

## Oponentský posudek disertační práce Mgr. Věry Hejhalové

*Zur Integration von Phrasemen in ein allgemeines und phraseologisches Wörterbuch Deutsch-Tschechisch aus kontrastiver und korpuslinguistischer Sicht*

Autorka se ve své disertační práci ze široka zabývá lexikografickým zachycením frazémů na základě moderního přístupu k lingvistice – jazykových korpusů, a to v osmi kapitolách. V kratičké úvodní kapitole 1 představuje obecný cíl své práce: popsat ve slovníku co nejpřehledněji frazeologii za pomoci hlavních moderních lingvistických nástrojů – jazykových korpusů, a to z kontrastivního hlediska: němčina vs. čeština. V tomto směru má autorka ke své práci přímo ideální podmínky: zejména má k dispozici korpus němčiny DeReKo, největší jednojazyčný korpus na světě, poněkud menší korpusy současné češtiny řady SYN a též paralelní německo-český korpus v rámci podprojektu Interkorp, oboje v rámci projektu *Český národní korpus*. Zdůrazňuje interdisciplinárnu své práce: spojuje v ní velmi podrobně a s hlubokou znalostí předmětu svého zkoumání aspekty frazeologie, korpusové lingvistiky a lexikografie. Na základě existujících korpusů a korpusových analýz navrhuje moderně pojaté lexikografické heslo, přičemž zkoumá „nové teoretické, resp. metodologické aspekty lexikografického zpracování frazémů na základě korpusů“. Praktické výsledky v podobě nově koncipovaných hesel budou zahrnuty do připravovaného *Velkého německo-českého akademického slovníku*.

V kapitole 2 představuje autorka svou metodu. Na vybraných zvířecích lexémech tvořících frazémové jádro „zvířecích frazémů“ hodlá zkoumat, jak jsou příslušné frazémy traktovány ve slovnících, zejména německých, přičemž si také připravila databázi takových frazémů, která je základem zmíněného slovníku a praktickým výstupem jejího studia teoretického. Dále uvádí stručnou charakteristiku použitých slovníků a dalších příruček. Zaměřuje se na „frazeoaktivitu“ názvů zvířat v obou jazycích a z příslušných frazémů se soustředí hlavně na přirovnání. Přehledně uvádí postup své analýzy v logicky konzistentním schématu na s. 15, přičemž klíčovou roli hraje inteligentní hledání v korpusu a na internetu (chytré hledání toho, co chci nalézt, je pro většinu lingvistů – troufám si napsat – dosud pouze nedosažený ideál!, neplatí to však o autorce). Výsledkem má být nově koncipované slovníkové heslo. V podkapitole 2.3 se pak doktorandka přehledně zabývá datovou základnou a softwarovými nástroji svého výzkumu – korpusy: mannheimským korpusem DeReKo a jeho ve světě unikátními nástroji jazykovědné analýzy (softwaru COSMAS a databázi CCDB). Stručně pak pojednává o projektu *Český národní korpus* a zmiňuje se i o internetu. Poté se v podkapitole 2.4 zabývá lexikografickými východisky a cíli, stručně charakterizuje slovníky, o nichž v práci píše: *Deutsches Universalwörterbuch*, *Duden – Redewendungen*, *Wörterbuch der deutschen Idiomatik, Band 11*; *Langenscheidts Großwörterbuch Deutsch als Fremdsprache*; Čermákův *Slovník české frazeologie a idiomatiky*, nedávno vydaný *Deutsch-tschechisches Wörterbuch der Phraseologismen und festgeprägten Wendungen* (dále jen DTWP) a internetovou databázi frazémů *Redensartindex*. Pár řádek věnuje i internetu. V další části kapitoly 2 uvádí autorka slovník, který by měl vzniknout i jako produkt její praktické lexikografické práce: *Velký německo-český akademický slovník* (dále jen GAWDT, mající obsahovat cca 130000 hesel), kde frazeologie má – zcela náležitě – hrát prominentní úlohu.

Kapitola 3 obsahuje teoretická východiska autorky pro zpracování frazémů ve slovníku. Velmi přehledná kapitola se zabývá frazeologií, frazeografií a korpusovou lexikografií. V této kapitole autorka nade vši pochybnost dokládá svou obeznámenost s předmětem svého studia, což je také potvrzeno důkladným soupisem relevantní literatury v závěru práce. Nejprve představuje frazeologii jako jazykovědné odvětví a její vztah k lexikologii (kap. 3.1.1), poté její vztah ke slovtvorbě (kap. 3.1.2) včetně (de)frazeologické derivace; dále vztah frazeologie k vnější a vnitřní syntaxi frazému (kap. 3.1.3) a pak se vyjadřuje i k postavení frazeologie v systému lingvistických disciplín. V podkapitole 3.2 se zabývá pojmem frazému a frazeologismu v literatuře, polylexikalitou a vztahem vymezení slova a frazému, přičemž neopomine ani

slovotvorbu. Jde jí o kriteriální vymezení frazému: více slov než jedno, ustálenost, idiomatičnost. V této souvislosti se dostává k rozdílu mezi idiomem a kolokací a volným spojením slov, přičemž uvádí případy, kdy se dané slovní spojení dá jen obtížně zařadit do příslušné kategorie. Na (částečné) idiomy se přitom – na to doktorandka nezapomíná! – dívá z vícejazyčného hlediska, vždyť jí jde v praktickém důsledku o německo-český slovník a hlavně o českého uživatele, jehož prvním jazykem je čeština. Naprosto souhlasím s jejím tvrzením, na s. 53 a 54: „Die rezeptiven Wörterbücher sollten grossen Wert auf Idiome legen, weil vor allem ihr Verständnis im Gegensatz zu den Kollokationen Probleme verursacht, dagegen sollten produktive Wörterbücher (Muttersprache → Fremdsprache) eher die Kollokationen den Idiomen vorziehen, weil man ohne Kollokationen einen fremdsprachlichen Text nicht produzieren kann. Die Idiome kann man dagegen meistens problemlos durch nichtidiomatische Umschreibungen ersetzen.“

Dále zařazuje ještě zvláštní podkapitulu o přirovnávacích frazémeh, jimž se ve své práci bude věnovat obšírněji později, a příslovím. V podkapitole 3.2.7 velmi přehledně shrnuje pravidla vymežující frazém.

V kapitole 3.3 se autorka dostává k frazeografii, čili k nesmírně důležitým praktickým problémům (může se totiž stát, že lexikograf je vynikající teoretik, který však své teoretické znalosti nedokáže adekvátně vyjádřit ve slovníku). Rozlišuje mezi produkcí a recepcí textů jakožto kritériem náležité koncepce slovníku a z hlediska hlavního předmětu svého zkoumání – moderního zachycení frazeologie ve slovníku – správně konstatuje, že „frazémy hrají důležitou roli především při recepci textů“. Stručně se pak zmiňuje o frazémeh ve slovníku GAWDT, na jehož přípravě se podílí, a poté přechází ke struktuře slovníkového hesla (kap. 3.3.4). Všímá si tu zejména nesnadných problémů, jak identifikovat příslušné heslové slovo frazému (Nennform), aby je uživatel slovníku mohl snadno a rychle ve slovníku nalézt, jeho variabilitou, nesnadností odlišení frazému od nefrazémových užití příslušného spojení a kritérii řazení jednotlivých frazémů (podle slovních druhů, abecedně...). Velmi podrobně se věnuje i aspektům stylistickým (3.3.4.3) a názorně srovnává stylistický popis ve *Slovníku české frazeologie a idiomatiky*, slovníku *Deutsches Universalwörterbuch, Duden – Redewendungen. Wörterbuch der deutschen Idiomatik, Band 11* a v připravovaném slovníku GAWDT. Věnuje se i opisu významu frazému ve výše zmíněných slovnících. Všímá si pochopitelně frazeologických ekvivalentů německých frazémů v jiném jazyce, zde v češtině, a jejich zachycením ve slovníku GAWDT. Za velmi důležité považuje autorka i příklady, a to zejména příklady vzaté z reálného jazykového úzu. Zde se výrazně projevuje význam jazykových korpusů obsahujících nepřeborné (doslova!) množství příkladů. Na konci lexikografické části doktorandka velmi výstižně charakterizuje gramatické aspekty frazémů, různé nepravidelnosti a omezení, frazeologická antonyma a synonyma, intonaci a etymologii a uvádí strukturu konkrétního hesla (phraseographisches Raster) s charakteristickým příkladem.

V kapitole 3.4 předvádí autorka suverénní přehled o korpusové lexikografii: zná rozdíl mezi přístupem *corpus-based* a *corpus-driven* a správně charakterizuje i další aspekty korpusů (přičemž uvádí nejmodernější literaturu o problematice korpusů). Velmi důležité je, že autorka chápe frazémy jako projevy zejména mluveného, spontánního jazyka. Nejsou jí cizí ani pojmy vyváženost a reprezentativnost, na něž klade velký důraz český lexikografický koryfeje František Čermák. Jejím úvahám o internetu jakožto korpusu není co vytknout. Zajímavou kapitolou pro mne byla kapitolka 3.4.2.5 o anotacích v korpusu, zejména morfosyntaktických. Autorka cituje na s. 89 v německém překladu Františka Čermáka: „... Annotationen ... deformieren damit [tj. jazykovědnou teorií autora příslušné anotace] den ursprünglichen ganz authentischen Text und somit ebenso seinen Aussagewert.“ Není mi tu jasné, jak anotace jakožto

k autentickému textu pouze přidaná informace může výpovědní hodnotu textu deformovat. Vždyť výzkumník může značkování textu zcela pominout, anotace slova není přece se slovem spjata jako neoddělitelné siamské dvojče. Proto mne zaráží, že s něčím pro jazykovědný výzkum tak důležitým a tak nápomocným, jako je právě morfosyntaktické značkování, se v korpusu DeReKo začalo až (sic!) v roce 2009!! (zdůvodnění vzápětí následuje a je až neuvěřitelné, uvědomím-li si existenci jiných nanejvýš sofistikovaných nástrojů pro práci s korpusem DeReKo, jako je třeba třeba COSMAS II). V Mannheimu v tomto směru zaspali dobu. Autorka se možná mohla k této situaci v Mannheimu kriticky vyjádřit, stejně jako neměla ponechat bez kritiky výše citovaný Čermákův výrok (s jinými citacemi z Čermáka jakožto největšího českého odborníka na frazeologii uváděnými v disertaci ovšem nelze než souhlasit). Pasáž o využití korpusů při zpracovávání zejména frazematických hesel v GAWDT je kritická a naprosto výstižná.

Klíčová kapitola 4 se zabývá frazeologií a vlivem korpusové lingvistiky na ni. Jde tu zvláště o problém variability: autorka píše o různých typech variant (topologické, gramatické a lexikální) a synonymech, uvádí definice z literatury a snaží se také systematizovat nejednoznačnou terminologii: v tomto směru čtenáři velmi pomáhá přehledná tabulka na s. 98. Jasně si uvědomuje variabilitu a synonymii jako lexikografický problém, který má závažné důsledky pro praktické pojetí hesla ve slovníku (jde o strukturu frazémového hesla, vykazuje-li frazém různé varianty) – problém opravdu nezáviděníhodný! Rovněž si je velmi jasně vědoma odlišnosti variability a jiných jevů: modifikace, flexe, nepravidelnosti, restriktce... Zvláštní podkapitolu věnuje i transformaci a jejím jednotlivým typům, jež představují při lexikografické práci problémy. Pozoruhodná je kapitolka o „chybách“ (kap. 4.1.2.5). Doktorandka se dále podrobně zabývá typy variability: variantami morfemickými, topologickými (s velmi názornými příklady!), gramatickými (například ne/přítomnost členu), přičemž uváděné příklady jsou opravdu velmi zajímavé, díky! V pasáži o gramatických variantách nelze než souhlasit s postojem autorky k neutrálnímu tvaru v souvislosti s číslem, velmi zajímavý je příklad # 34 na s. 112 nahoře. Dalším zajímavým a autorkou také poutavě pojatým problémem je negace rozebraná v kap. 4.1.3.5 (s. 113–115) a též kratičká následující kapitolka o alternujících valenčních strukturách. Poměrně hodně prostoru věnuje autorka pochopitelně lexikálním variantám a sémantické variabilitě: uvádí výstižné příklady i podobu dotazů kladených korpusu DeReKo. Příklady si opravdu zaslouží podrobné prostudování, jsou velmi cenné! Na s. 123–127 je vynikající rozbor frazémů s nenápadně polysémiím idiomatickým jádrem *eine Fliege an der Wand*. Bez korpusů, jak píše autorka, by nemohla postihnout všechny významy, o nichž referuje! Naprosto souhlasím se závěrem kapitoly 4.1 na s. 127: poslední odstavec této podkapitoly by doktorandka měla zdůraznit během prezentaci při obhajobě! Opět velice zajímavá je kapitola 4.2 o dalších frazeologických problémech: modifikacích, transformacích, antonymech, nepravidelnostech, omezeních a preferencích – vše doloženo názornými příklady.

Podrobněji se pak autorka věnuje (kap. 4.2.5) specifickým frazeologickým přirovnáním jako významné součásti frazeologie: jejich vydělením, heslovým slovem frazému (Nennform) a dalšími aspekty v přirovnávacích frazémích, kde autorka zdůrazňuje hlavně přehlednost a uživatelskou přívětivost. Podrobně rozebírá dosavadní lexikografickou praxi v různých slovnících a správně naznačuje, jak je důležité přijít s jednotným systémem přehledného řazení frazémů spjatých s jedním klíčovým slovem (zde *Hund*): přehledně pojednáno na s. 147. Autorku chválím i za kapitolku *Přirovnání jako předmět vyhledávání v korpusu / analýzy korpusu* s názornými příklady dotazů a jejich výsledků. Kapitolu 4.2.6 věnuje autorka frazeologickým termínům a propriím, opět dokládá obeznámenost s relevantní literaturou a

správně zdůrazňuje, že: „Sowohl für das produktive als auch für das rezeptive Wörterbuch sind die Termini und Onyme nicht weglassbar“, termíny a onyma však nelze nahradit opisem, jak výtečně dokládá na s. 152–153, kde píše o souhvězdých termínech. Z toho autorce vyplývá, že termíny a onyma nebudou v GAWDT zařazeny do frazeologie, ale budou ve slovníkovém hesle zařazeny mezi kolokace a frazémy.

Kapitola 5 se věnuje korpusové lingvistice a jejímu přínosu pro lexikografické účely. Nejprve doktorandka uvádí nesporné výhody hledání v korpusu s pěkně a opět názorně vyvedenými příklady, poté naopak nevýhody a nedostatky vyhledávacích programů. Na s. 163 se opět v celé své naléhavosti, ba nahotě projevuje absence značkování, zejména disambiguace v korpusu DeReKo: příklad *Dachs* mluví za všechno. Obtíže plynoucí z této absence jsou zřejmé: obrovská pracnost výzkumné práce! Na s. 164 se ukazuje, jak by bylo bývalo dobré, kdyby se ortografická slova *aus* a *hält / hielt* byla značkovala jako příslušející jedinému lexému *aushalten*. Obava, že se autentická textová data takovým značkováním „deformují“ a že značkování snad zmate výzkumníka, je absurdní! Složitější jsou lexikální homonyma (např. *Kiefer*, *Strauss*, (der/das) *Verdienst...*, *travička...*), jež lze disambiguovat bohužel asi až na základě sémantiky kontextu, nikoli jejich formálních vlastností.

V kapitole 5.3 rozebírá doktorandka kookurenční analýzu korpusu DeReKo, parametrizaci různých funkcí – velmi poučné a díky podrobným příkladům jasné a přehledné! V kapitole 5.4 si autorka stýská na zanedbatelné zastoupení mluvených dat v dnešních korpusech a opět na přítomnost homonymie v datech (zde jde o homonymii apelatív a proprií). To je hlavně v němčině vzhledem k jejímu pravopisu substantív opravdu problém, ale za pokus o automatické značkování to jistě stojí, jinak musí výzkumník pracovat s klopotným operátorem NICHT a to je nesmírně pracné. Vždyť toto odfiltrování nežádoucího kontextu lze svěřit programu!

Velmi náležitě a kriticky se autorka vypořádává s údaji o četnosti (dlouhá kap. 5.5), přičemž největší problém pro zadávání korpusových dotazů a jejich frekvenční výsledky tu představuje polylexicalita a variabilita frazémů. Autorka popisuje tři přirozené koncepce: maximalistickou, střední a minimalistickou a s velkým porozuměním věci vystihuje problémy spjaté s každou z nich. V pasáži o minimalistické koncepci se zmiňuje o homonymním tvaru *sein* a správně si stěžuje na absenci úplného morfologického značkování mannheimských korpusů. Předvádí dále složitost hledání frazémů na internetu i v souvislosti s absencí značkovacích programů. Z doložených problémů jí jasně vyplývá nespolehlivost a nízká výpovědní hodnota nalezených dat a souvisejících frekvenčních údajů. Data získaná rešeršemi je nutno pečlivě projít a teprve po namáhavé „ruční“ analýze lze dospět k opravdu spolehlivým frekvenčním údajům.

V kapitole 6 vyvozuje autorka lexikografické důsledky ze své dosavadní praktické lexikografické práce včetně práce s korpusem. Zabývá se tu formou frazému, jeho významem, jeho zachycením ve slovníku, vyjevuje nedostatky v systému odkazování ve slovnících a další nedostatky nejednotného popisu a koncepce slovníků v přístupu k frazémům. V kapitole 6.2 se kriticky zabývá slovníkem DTWP, přičemž mezi důležité aspekty řadí uspořádání frazémů (to zdaleka není problém snadný, jak dokládá na s. 192!), přehlednost a přívětivost slovníku. Zmiňuje se o absenci synonymie v DTWP a neprovázanosti lexikálních variant téhož frazému. Vyjadřuje se i k příkladům a kritizuje fakt, že některé časté frazémy se v tomto slovníku nenevidují. V podkapitole 6.3 se pak zabývá možnostmi a mezemi frazeografického popisu, jež demonstruje na vybraných problémech: abecedním pořadí frazémů v souvislosti s číslem substantív, transformacích (zejména u přirovnání) a příkladech. Zde naprosto správně zdůrazňuje, že „Die Belege selbst sind jedoch keineswegs ohne weitere Bearbeitung als

lexikographische Beispiele anzusehen“ (s. 198), doklady přitom musí pocházet z úzu, nemohou ovšem vplynout do slovníku jako hotové jednotky (tamtéž). Samozřejmě! Zajímavým problémem je i počet a délka příkladů (šíře kontextu), sám se přimlouvám za jejich co největší počet a délku! V podkapitole 6.3.4 se autorka ještě vyjadřuje k frazeologickým ekvivalentům oproti překladům příkladů. Konstatuje ohledně slovníku DTWP, že „nalezení ekvivalentu bohužel neimplikuje jistotu jeho užití ve všech autentických kontextech“, a kritizuje tu přístup užitý v DTWP. Ukazuje dále, že ekvivalentů nelze ve většině kontextů náležitě užít, a proto ve slovníku GAWDT chce autorka užívat co nejvíce systémových i překladových ekvivalentů, opisů, jak vyrozumívám, příkladů, a to v co největší míře. S tím lze jen souhlasit. Dále se rovněž zabývá vztahem lemmatu a frazému, frazeologickým jádrem a jeho lexikografickým zpracováním. Za velmi zajímavou pokládám tabulku frazémů, jež byly odhaleny v korpusech, a nebyly přitom zachyceny ve zkoumaných slovnících.

Po výstižném shrnutí v kapitole 7 následuje v kapitole 8 představení obecných výsledků a požadavků na další výzkum (a patrně i autorčiny plány).

V dodatku je práce vybavena seznamem zkratk, bohatou a relevantní literaturou a konečně třemi přílohami: soupisem zvířecích lexémů, strukturou frazémového hesla a velmi zajímavými zpracovanými lemmaty – jen příklady, více na příloženém CD.

Celkově lze říci, že posuzovaná disertační práce má vynikající kvalitu. Autorka prokazuje, že se ve studované problematice výtečně vyzná, a to jak po stránce teoretické, tak prakticky lexikografické, s lexikografickou prací má zjevně velké zkušenosti. Nejen po obsahové, ale i formální stránce je práce výtečná, vše je přehledné, příklady názorné. Pokud mohu soudit, němčina autorky je bezvadná. Doporučuji práci k vydání, bylo by dobré, kdyby mohla vyjít i česky. Za jediný zápor považuji pár chyb, na něž jsem při četbě narazil, níže jich několik namátkou uvádím (před publikací by se měly odstranit, ve svém exempláři práce jsem je snad všechny vyznačil tužkou):

s. 46, 4. řádek shora  
zeugen → zeugt

s. 50, 10. řádek zdola  
fogende → folgende

s. 58, 2. řádek shora  
dahingehen?

s. 75, 1. řádek zdola  
geplannt → geplant

s. 117, # 52  
sich3?

s. 123, 9. řádek shora  
flach → platt

s. 127, 14. řádek shora  
dtto

s. 148, 10. řádek shora  
Gestaltung → Gestaltung

s. 177, 14. řádek zdola  
Phänomes → Phänomens

s. 183, 7. řádek zdola  
Intertnet → Internet

## **Obecný závěr**

Na závěr konstatuji, že oponovaná disertační práce je vynikající a nanejvýš přesvědčivě dokazuje schopnosti doktorandky samostatně vědecky pracovat a řešit náročné odborné problémy přinejmenším v oblasti lexikologie, lexikografie a korpusové lingvistiky. Bezvýhradně tedy doporučuji, aby jí na základě této práce a ostatních splněných požadavků kladených na doktorské studium byl udělen titul Ph.D.

V Praze 1. srpna 2012

doc. RNDr. Vladimír Petkevič, CSc.

Ústav teoretické a počítačn<sup>í</sup> lingvistiky FFUK