ÚSTAV ČESKÉHO JAZYKA A TEORIE KOMUNIKACE
FF UK V PRAZE

Posudek na magisterskou diplomovou práci

Autor DP: **Bc. Martina Vokáčová**

Název práce (česky/anglicky): **Vliv gramatických profilů českých substantiv na jejich osvojování nerodilými mluvčími /** The influence of grammatical profiles of Czech nouns on their acquisition by non-native speakers

Vedoucí práce: **Mgr. Eva Lehečková, Ph. D.**

Oponent: **Mgr. Marek Nagy, Ph. D.**

Hodnoticí kritéria	Hodnocení
Soulad názvu práce s obsahem, schopnost formulovat cíl a závěry práce	2
Teoretická podkladová koncepce, metodologie, terminologie	1-2
Schopnost analýzy excerpovaného materiálu/předmětu výzkumu a následné syntézy	2
Interpretace a argumentace	2-3
Sekundární literatura, tvůrčí aplikace, schopnost kritického komentáře	1-2
Struktura práce, výstavba textu, adekvátnost použitých stylistických prostředků, dodržení ortografické a gramatické normy	3
Bibliografický a poznámkový aparát, jednotný způsob citací	2
Výsledná známka	**Velmi dobře**

Hodnocení písemné verze práce:

Diplomová práce Martiny Vokáčové referuje o autorčině původním empirickém výzkumu, který vycházel z premis a postupů usage-based orientované kognitivní lingvistiky. Výzkum byl zaměřen na problematiku osvojování druhého jazyka, konkrétněji na testování hypotézy o vlivu faktoru pádové frekvence českých substantiv na úroveň chybovosti v produkci studentů češtiny jako cizího jazyka. Tato hypotéza umožňuje ve výzkumném programu, ke kterému se jeho autorka hlásí, konkrétní predikci: čím robustnější je mentální reprezentace gramatického profilu substantiva v mysli jeho uživatele, tím menší je pravděpodobnost, že v případech užití tomuto profilu odpovídajících jazykových jednotek bude mluvčí chybovat. Potvrzení platnosti

dané hypotézy by jednak posílilo diskutovanou relevanci předpokladu o mentální reprezentaci gramatických profilů v mentálním lexikonu; o pozitivních důsledcích takového zjištění pro didaktiku cizích jazyků pak netřeba pochybovat. Toto – rozumně opatrné! – zjištění je také závěrem autorčina zkoumání.

Na základě předložené práce a dat v ní obsažených se domnívám, že Martina Vokáčová na jedné straně realizovala přínosný a smysluplný výzkum, na straně druhé o něm však referuje se zjevnými textovými nedostatky, které výsledný dojem narušují nejen po stránce formální, ale také po stránce věcné/výzkumné. Napětí mezi těmito dvěma složkami badatelské práce je náplní následného kritického komentáře. Nepoměr mezi níže představenými klady a zápory práce neodpovídá přitom jejich **faktickému** poměru v celku autorčina bádání, je motivován spíše mojí snahou soustředit autorčinu pozornost nedostatky, kterým by bylo vhodné se vyhnout při předpokládaném pokračování jejího zajímavého výzkumu.

***

Úvod diplomové práce je povedený. Je napsán svižným, čtivým stylem, autorka s příjemnou lehkostí srozumitelně zdůvodňuje motivaci výzkumu a jeho primární cíl. Cíl se přitom jeví jako přiměřený a v rámci diplomové práce realizovatelný. Dobré sebe/zadání badatelského úkolu.

V následujících kapitolách autorka na obecnější úrovni kontextualizuje rámec a východiska svého specifického výzkumu. Představuje perspektivu usage-based a (radikálně korpusové) kognitivní lingvistiky a seznamuje stručně čtenáře s některými z vybraných základních konceptů, které jsou pro tento přístup i autorčinu práci určující (frekvence a její kognitivní efekty v reprezentaci a zpracování jazyka, pojem gramatického profilu). Co v této části práce však podle mého názoru nevhodně absentuje, je podrobnější teoretické a argumentační představení konkrétní problematiky, které se práce věnuje, a jejímu instrumentáriu. Tedy podrobnější „logické“ provázání autorčina výzkumu s již realizovanými výzkumy v dané oblasti a podrobnější referenci o fenoménech a konceptech, které poskytují zadání práce oporu. Diplomová práce není útvarově totožná s odborným článkem, pohybuje se spíše na pomezí článku a odborné monografie, proto je dle mého názoru vhodné věnovat čas a prostor budování co nejdetailnější východiskové a argumentační báze, na jejímž pozadí bude možné výsledky výzkumu interpretovat, rozvíjet a obhajovat. Dobré pojmové vymezení a ucelený přehled dosavadních relevantních zjištění o klíčových konceptech (nástrojích) autorčina výzkumu je předpokladem porozumění jeho teoretické opodstatněnosti. To samé se týká i popisu sledovaných jazykových jednotek. Takovým pars pro toto je např. způsob, jakým autorka zachází s klíčovým termínem práce, **gramatickým profilem**. Na s. 13 se jím rozumí „substantivní třídy/klastry lemmat se stejnou frekvencí“, zatímco na s. 43 je „chápán jako soubor dvou až tří nejfrekventovanějších tvarů“. Je třeba být v tomto ohledu důslednější.

Slabší textová a argumentační provázanost mezi východisky práce, cílovou hypotézou a jednotlivými metodologickými postupy bohužel v jistém smyslu poznamenává celé vyznění práce. Možná že právě jen z důvodu této **textové** absence nakonec syntetizující část práce vyznívá poněkud neorganizovaně a výčtově-deskriptivně. V empirickém výzkumu je nutné usilovat o dobrou logickou strukturaci, maximální explicitnost a soustředěnost na klíčové

otázky a relevantní faktory zkoumání. Právě v tomto místě tedy pociťuji jistou slabina předložené práce. Dobře vystavěná „mapa“ problematiky umožňuje nejen předjímat možné nejasnosti, pomáhá zároveň strukturovat text v závěrečné interpretační části.

Čemu naopak autorka věnuje – naprosto správně! – velkou pozornost, je popis metodiky jejího konkrétního výzkumu. Tedy stavbě vzorku a kritériím pro kódování. Cenná je zejména její reflexe problémů, které se týkají využitelnosti obou sledovaných korpusů. Přesto se domnívám, že i v této pasáži by bylo vhodné jednak doplnění některých údajů, jednak lepší reflexe problematiky spojené s kódováním chybovosti, zejména s jejich dopady při vyhodnocování. V textu je např. dobře popsána a zdůvodněna volba cílových sledovaných lexémů. Přesto bych navrhoval doplnit ji o podrobnější lexikálně-sémantickou charakteristiku. Část výzkumu sleduje (ne úplně dobře zdůvodněně) klastrování lexémů dle sémantických a funkčních tříd, proto je podle mého názoru dobré mít lexémy kódovány vzhledem k těmto jejich vlastnostem. (Například i proto, že pro akvizici lexikálních jednotek v oblasti osvojování prvního jazyka některé výzkumy označují jako relevantní faktor vztah mezi představitelností denotátu a akvizicí morfologických charakteristik dané jednotky). Stejně tak bych očekával větší explicitnost při řešení nejednoznačností z hlediska kódování chybovosti (viz příklad s lexémem *doba*).

Po náležitě deskriptivní analytické části, která zachycuje vztahy ve sledovaném systému, přichází – jak už bylo uvedeno výše – nevhodně stručná část interpretační, diskuzní a syntetická. Je nutné říct, že autorka si vytyčila jako cíl identifikaci korelace mezi dvěma jevy, předpokládal bych ale lépe komentované zdůvodnění konkrétních nálezů (a jejich případných odchylek). Vzhledem ke stavbě vzorku (resp. způsobu kódování) by také možná bylo na místě věnovat při analýze chyb pozornost nejen rysů shodným, ale i idiosynkratickým rysům jednotlivých jednotek. I proto závěrečný oddíl podle mého názoru jen málo vytěžuje zjištění, která si autorka připravila v předchozích materiálových kapitolách.

Navzdory uvedeným výtkám se však domnívám, že tyto obsahové nedostatky jsou dílem časové tísně, nikoliv neschopnosti promyslet lépe závěry, a proto i v tomto ohledu hodnotím práci jako zajímavou, byť vlastně neuzavřenou, nedotaženou/nevytěženou.

Stejnou příčinu, časovou tíseň, mohou mít i nedostatky formální, kterých je však v práci opravdu hraniční množství a to zbytečně znehodnocuje dojem z ní, v některých případech i orientaci v obsahu. Prakticky na každé straně (hyperbola, ale mírná) nalezneme nějaký ortografický nedostatek (chybějící slovo, mezera, interpunkce, záměna výrazů atd.). Trochu příznačné je, že takový typ chyby nalezneme dokonce i v autorčině dodatečném erratu... Tabulka 3.1 obsahuje logický nesmysl (při srovnání s citovaným zdrojem se ukazuje, že v něm je algoritmus zapsán správně), v bibliografii chybí odkazovaný zdroj (Roeper 2007) , autorka cituje ze druhé ruky (byť jen v obecné části), používá zkratku, kterou nezavádí (KWIC), u tabulek zapomene použít legendu a nepoužívá hlavičky, v závěru práce uvede chybně jedno z hlavních zjištění výzkumu („třídu s1-p1-s1 (obsahující lemmata muž , autor , lékař)“ – má tam být p2), občas nedodržuje vlastní typografickou konvenci. Ze čtenářského hlediska by také bylo vhodnější neuvádět nejprve společně všechny komentáře a až poté souhrnně příklady, protože v mnoha případech je pak kvůli hutnosti informací nutné se v textu

vracet a znova usouvztažňovat popis a jeho ilustraci. Rovněž bych doporučil zvážit některé formy vizualizace výsledků, např. u klastrování. Tabulky nevyznívají v tomto případě tak přehledně a existují na to pěkné a percepčně funkčnější „přátelštější“. Atd.

I přes tato textová selhání vnímám práci Martiny Vokáčové jako nadějný příspěvek hned k několika tématům, doporučuji ji k obhajobě a hodnotím stupněm **velmi dobře**.

Náměty pro obhajobu (formou otázek):

1. Jsou podle Vás frekvenční údaje korpusu SYN2015 dobrým indikátorem inputu, tj. zdroje mentální reprezentace pro zkoumanou cílovou skupinu? Proč? Nabízí se alternativa?
2. Jakou úlohu ve zkoumání problematiky připisujete gramatickým profilům odvozeným z korpusů Merlin a CZE-SL? Pokud byly sestaveny na základě frekvencí náležitých, nikoliv užitých tvarů (chápu to správně?), proč?
3. Nabízí se podle Vás při uvažování o zdrojích transferu i jiné hledisko než hledisko typové frekvence? Pomohlo by kódování jiných vlastností lexémů? Pokud ano, kterých?

V Hradci Králové dne 7. 9. 2016