jaké vydává zvuky, také jak funguje, k čemu ji používáme, kde ji můžete najít... a cokoliv, co jiného vás ještě napadne.**

Teď máte celou vyučovací hodinu na to, abyste mi psali ke každému slovu, **co znamená**. Nezapomeň, že mě zajímá, co o tom víš a co si myslíš **zrovna a jen ty**. Prosím, jste velké děti, jistě **zvládnete neopisovat**. Otočte se ke kamarádovi zády a pište každý sám. Není to písemka, zapojte fantazii a **pište to, co máte v hlavě**. Neradte se mezi sebou, prostě napiš, co víš zrovna ty.

Pokud nevíte k některému slovu **nic, nevíte, co napsat**, vůbec to **nevadí**. Zkus si vzpomenout, ale pokud to nepůjde, můžeš ho **vynechat úplně**, nechat místo u slova **prázdné**. Jakmile ale **půjdeš k dalšímu slovu, už se prosím nevracej, nelistuj sešitem zpět**, jdi dál a nech vše tak, jak jsi to napsal.

Nemusíte spěchat, ale snažte se psát **čitelně**.

Prosím, teď začněte.

6.7 Příloha 7 (Zadání úkolu – sémantická úloha)

Činnost, která nás čeká, je velice snadná. Experiment má několik částí.

Na začátku každé části ti položím **otázku**, na kterou budeš odpovídat. **První část je cvičná, zkusíme si jednu otázku spolu a jednu si zkusíš ty sám.**

Na obrazovce se budou **objevovat slova, postupně jedno za druhým. První slovo se vždy objeví na kratičký čas. U druhého slova budeš odpovídat na tu položenou otázku.**

Jakmile odpovíš, **klepneš na myš, kterou celou dobu držíš v ruce**, a zobrazí se další dvojice slov. **Myš – červené tlačítko – NE, zelené ANO.**

Na obrazovce se budou objevovat symboly na velmi krátký čas, **hvězdička ti ukazuje, kam se dívat.**

Snaž se zároveň sledovat to **slovo, na které neodpovídáš**, protože děláme experiment, ve kterém testujeme, jakou máš paměť, a já se tě pak na konci budu ptát na některá slovíčka – zda jsi je na počítači během experimentu už viděl, nebo ne.

Polož si palce na tlačítka a stiskni odpověď tak rychle, jak můžeš. **Všechno čteš potichu**, nemusíš nic říkat.

Pokud nestihneš odpovědět nebo odpovíš špatně, nic se neděje, pokračujeme dál. Je to celkem rychlá práce, dobře se soustřeď.

Některé části úkolu jsou trošku delší, ale to jistě zvládneš – na konci se na monitoru objeví slovo „**HOTOVO**“.

Za každou část dostaneš odměnu v podobě samolepky a na konci dostaneš velkou odměnu v podobě školních potřeb.

6.8 Příloha 8 (Příklady popisů pojmů dětmi)

(Jde o přesný opis odpovědí dětí, tedy včetně pravopisných a gramatických chyb, použití malých a velkých písmen apod.)

Dívka (8 let)	POMERANČ	<i>Je to ovoce. Má oranžovou barvu.</i>
	MEČ	<i>Je to železné i kovové. Má to rukojeť a používalo se to do bitvy.</i>
Dívka (8 let)	POMERANČ	<i>JE K JÍDLU, JE DOBRÝ, JE ORANGOVÍ (obrázek)</i>
	MEČ	<i>OSTRÝ, BOJUJE SE S NÍM, není na jídlo, není na HRANÍ. MAJÍ HO RYTÍRI (obrázek)</i>
Chlapec (8 let)	POMERANČ	<i>ovoce, oranžová</i>
	MEČ	<i>rytíř</i>
Chlapec (8 let)	POMERANČ	<i>je ovoce které se může jíst je orangové vnitřek je orangový</i>
	MEČ	<i>je ostrá zbraň rukojet je dřevěná čepel je ostrá kovová stříbrná</i>

6.9 Příloha 9 (Sémantické rysy uvedené dětmi ve věku 8–10 let pro dva vybrané pojmy)
(V tabulce nejsou zahrnuty sémantické rysy s produkční frekvencí 1, tzn. rysy, které byly zmíněny pouze jedním účastníkem)

pojem	rys	produkční frekvence	typ rysu
TYGR (popsán 27 účastníky)			
	<i>je zvíře</i>	19	taxonomické rysy
	<i>je savec</i>	2	taxonomické rysy
	<i>je šelma</i>	9	taxonomické rysy-části
	<i>je masožravec</i>	8	taxonomické rysy-části
	<i>je kočkovitá šelma</i>	6	taxonomické rysy-části
	<i>není věc</i>	2	taxonomické rysy-části
	<i>má srst</i>	4	percepční-vizuální rysy_části
	<i>má nohu/y</i>	3	percepční-vizuální rysy_části
	<i>má zuby</i>	3	percepční-vizuální rysy_části
	<i>má čtyři nohy</i>	3	percepční-vizuální rysy_části
	<i>má drápy</i>	2	percepční-vizuální rysy_části
	<i>má kůži</i>	2	percepční-vizuální rysy_části
	<i>je pruhovaný</i>	9	percepční-vizuální rysy_povrch
	<i>je oranžový</i>	5	percepční-vizuální rysy_barva
	<i>je černě pruhovaný</i>	4	percepční-vizuální rysy_barva
	<i>má oranžovou srst</i>	2	percepční-vizuální rysy_barva
	<i>má oranžovou kůži</i>	2	percepční-vizuální rysy_barva
	<i>má ostré zuby</i>	2	percepční-hmatové rysy
	<i>je nebezpečný</i>	7	encyklopedické rysy
	<i>je živý</i>	6	encyklopedické rysy
	<i>živí se masem</i>	4	encyklopedické rysy
	<i>v zoo je</i>	4	encyklopedické rysy
	<i>loví</i>	2	encyklopedické rysy
	<i>je zlý</i>	2	encyklopedické rysy
	<i>je v africe</i>	2	encyklopedické rysy
	<i>v savaně je</i>	2	encyklopedické rysy
KYTARA (popsána 21 účastníky)			
	<i>je hudební nástroj</i>	16	taxonomické rysy
	<i>má strunu/y</i>	7	percepční-vizuální rysy_části
	<i>je ze dřeva</i>	7	percepční-vizuální rysy_části
	<i>je hnědý</i>	3	percepční-vizuální rysy_barva
	<i>je bílý</i>	2	percepční-vizuální rysy_barva
	<i>hezký zvuk vydává</i>	2	percepční-zvukové rysy
	<i>zvuk vydává</i>	2	percepční-zvukové rysy
	<i>hraje se na to</i>	11	funkční rysy
	<i>brnká se na to</i>	2	funkční rysy
	<i>vyrábí se</i>	2	encyklopedické rysy
	<i>hraje se na to na koncertě</i>	2	encyklopedické rysy
	<i>hraje se na to rukama</i>	2	encyklopedické rysy

6.10 Příloha 10 (Srovnání sémantických rysů uvedených dětmi a slovníkové definice)

PAPOUŠEK	slovníková definice ¹⁵⁷	odpovědi participantů (frekvence)
<i>je barevný</i>	<i>pestře zbarvený</i>	6
<i>žíví se buráky</i>		1
<i>je býložravec</i>		1
<i>je část přírody</i>		1
<i>je červený</i>		1
<i>je doma</i>		2
<i>je domácí</i>		2
<i>má dvě nohy</i>		1
<i>žíví se hmyzem</i>		1
<i>má (= vlastní) hračky</i>		1
<i>chová se</i>		1
<i>žíví se jablky</i>		1
<i>je k jídlu</i>		1
<i>je krásný</i>		1
<i>má křídlo/a</i>		1
<i>létá</i>		4
<i>člověk to učí mluvit</i>		1
<i>je malý</i>		2
<i>žíví se malým hmyzem</i>		1
<i>je mazlíček</i>		1
<i>mluví</i>		7
<i>je moudrý</i>		1
<i>je nejedlý</i>		1
<i>má nohu/y</i>		1
<i>má oranžové nohy</i>		1
<i>je oranžový</i>		1
<i>má peří</i>		2
<i>žíví se potravou</i>		1
<i>je pták</i>	<i>pták</i>	12
<i>má různé druhy</i>		2
<i>je různobarevný</i>		4
<i>opakuje řeč</i>	<i>schopný napodobit zvuky lidské řeči</i>	4
<i>je tvor</i>		1
<i>není v České republice</i>		1
<i>je v džungli</i>		1
<i>je v pralese</i>		2
<i>je v tropickém deštném pralese</i>	<i>tropický</i>	1
<i>je velký</i>		2
<i>musí mít velký prostor</i>		1

¹⁵⁷ Červená V. et al.: *Slovník spisovné češtiny pro školu a veřejnost*, 2005.

<i>je vzácný</i>		1
<i>je zelený</i>		1
<i>má zobák</i>		2
<i>zpívá</i>		1
<i>je zvíře</i>		10
<i>vydává zvuk/y</i>	<i>s pronikavým hlasem</i>	2
<i>je živý</i>		3

TRIANGL	slovníková definice ¹⁵⁸	odpovědi participantů (frekvence)
<i>cinká</i>		3
<i>udírá se do toho</i>	<i>bicí</i>	3
<i>je hlasitý</i>		1
<i>vydává hudbu</i>		1
<i>je hudební nástroj</i>	<i>hudební nástroj</i>	18
<i>má kovovou tyč</i>		2
<i>je lidský produkt</i>		1
<i>vydává melodii</i>		1
<i>cinká se na to</i>		1
<i>hraje se na to</i>		4
<i>je stříbrný</i>		1
<i>má šňůru</i>		1
<i>má tvar trojúhelníku</i>	<i>v podobě trojúhelníku</i>	7
<i>má tyč</i>		6
<i>udírá se do toho tyčí</i>		1
<i>cinká se na to tyčí</i>		1
<i>hraje se na to tyčí</i>		1
<i>používá se v písních</i>		1
<i>věc je</i>		2
<i>je z kovu</i>	<i>kovový</i>	3
<i>drží se za šňůru</i>		1
<i>je ze železa</i>		5
<i>vydává zvuk/y</i>		2
<i>má železnou tyč</i>		2

¹⁵⁸ Červená V. et al.: *Slovník spisovné češtiny pro školu a veřejnost*, 2005.

6.11 Příloha 11 (Srovnání sémantických rysů uvedených českými dětmi a anglickými mluvčími)

JELEN	produkční frekvence ¹⁵⁹	K. McRae et al. ¹⁶⁰ MOOSE	produkční frekvence
<i>je býložravec</i>	2	<i>an herbivore</i>	8
<i>není člověk</i>	1	-	
<i>má čtyři nohy</i>	1	<i>has four legs</i>	12
<i>má dva parohy</i>	1	-	
<i>je hezký</i>	1	-	
<i>má hezký ocas</i>	1	-	
<i>je hlasitý</i>	1	-	
<i>má hnědou srst</i>	1	-	
<i>je hnědý</i>	7	<i>is brown</i>	10
<i>má huňatou srst</i>	1	-	
<i>má kopyta</i>	3	<i>has hooves</i>	5
<i>je král</i>	1	-	
<i>je král zvířat</i>	1	-	
<i>loví se</i>	1	<i>hunted by people</i>	17
<i>má malý ocas</i>	1	-	
<i>maso je chutné</i>	1	-	
<i>maso je tmavé barvy</i>	1	-	
<i>maso je zvěřina</i>	1	-	
<i>mládě je kolouch</i>	1	-	
<i>střílí to myslivec</i>	1	-	
<i>má na hlavě parohy</i>	1	-	
<i>má nohu/y</i>	1	<i>has legs</i>	14
<i>samci bojují o samici</i>	1	-	
<i>má ocas</i>	3	-	
<i>má parohy</i>	16	<i>has antlers</i>	23
<i>shazuje parohy</i>	1	-	
<i>je plachý</i>	2	-	
<i>je podobný losovi</i>	1	-	
<i>je rychlý</i>	2	-	
<i>je samec</i>	1	-	
<i>je samice laň</i>	1	-	
<i>je savec</i>	1	<i>a mammal</i>	9
<i>je silný</i>	1	-	
<i>má srst</i>	3	<i>has fur</i>	7
<i>je trochu podobný losovi</i>	1	-	

¹⁵⁹ Je nutné mít na mysli, že produkční frekvence ukazuje, kolik participantů daný rys pro pojem zmínilo. Není proto možné toto číslo srovnávat mezi dvěma databázemi, které obsahují data od jiného počtu účastníků výzkumu. Přesto je i v této tabulce toto číslo užitečné, protože vypovídá o frekvenci jednotlivých rysů – tzn. o jejich důležitosti – v rámci daného pojmu. Například je v obou databázích mnohými účastníky zmíněn rys *má parohy*, a to jak u *jelena*, tak u *losa*. Informace, že *má nohy*, však byla zmíněna mnohými dospělými, avšak pouze jedním dítětem.

¹⁶⁰ **Mcrae, K. et al.:** *Semantic feature production norms for a large set of living and non-living things*, 2005.

<i>je v lese</i>	13	<i>lives¹⁶¹ in woods</i>	14
<i>v zimě shazuje parohy</i>	1	-	
<i>má velké parohy</i>	1	-	
<i>je velký</i>	1	<i>is large</i>	27
<i>je z masa</i>	1	-	
<i>je zvíře</i>	17	<i>an animal</i>	17
<i>vydává zvuk/y</i>	1	-	
<i>je živý</i>	10	-	
-		<i>has hair</i>	5
-		<i>eaten for meat</i>	5
-		<i>lives in wilderness</i>	8

DÝKA	produkční frekvence	K. McRae et al. ¹⁶² KNIFE	produkční frekvence
<i>brousí se</i>	1	-	
<i>je dlouhý</i>	1	-	
<i>má dřevěnou rukojeť</i>	1	-	
<i>chrání</i>	1	-	
<i>má koženou rukojeť</i>	1	-	
<i>je krátký</i>	1	-	
<i>je krutý</i>	1	-	
<i>je krvavý</i>	1	-	
<i>používá to loupežník</i>		-	
<i>Karaba</i>	1		
<i>používá to loupežník</i>	1	-	
<i>používá to pytlák</i>	1	-	
<i>je malý</i>	1	-	
<i>je nářadí</i>	1	-	
<i>je nástroj</i>	2	<i>is utensil</i>	9
<i>je nebezpečný</i>	2	<i>is dangerous</i>	14
<i>je nůž</i>	5	-	
<i>je ostrý</i>	15	<i>is sharp</i>	29
<i>je ostřejší než klasický nůž</i>	1	-	
<i>má ostří</i>	1	<i>has a blade</i>	11
<i>je podobný noži</i>	5	-	
<i>má rukojeť</i>	4	<i>has a handle</i>	14
<i>je řezný</i>	1	-	
<i>je špičatý</i>	2	-	
<i>bojuje se tím</i>	1	-	
<i>krájí se tím</i>	1	<i>used for cutting</i>	25
<i>zabíjí se tím</i>	4	<i>used for killing</i>	7
<i>je tupý</i>	1	-	
<i>je v dílně</i>	1	-	
<i>je v plátěné pochvě</i>	1	-	
<i>je v pohádce</i>	1	-	
<i>je v pohádce Tajemství staré bambitky</i>	1	-	
<i>je v pochvě</i>	1	-	
<i>je ve filmu</i>	1	-	

¹⁶¹ V našem případě uvádíme také samostatný rys živý je.

¹⁶² **Mcrae, K. et al.:** *Semantic feature production norms for a large set of living and non-living things*, 2005.

<i>používalo se ve středověku</i>	1	-	
<i>je věc</i>	6	-	
<i>je velký</i>	2	-	
<i>je větší než klasický nůž</i>	1	-	
<i>je vojenský</i>	1	-	
<i>vraždí</i>	2	-	
<i>je z kovu</i>	1	<i>made of metal</i>	7
<i>je z kůže</i>	1	-	
<i>zabíjí</i>	1	-	
<i>zbraň je</i>	8	<i>is weapon</i>	11
<i>ze železa je</i>	1	-	
-		<i>made of steel</i>	8
-		<i>made of stainless steel</i>	5
-		<i>is shiny</i>	5
-		<i>used by butchers</i>	5
-		<i>is serrated</i>	8
-		<i>found in kitchens</i>	8
-		<i>used with forks</i>	6
-		<i>a cutlery</i>	5

6.12 Příloha 12 (Přehled pojmových párů a míry sémantické podobnosti mezi pojmy v páru)

otázka	primové slovo (prime)	cílové slovo (target)	vysoká podobnost	primové slovo (prime)	cílové slovo (target)	nízká podobnost
Roste to?	SUP	OREL	0.877	HUSA	OREL	0.430
	HOLUB	PAPOUŠEK	0.627	LABUŤ	PAPOUŠEK	0.456
	HOLUB	RACEK	0.742	HUSA	RACEK	0.409
	VIŠEŇ	TŘEŠEŇ	0.867	ANANAS	TŘEŠEŇ	0.479
	POMERANČ	MANDARINKA	0.898	JAHODA	MANDARINKA	0.400
	VIŠEŇ	JABLKO	0.782	POMERANČ	JABLKO	0.453
	KVĚTÁK	FAZOLE	0.784	CIBULE	FAZOLE	0.414
	SALÁT	OKURKA	0.742	MRKEV	OKURKA	0.411
	KVĚTÁK	CUKETA	0.801	CIBULE	CUKETA	0.459
	SALÁT	HRÁŠEK	0.701	MRKEV	HRÁŠEK	0.340
	KOČKA	KŮŇ	0.604	DELFIN	KŮŇ	0.414
	KOČKA	KRÁLÍK	0.689	DELFIN	KRÁLÍK	0.268
	ZAJÍC	JELEN	0.689	OPICE	JELEN	0.427
	KOPRETINA	SEDMIKRÁSKA	0.877	LÍPA	SEDMIKRÁSKA	0.063
	LÍPA	BOROVICE	0.629	LEKNÍN	BOROVICE	0.131
	RŮŽE	TULIPÁN	0.769	ČTYŘLÍSTEK	TULIPÁN	0.274
Vyrábějí to lidé?	SEDAČKA	ŽIDLE	0.672	LAMPA	ŽIDLE	0.101
	SEDAČKA	POHOVKA	0.750	VĚŠÁK	POHOVKA	0.089
	DVEŘE	SKŘÍŇ	0.249	LAMPA	SKŘÍŇ	0.077
	HRNEK	SKLENICE	0.603	VAŘEČKA	SKLENICE	0.037
	LŽÍCE	STRUHADLO	0.179	NABĚRAČKA	STRUHADLO	0.049
	KONVICE	HRNEC	0.412	VAŘEČKA	HRNEC	0.025
	REVOLVER	PISTOLE	0.821	MEČ	PISTOLE	0.278
	MEČ	DÝKA	0.762	PRAK	DÝKA	0.147
	REVOLVER	DĚLO	0.852	OŠTĚP	DĚLO	0.225
	TROLEJBUS	VLAK	0.702	LOŽ	VLAK	0.065
	LETADLO	VRTULNÍK	0.695	TRAKTOR	VRTULNÍK	0.091
	TROLEJBUS	AUTOBUS	0.676	LOŽ	AUTOBUS	0.104
	ČEPICE	KLOBOUK	0.806	PYŽAMO	KLOBOUK	0.0
	ŠATY	TRIKO	0.700	BOTA	TRIKO	0.231
	ŠATY	SUKNĚ	0.855	BOTA	SUKNĚ	0.198
	KYTARA	HOUSLE	0.880	TRIANGL	HOUSLE	0.547
	VARHANY	HARFA	0.900	HOUSLE	HARFA	0.672

	LOPATA	RÝČ	0.607	KOŠTĚ	RÝČ	0.062
	SEKERA	PILA	0.201	KOŠTĚ	PILA	0.033
	LOPATA	HRÁBĚ	0.289	SEKERA	HRÁBĚ	0.066

7. Anotace

Sémantické rysy a jejich úloha v pojmové reprezentaci u dětí

Osvojování jazyka u dětí je jedním z aktuálně diskutovaných témat v oblasti psycholingvistiky. Značný prostor pro budoucí výzkum lze spatřovat v oblasti vývoje slovní zásoby u česky mluvících dětí, přičemž v našem případě nás zajímá především význam, tj. obsah osvojovaných slov (pojmu), a role tzv. sémantických rysů v mentální reprezentaci.

Zamýšleným cílem výzkumu je přinést nové informace z výše zmíněné oblasti, potvrdit nebo vyvrátit některá dosavadní teoretická tvrzení a porovnat výsledky zahraničních výzkumů s daty, která budou získána na materiálu českého jazyka. Obdobné výzkumy byly prováděny v různých světových jazycích, zatím ale neexistuje mnoho prací, které by se věnovaly dané problematice v českém jazykovém prostředí. V rámci naší práce byla připravena obsáhlá databáze sémantických rysů pro vybrané pojmy. Tato databáze byla statisticky zpracována a následně byla data analyzována a interpretována na základě teorií o vývoji řečové kompetence dítěte. Tento materiál, získaný od dětí ve věku 8–9 let (nižší stupeň základní školy) vyrůstajících v českém jazykovém prostředí, byl využit v další fázi výzkumu, ve které byl proveden experiment s účastníky spadajícími do stejné věkové kategorie: v sémantické úloze založené na jevu zvaném sémantický priming byl pozorován vliv rysově podobnosti dvou pojmů na rychlost rozhodování.

Výsledky výzkumu rozšiřují škálu dosud publikovaných informací v této vědecké oblasti v českém prostředí. Tento výzkum může přinést cenné poznatky v otázkách akvizice jazyka u dětí. Získaná data mohou být také prakticky přínosná nejen pro pedagogy, psychology a logopedy, ale například i pro rodiče.

Semantic Features and Their Role In Conceptual Representation In School Age Children

Language acquisition is one of the currently much discussed topics in the field of psycholinguistics. Considerable space for future research can be seen in the development of vocabulary in Czech-speaking children. In our case, we are mainly interested in the meaning, i.e. the content of acquired words (concepts), and the role of so-called semantic features in mental representation.

The intended goal of our research is to bring new information from the above-mentioned area, to confirm or disprove some existing theoretical statements and to compare the results of foreign research with data obtained using the Czech language material. Similar research has been conducted in various world languages, but so far there are not many papers that address the issue in the Czech language environment. As part of our work, a comprehensive database of semantic features for selected concepts has been prepared. This database has been statistically processed and subsequently the data has been analyzed and interpreted on the basis of theories about the development of the child's speech competence. This material, obtained from children aged 8-9 (lower primary school) growing up in a Czech language environment, has been used in the next phase of research, in which an experiment with subjects belonging to the same age category has been performed: in a semantic task based on the phenomenon called semantic priming, the effect of featural similarity of two concepts on decision in a speeded task has been observed.

The results of the research expand the range of information published so far in this scientific field in the Czech environment. This research can provide valuable insights into children's language acquisition issues. The data gathered can also be practically beneficial not only for teachers, psychologists and speech therapists, but also for parents, for example.