

Richard Změlík

Kvantitativně-korpusová analýza a literární věda

Model a realizace autorského korpusu a slovníku Jana Čepa
v kontextu zahraniční a české autorské lexikografie

[illegible]

Richard Změlík

Kvantitativně-korpusová analýza a literární věda

Model a realizace autorského korpusu a slovníku Jana Čepa
v kontextu zahraniční a české autorské lexikografie

Recenzovali: doc. Mgr. Erik Gilk, Ph.D.
PhDr. Karel Piorecký, Ph.D.

Publikace vznikla v rámci projektu

Inovace bohemistických studií v mezioborových kontextech (CZ.1.07/2.2.00/28.0178).

Tento projekt je spolufinancován Evropským sociálním fondem
a státním rozpočtem České republiky.

Neoprávněné užití tohoto díla je porušením autorských práv a může zakládat
občanskoprávní, správněprávní, popř. trestněprávní odpovědnost.

1. vydání

© Richard Změlík, 2015

© Univerzita Palackého v Olomouci, 2015

ISBN 978-80-244-4766-7

Motto:

*Without indulging the anthropomorphic hopes or fears
that lie behind many popular accounts of “artificial intelligence”
and virtually mysticize the machine, we must recognize
that a number of existing computer programs [...]
make it possible to unite classes in super-classes,
to form categories of categories with an increasing degree of generality.**

(John F. Burrows: *Computers and the Study of Literature* 1992, s. 170)

* Aniž bychom se oddávali antropomorfním nadějím či obavám, které stojí za mnoha všeobecně rozšířenými výklady „umělé inteligence“ a které činí ze stroje cosi takřka mystického, musíme uznat, že řada existujících počítačových programů [...] umožňuje spojovat skupiny do vyšších skupin, vytvářet kategorie kategorií se stále rostoucí mírou generalizace.

Obsah

Předmluva.....	9
1. Autorské slovníky z hlediska základní metodologie a typologie.....	11
1.1 Úvodem	11
1.2 Metodologická východiska autorské lexikografie	14
1.3 Vybrané příklady autorských slovníků (stručný historický přehled)	37
1.3.1 Cizojazyčné autorské slovníky	38
1.3.1.1 Konkordanční slovníky	39
1.3.1.2 Výkladové slovníky (v redukované podobě)	48
1.3.1.3 Výkladové slovníky (v neredukované podobě)	55
1.3.1.4 Strojově zpracované slovníky.....	66
1.3.2 České autorské slovníky	67
1.3.2.1 Strojově zpracované slovníky (frekvenční).....	68
2. Autorský korpus a slovník Jana Čepa: cíle, koncepce, způsob realizace a předpoklady využití v literárněvědném kontextu	77
2.1 Úvodem	77
2.2 Základní metodologická východiska	79
2.3 Návrh a organizace autorského korpusu.....	80
2.3.1 Relativní chronologie vzniku jednotlivých povídkových souborů.....	86
2.3.1.1 Strukturace korpusu fikčních narativů (model relativní chronologie)	95
2.3.2 Základní osnova frekvenčního slovníku.....	100
2.3.3 Segmentace na úrovni jednotlivých (sub)korpusů	102
2.3.4 Kontext a sémantika.....	107
2.3.5 Strukturace a hierarchizace autorského korpusu	114
2.4 Závěrem	116
3. První tři Čepovy povídkové knihy z hlediska kvantitativně-korpusové analýzy tematických oblastí	121
3.1 Metodologie	123
3.2 Vyhodnocení	131
3.3 Kolokace vybraných lemmat tematických oblastí.....	144
3.4 Strukturace sémantických tříd substantiv a jejich vztahy, základní situace adjektiv a sloves	166
3.5 Literárněvědné aspekty (vypravěč, tematika, motivika, čas a prostor)	195
3.6 Synekdocha jako ústřední organizující princip.....	210
4. Závěr	215
Summary.....	223
Ediční poznámka.....	225
Bibliografie	227
Jmenný rejstřík.....	241

Předmluva

Předkládaná publikace si klade dva základní cíle: na konkrétním příkladu autorského korpusu Jana Čepa prezentovat nejen první badatelské výsledky praktické povahy, které směřují k vytvoření základního autorského frekvenčního slovníku, ale především reflektovat teoretickou rovinu, jež je vymezena otázkou možného zapojení kvantitativně-korpusové analýzy do kontextu literárněvědného bádání. Volba autora byla určena výzkumným zadáním na katedře bohemistiky Filozofické fakulty Univerzity Palackého v Olomouci. I když při koncepci autorského korpusu byla primárním požadavkem jeho praktická využitelnost v literární vědě, je nejen samozřejmé, ale i užitečné, že výsledný produkt může najít uplatnění i v lingvistice, zejména v oblasti analýzy autorského jazyka či stylu. Součástí monografie je rovněž první verze Čepova frekvenčního slovníku, který vychází z autorova korpusu, jenž zahrnuje Čepovu beletrii, tzn. povídkové soubory vydávané v letech 1926–1953 (máme na mysli první vydání), spolu s časopiseckými povídkami, které nebyly Čepem zařazeny do žádného knižního souboru.

Předkládaná publikace by měla plnit několik funkcí. Především jde o představení specifického projektu autorského korpusu Jan Čepa, s čímž je spojen způsob řešení jeho jednotlivých částí a segmentů. Zejména však jde o zkoumání možností, které takováto koncepce nabízí literární vědě. Nezbytné je ovšem hned zpočátku zdůraznit, že prezentovaný model i metodologické kroky neznamenaají popření či vytěsnění standardních literárněvědných postupů. Smyslem této práce je snaha o zapojení moderních kvantitativních metod aktuálně využívaných v lingvistice do oblasti literární vědy obecně, nikoli substituovat vlastní literárněvědné postupy. Jak výstižně konstatoval Franco Moretti, samotná kvantitativní analýza bez relevantní interpretace ztrácí smysl.¹ Jelikož se jedná o postupy, které tradičně nenáleží k literárněvědným metodám, jde rovněž o svého druhu experiment, na jehož základě se pokoušíme promýšlet funkčnost možného propojení konkrétních literárněvědných metod s moderními pracovními technikami kvantitativní a korpusové analýzy. Cílem nebylo vytvořit čepovskou monografii v pravém slova smyslu, jakou je například kniha Tomáše Kubíčka, ale v prvé řadě monografii „o metodě“. Neusilujeme přinést v pravém slova smyslu faktografické či literárněhistorické poznatky o Čepově díle, i když se domníváme, že učiněné analýzy prvních

¹ „Jak jsem již uvedl, kvantitativní výzkum si klade za cíl poskytnout svá data následné specifické interpretaci, a toto je přirozeně také jeho mez: poskytnout data, nikoli interpretaci samotnou.“ (Moretti 2014, s. 16)

tří Čepových povídkových knih mohou přinést inovativní (empiricky fundovaný) pohled na strukturní mechanismus, který je spojuje, a současně na jeho pozadí odkrýt a vysvětlit jedinečnou tvůrčí strategii, jež se v těchto jednotlivých knihách i celém souboru *Zeměžluč* realizuje. Primárním cílem (jak bylo konstatováno) je zkoumat možnosti funkčního propojení kvantitativních metod s metodami literárněvědnými.

Monografie je strukturována do tří oddílů, které představují relativně samostatné, avšak vnitřně provázané části, jež prezentují vlastní model slovníku v kontextu zahraniční i české autorské lexikografie a současně ukazují první konkrétní výsledky jmenovaného projektu a jejich interpretaci. Vlastní realizace slovníku se neobešla bez moderních technologických postupů, jež v tomto případě představují především počítačově zpracovaná jazyková data. Dílčí kroky a postupy jsou popsány a vysvětleny na příslušných místech této knihy.

Celý projekt by nebylo možné realizovat bez přispění těch, kteří byli autorovi nápomocni cennými konzultacemi či technickým zabezpečením. Na tomto místě bychom proto rádi poděkovali Františku Čermákovi (ČNK), Michalu Křenovi (ČNK), Petru Plecháčovi (ÚČL AV ČR) a Janě Hajdové (FF UP).

U cizojazyčných citací je vždy v poznámce pod čarou připojen český překlad. Dané řešení spočívá v usnadnění porozumění cizojazyčným pasážím, které v textu uvádíme v originálu. Překlad je motivován praktickými potřebami, není tedy normativní a jedinou možnou variantou. Za korekce překladů děkuji Ondřeji Kolářovi, Jaroslavě Kolářové, Filipu Krajníkovi (FF MU) a Martinu Lukášovi (FF UP).

Korpus Čepovy beletrie je volně přístupný na adrese WWW: <slovníkcep.g6.cz/index.php>. Součástí monografie je CD, které obsahuje *Frekvenční slovník Jana Čepa (beletrie)* a *Frekvenční slovník Jana Čepa (hapaxy)*.

1. Autorské slovníky z hlediska základní metodologie a typologie

1.1 Úvodem

Smyslem autorského slovníku je zmapovat jazykovou (převážně lexikální) situaci v konkrétní a jasně vymezené oblasti, která je definována slovesnou tvorbou jednoho autora. V tom spočívá první výhoda takového slovníku oproti slovníkům, které reflektují a zpracovávají národní jazyk jako celek. V případě autorského slovníku se totiž jedná o soubor jazykových jednotek, jevů a prostředků, který lze považovat za definitivní a uzavřený, což v případě národního jazyka přirozeně neplatí.² Uzavřenost autorského díla tedy představuje významný aspekt pro lexikografickou práci, která tak může podat ucelený, kompletní a systematický přehled o vymezené jazykové situaci. Současně zde má svůj význam i to, že jazyk konkrétního autora je projevem fixovaným, a tedy nejen dobře dostupným a technicky zpracovatelným, ale i do velké míry koherentním (viz Mattausch 1990, s. 1549, 1550, 1553), což vyplývá z příslušnosti jazykových prostředků do oblasti definované jejich užitím konkrétním autorem a určitým způsobem (tvorba literární fikce). Tato koherence ovšem není synonymem unifikace a nivelizace; naopak lze předpokládat, že přehled celkové tvůrčí situace bude vykazovat řadu odchylek, ruptur a modifikací. A právě ve spojení výše zmíněné jednoty zaručené autorským subjektem a rozdílnosti v rámci celku umělecké produkce (diference textů) dostává autorský slovník svůj specifický smysl. Jak bylo řečeno, autorský slovník tak podává nejen relativně úplný přehled o jazykových prostředcích, jež určitý autor použil,³ ale současně (z povahy materiálu a situace) dokáže podat konzistentní model způsobu zacházení s jazykem, jenž není narušován například mnohočetností uživatelů (naopak se tu projevuje již zmíněná tendence ke koherentnosti). Takový slovník tedy na první pohled může plnit několik funkcí. Informuje nejen o slovní zásobě daného autora, o jeho zacházení s jazykem, které systematicky kategorizuje a uspořádává do zvoleného lexikografického formátu, ale rovněž poskytuje přístup ke způsobům, jakými je s jazykem zacházeno, jakými prostředky disponuje z hlediska kvalitativního i kvantitativního, zjišťuje typy strukturních

2 Viz pozn. 51.

3 Úplnost autorského korpusu závisí na několika faktorech. Jak bude řečeno dále, jedním z kritérií je volba typu autorského slovníku. Podstatnou úlohu sehrává i to, do jaké míry je autorovo dílo kompletní, popřípadě jsou-li známé a dostupné rukopisy.

kombinací lexikálních jednotek a vymezuje je vůči společnému základu (tj. národnímu jazyku); dále umožňuje systematicky zkoumat a poznávat zákonitosti konkrétního uměleckého stylu s ohledem na vývojová tvůrčí období či žánry, v neposlední řadě dovoluje prohloubit obecné znalosti o jazykové situaci určitého období (Ibidem, s. 1552). Autorský slovník je tedy na jedné straně mapováním uzavřené jazykové (umělecké) situace vybraného autora a současně sondou do situace národního jazyka daného období, takže se přirozeně stává součástí širších lexikografických projektů zaměřených právě na národní jazyk jako celek. Jeho těžiště spočívá především v tom, že doplňuje a obohacuje poznání týkající se stratifikace a funkčního rozvrstvení národního jazyka. Autorské slovníky však mají ještě další významy a možnosti využití, které se odvíjely od metodologií, jimiž byly zpracovány, a zejména od možnosti využít moderních technologií a spolu s nimi i statistických výpočtů speciálně navržených pro kvantitativní analýzu uměleckého jazyka a textů.

Jeden z hlavních důvodů sestavování prvních autorských slovníků spočívá bezpochyby ve významu autora, jehož slovesné umění představuje zásadní a výrazný vklad do vývoje literatury domácí či světové.⁴ Příkladem jsou autoři tzv. klasické literatury.

4 Ve výsledku se většinou nejedná pouze o oblast slovesného umění, ale o celou oblast jazykové situace daného období, vůči které je takové významné dílo (či soubor děl) chápáno jako její reprezentativní projev. Dobře to dokládá například situace kolem *Słowniku języka Adama Mickiewicza*: „Na tym owa wewnętrzna historia polegała, że od pierwszych dzieł Mickiewicza jawne było i oczywiste, iż poeta ten przemienił do rdzenia poezję narodową, ponieważ do podobnego rdzenia przemienił słowo poezji tej służące. Ponieważ nowe słowo poezji narodowej wywiódł z samego gruntu podstawowego, z tego co w języku narodowym jest wspólne i niezmiennie, jakiegokolwiek ten język pełni funkcje i zadania. Nie rola romantyka, pod którego piórem słowo łśni niespodzianką i samowolnym jego użyciem, lecz rola prawodawcy stała się jego udziałem.” (Wyka 1962, s. V)

[„Tato vnitřní historie (autor má na mysli projekt Mickiewiczova slovníku – pozn. R. Z.) spočívala v tom, že od prvních Mickiewiczových děl bylo zjevné a samozřejmé, že tento básník od základu proměnil národní poezii, neboť od podobného základu proměnil slovo sloužící této poezii. Jelikož nové slovo národní poezie vyzvedl ze samého základu, z toho, co je v národním jazyce společné a neměnné bez ohledu na to, jaké tento jazyk plní funkce a úkoly. Jeho údělem nebyla role romantika, pod jehož perem se slovo třpytí neočekávaností a svým svébytným užitím, ale role zákonodárce.”]

Analogická je situace i u dalších prací tohoto typu: „Za wydaniem słownika Kochanowskiego przemawia jednak, po pierwsze, pozycja tego poety w literaturze polskiej: był on największym na przestrzeni paru wieków (do Mickiewicza) poetą polskim, na jego utworach wzorowało się wielu poetów i pisarzy, są te utwory i pozostaną lekturami szkolnymi.” (Kucala 1994, s. V)

[„Pro vydání slovníku Kochanowského hovoří jednak postavení básníka v polské literatuře: po řadu staletí byl největším polským básníkem (do Mickiewicze), jeho díla představovala vzor pro řadu básníků a spisovatelů, tato díla jsou a zůstanou školní četbou.”]

Podobně je tomu i v případě dalšího polského slovníku, který excerpuje lexikum Cypriana Norwida: „Prace nad *Słownikiem Cypriana Norwida* zostały rozpoczęte przy Wydziale Polonistyki Uniwersytetu Warszawskiego jesienią 1983 r. Motywem ich podjęcia było przekonanie o ważności spuścizny Norwida dla kultury polskiej, ważności przede wszystkim w zakresie treści, jakie ona niesie, ale także formy – języka poetyckiego, stylu epistolarnego i publicystycznego Norwida. Ta forma – trudna, a bardzo świadomie przez Norwida konstruowana – stanowi o potrzebie prac słownikowych i szerzej – prac nad językiem autora „Rzeczy o wolności słowa.” Wydają się one konieczne po to, by pełniej zrozumieć treści zawarte w jego utworach, a także, by rozpoznać i ocenić blask i cienie jego kunstownego języka poetyckiego, oraz osobliwości różnych gatunków jego prozy.” (Chojak – Puzynina – Teleżyńska – Wiśniewska 1988, s. 1)

V 19. století vznikla řada slovníků věnovaných takovým autorům, jakými byli Homér, Aischylós, Sofoklés, Tacitus a mnozí další.⁵ Způsob, jakým jsou tyto slovníky obvykle sestaveny, zdůrazňuje především slovní zásobu spolu s lexikálními významy a kontexty, v jakých je daných slov užito. Vedle lexikografické či lingvistické funkce plnila tato díla úlohu reprezentativní a didaktickou, neboť systematickým zpracováním lexika vybraných autorů poskytovala schematický obraz vyspělého klasického jazyka a stylu. Rovněž při studiu antické literatury se tyto práce ukazují jako velice užitečné.⁶

Jaký vliv měl jazyk klasických básníků na soudobou poezii, ukazuje i příklad z historie českého básnictví. V první polovině 19. století došlo ke známému sporu o českou prozodii, ve kterém šlo o ideové prosazení dvou metrických systémů, z nichž první (časoměrný) byl inspirován latinskými verši klasiků (k obhájcům a propagátorům tohoto systému patřili například Václav Thám, František Palacký, Pavel Josef Šafařík, Václav Stach), zatímco sylabotónický verš respektoval fonologickou povahu národního jazyka (stoupenci sylabotónismu v českém básnictví byli kupříkladu Josef Dobrovský či Antónín Jaroslav Puchmajer).

[„Práce na *Slovníku Cypriana Norwida* započaly na podzim roku 1983 na Fakultě polonistiky Varšavské univerzity. Motivem započetí bylo přesvědčení o významu Norwidova dědictví pro polskou kulturu, o důležitosti zejména obsahu, který toto dědictví přináší, ale též formy – básnického jazyka, epistolárního a publicistického stylu. Tato forma – obtížná a Norwidem velice uvědoměle konstruovaná – dokládá potřebu slovníkových prací a šířeji – prací o autorově jazyce ‚Řeči o svobodě slova.‘ Ukazují se jako nezbytné proto, abychom lépe porozuměli obsahu jeho tvorby, a také proto, abychom byli schopni rozpoznat a ocenit světla a stíny jeho uměleckého básnického jazyka a také specifičnosti různých žánrů jeho prózy.“]

5 *Lexicon Aeschyleum* (1830), *Lexicon Homericum* (1885), *Lexicon Taciteum* (1830).

6 Christa Dill ve své studii *Lexika zur einzelnen Schriftstellern* (1959) rovněž uvádí jako jeden z důvodů vzniku nejstarších autorských slovníků *de facto* hledisko hermeneutické: „Das semasiologische Wörterbuch zu einzelnen Dichtern ist nicht Neues. Für den Altphilologen waren solche Wörterbücher stets ein notwendiges Hilfsmittel zum richtigen Verständnis der Texte, die in einer heute nicht mehr lebendigen Sprache überliefert sind. Diese Wörterbücher sind entweder einsprachig, wie z. B. das lat.-lat. *Lexicon Petronianum* von J. Segebade und E. Lommatzsch (Leipzig 1898), oder zweisprachig, wie das griech.-lat. *Lexicon Plutarcheum* von D. Wyttenbach (Leipzig 1843) oder wie das lat.-deutsche *Vollständiges Lexikon zu den pseudo-cäsarianischen Schriftwerken* von S. Preuss (Erlangen 1884). Eine ganz moderne Arbeit auf diesem Gebiet ist das im Akademie-Verlag erscheinene *Polybios-Lexikon* von Mauersberger.“ (Dill 1959, s. 342)

[„Semasiologické slovníky věnované jednotlivým básníkům nejsou ničím novým. Pro starší filology byly takové slovníky nezbytnou pomůckou pro správné porozumění textům, které jsou dědictvím dnes již neživého jazyka. Tyto slovníky jsou buď jednojazyčné, jako např. latinsko-latinský *Lexicon Petronianum* J. Segebadeho a E. Lommatzsch (Lipsko 1843), nebo dvojjazyčné, jako řecko-latinský *Lexicon Plutarcheum* D. Wyttenbacha (Lipsko 1843) nebo jako latinsko-německý *Úplný lexikon pseudocísařských písemných prací* S. Preusse (Erlangen 1884). Poměrně moderní práci z této oblasti je v Akademickém nakladatelství vydaný Mauersbergerův *Polybios-Lexikon*.“]

1.2 Metodologická východiska autorské lexikografie

O charakteru a zaměření prvních autorských slovníků pojednává řada lexikografů, mezi mnohými i Josef Mattausch, který ve studii *Textlexikographische Aspekte im Autorenwörterbuch – am Beispiel des Goethe-Wörterbuchs* (1991) konstatuje: „Neben den großen Leistungen der klassischen Philologie begegnen uns seit der Mitte des vorigen Jahrhunderts und besonders in seinem letzten Drittel bedeutende Lexika zu Dante, Shakespeare und zu den französischen Klassikern. Diese Entwicklung setzte sich in unserem Jahrhundert in bemerkenswerter Weise in den ost- und südosteuropäischen Ländern fort, wo nach dem Zweiten Weltkrieg wesentliche Nationalautoren wie Puschkin, Ševčenko, Gorki, Mickiewicz, Pasek, Petöfi, Juhász, Eminescu u. a. Gegenstand gewichtiger Wörterbuchunternehmungen, meist nationaler Akademieprojekte, wurden.“⁷ (Mattausch 1991, s. 713) Výrazněji se ovšem podle Mattausche začalo nad problematikou autorských slovníků diskutovat zejména od 60. let 20. století, a to především v souvislosti s přípravami ke slovníkům Puškina, Gorkého a Juhásze (Ibidem, s. 714–715). V této době zaznamenáváme také první práce v českém prostředí, kde je zájem motivován zejména novými možnostmi kvantitativní lingvistiky, jež od 60. let 20. století zaznamenává výrazný vzestup jak v oblasti teoretické, tak aplikační.⁸ V 50. a 60. letech 20. století vznikají projekty velkých autorských slovníků; v této souvislosti se nejčastěji uvádí například Puškinův slovník nebo slovník Maxima Gorkého (Ibidem 1991, s. 715). Hlavním smyslem takových projektů je systematicky a pokud možno v úplnosti zpracovat autorské lexikum, které svým způsobem reprezentuje kulturní vyspělost národa.

K základním předpokladům lexikografické práce v obecném smyslu⁹ podle Mattausche náleží otázka vymezení rozsahu excerpovaného slovesného materiálu, který – jak jsme již naznačili – může mít ve výsledku nejen různý rozsah, ale i kvalitu. Této problematice se Mattausch podrobněji věnuje ve studii *Das Autoren-Bedeutungswörterbuch* (1990), kde uvádí, že autorský korpus je možné koncipovat různými způsoby; lze vymezit a specifikovat pouze určitou část autorského korpusu, definovanou například žánrovým rozsahem, a zkoumat lexikální či sémantické zákonitosti s ohledem na funkční potenciál

7 [„Vedle velké oblasti významných děl klasické filologie se od poloviny minulého století, a zejména v jeho poslední třetině setkáváme s významnými slovníky k Dantovi a Shakespearovi a se slovníky francouzských klasiků. Tento vývoj pokračoval pozoruhodným způsobem i v našem století a v zemích východní a jihovýchodní Evropy, kde se po druhé světové válce stalo prioritou koncipovat projekty–slovníky (většinou se jednalo o akademické projekty) významných národních autorů, jakými jsou Puškin, Ševčenko, Gorkij, Pasek, Petöfi, Juhász, Eminescu a jiní.“]

8 Podrobněji se problematice zahraniční kvantitativní lexikografie věnujeme v kapitole 1.3.1.4, na české autorské slovníky zpracované pomocí výpočetní technologie se soustředíme v kapitole 1.3.2.1.

9 Oblast lexikografie, která se – jak zdůrazňuje Mattausch a jiní – zabývá autorskými slovníky, náleží pod obecnou lexikografii (Mattausch 1991, s. 715; Roelcke 1994, s. 2; Wiegand 1986, s. 163).

takové skupiny. Mattausch jako příklad uvádí Merguetův Ciceronův slovník,¹⁰ který zpracovává pouze vybranou část Ciceronových textů: „Merguet 1877/84 z. B. enthält die Ciceronianischen Reden, nicht aber die philosophischen und rhetorischen Schriften und die Briefe.“¹¹ (Mattausch 1990, s. 1550) Jak uvidíme dále, je možné v případě autorských slovníků volit různé jeho formy, způsoby zpracování, metody i rozsah excerpovaného materiálu s ohledem na specifická kritéria, která mohou být dána rovněž praktickou povahou, kdy je potřeba vydat autorský slovník v redukované formě, nebo odborným zřetelem, jenž naopak může vyžadovat komplexní zpracování slovesného díla s důrazem na konkrétní metodologii.¹² V pozadí metodologické volby a formálně-typologických kritérií slovníku stojí v první řadě jazykové aspekty. Případné další nároky kladené na tyto práce jsou upozaděny, s čímž souvisí i jejich literárněvědné využití. Mezi zahraničními i českými autorskými slovníky jsme se dosud nesetkali s projektem, který by kromě přirozených lingvistických kritérií byl soustředěn rovněž na některé ze základních literárněvědných kategorií (žánr, narativní kategorie, tvůrčí etapy nebo historickovývojové fáze apod.). Takový záměr zřejmě v minulosti představoval nemalý problém, který souvisel s dodnes velmi diskutabilní otázkou specifické formalizace příslušných literárněvědných kategorií, jež v našem případě vyžadují speciální schematizaci vyžadovanou slovníkovým formátem.

K jednomu ze základních typů autorských slovníků náleží výkladový slovník. Je-li excerpováno celé autorovo dílo a maximálně zpracovaný a uspořádaný interpretační a výkladový aparát, vytváří podobu tzv. thesauru,¹³ jehož základní princip se dle Mattausche zakládá na „Vollständigkeit in der Wortaufnahme, Vollständigkeit in den Bedeutungserklärungen, vollständige oder selektive (repräsentative) Textbelegung“¹⁴ (Mattausch 1991, s. 716).

10 Celý název zní *Lexikon zu den Reden des Cicero, mit Angabe sämtlicher Stellen* (1877).

11 [„Merguet například zpracovává Ciceronovy řeči, nikoli ale jeho filozofické a rétorické pojednání a dopisy.“]

12 V dalším výkladu se seznámíme s různými typy autorských slovníků a uvedeme, že jedním z rozhodujících kritérií je zde především metodologická otázka, resp. způsob, jakým je slovník koncipován.

13 „Ein thesaurisch angelegtes Wörterbuch – als Werk-, Werkgruppen- oder Gesamtwerkthesaurus – beruht auf dem Prinzip der Vollständigkeit, sowohl was die Ausschöpfung des zugrunde gelegten Textkorpus betrifft (Totalexcerption) als auch hinsichtlich seiner lexikographischen Behandlung (Vollständigkeit in Lemma- und Belegbestand, nicht selten auch in der Bedeutungsdarstellung).“ (Mattausch 1990, s. 1550)

[„Slovník typu thesauru, jakým je slovník excerpující jedno dílo, skupinu děl nebo všechna díla, se zakládá na principu úplnosti, a to jak s ohledem na vyčerpání textového korpusu, který je jeho základem (totální excerpce), tak s ohledem na jeho lexikografické zpracování (komplexnost lemmat, jejich umístění v korpusu, nezřídka také s jejich významovým určením).“]

14 [„úplnosti slovní zásoby, úplnosti, co se týče vysvětlení významu, úplném nebo výběrovém (reprezentativním) výskytu pojmu v korpusu“]

O klasifikaci autorských slovníků pojednává řada převážně zahraničních lexikografických prací, jež se na tuto problematiku zaměřují (např. Herbert Ernst Wiegand, Christa Dill, Manfred Gräfe, Thorsten Roelcke aj. – viz dále). Jestliže jsme v úvodu konstatovali, že první autorské slovníky vznikaly již koncem 19. století a jejich významný nárůst lze zaznamenat od poloviny 20. století, neznamená to zcela automaticky, že produkce tohoto typu slovníku je všeobecně rozšířená. Ještě v 80. letech 20. století Josef Mattausch v jiné své studii nazvané *Kommunikativ-pragmatische Sprachbetrachtung und Lexikographie: Zur Spezifik des Autorenwörterbuchs* (1982) uvádí, že zájem o problematiku autorských slovníků je v německy mluvících zemích poměrně malý: „Es verwundert angesichts dieser Verschiedenartigkeit der Gegenstandsbereiche und Interessenlagen nicht, daß das Autorenwörterbuch als eigener lexikographischer Typus – trotz zunehmender theoretischer Reflexion der Wörterbucharbeit – bisher nur sporadisch, im deutschsprachigen Raum fast keine Aufmerksamkeit gefunden hat.“¹⁵ (Mattausch 1982, s. 304) Také u nás došlo k částečnému přerušení vývoje, který byl započat v 60. letech 20. století úsilím o sestavení moderního autorského slovníku za využití tehdy aktuálních a progresivně se vyvíjejících technologií. Teprve v 90. letech se znovu objevily české autorské slovníky využívající metod kvantitativní nebo postupně se formující korpusové lingvistiky.

Vrátíme-li se však zpět k metodologickým a taxonomickým otázkám, je nezbytné představit několik základních pohledů a přístupů, které nacházíme v pracích zahraničních lexikografů. Autorem prvního zde prezentovaného modelu taxonomie autorských slovníků je již několikrát zmíněný J. Mattausch, jehož dělení výkladového typu autorského slovníku je určeno rozsahem zpracovaného (excerpovaného) slovesného materiálu.

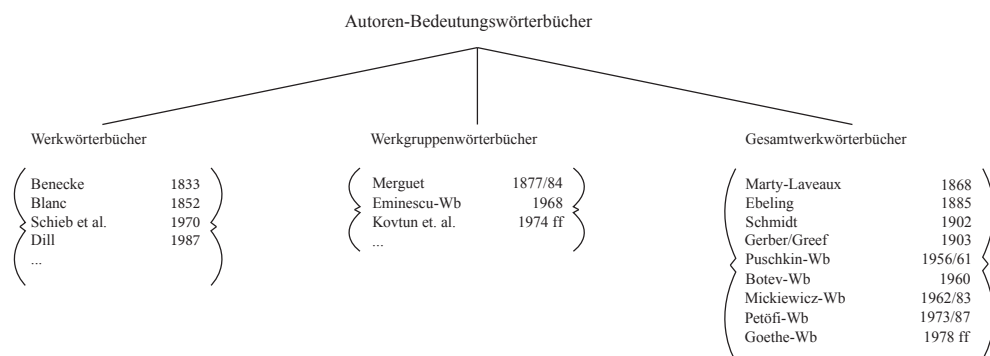


Schéma 1: Taxonomie výkladových autorských slovníků podle J. Mattausche zohledňující rozsah excerpovaného materiálu dle formálních hranic. Schéma přebíráme ze studie J. Mattausche *Das Autoren-Bedeutungswörterbuch* (Mattausch 1990, s. 1549).

¹⁵ [„Vzhledem k rozmanitosti tematických oblastí a zájmu o ně neudivuje, že i přes rozvíjející se teoretickou reflexi lexikografické činnosti je autorskému slovníku coby vlastnímu lexikografickému typu věnována malá pozornost, v německy mluvících zemích téměř žádná.“]

Dle uvedeného schématu (Schéma 1) jsou autorské výkladové slovníky (*Autoren-Bedeutungswörterbücher*) členěny do tří základních skupin: slovníky založené na zpracování jednoho díla (*Werkwörterbücher*), skupiny děl (*Werkgruppenwörterbücher*) a slovníky, jejichž bázi tvoří kompletní autorova slovesná tvorba (*Gesamtwerkwörterbücher*). Jedná se tedy nikoli o typologicky odlišné slovníky, ale o modifikace téhož slovníkového typu. Při volbě konkrétního způsobu zpracování ovšem musí být zachována patřičná míra reprezentativnosti lexikografického produktu, zejména jeho metodologické oblasti. To se týká především kvality metajazykového aparátu (komentářů), který musí vycházet ze základních lexikografických kritérií. K těm náleží přiměřenost vzhledem k rozsahu a způsobu komentáře, dále jeho formálně-stylistické aspekty, tj. odbornost a explikativnost a podobně i jednoznačnost výkladu, který musí především jasně odlišit a definovat sémantické rozdíly. Tato skutečnost je velice důležitá, neboť se týká specifického materiálu, kterým je umělecký, poetický jazyk, jenž disponuje řadou rysů, jež jsou podmíněny nejen vlastním autorským stylem, ale např. žánrem, historickým kontextem, imanentními vývojovými aspekty apod. Autorské lexikum je ve slovníku výkladového typu reflektováno na úrovni systému národního jazyka a současně na bázi systému utvářeného literárními texty, slovy Jurije M. Lotmana tzv. sekundárního modelujícího systému.

Obrátíme-li pozornost k uvedené taxonomii autorských výkladových slovníků, je pro nás důležité to, co Mattausch konstatuje o každém z výše uvedených modelů. Ve všech případech platí, že je lze realizovat v redukované formě, což se týká nejen rozsahu excerpovaného materiálu (velikost autorského korpusu), ale i způsobu realizace výkladového (metajazykového) aparátu. Následující Mattauschovo schéma ukazuje kritéria těchto možností, jež tvoří škálu odstupňovanou z hlediska zvyšující se intenzity (zleva doprava) metajazykových informací. Poslední (pravý) sloupec představuje takový typ výkladového autorského slovníku, který se blíží thesauru v pravém slova smyslu. Ovšem jak bylo řečeno, i pro redukované verze autorského slovníku platí zásady o způsobu, jakým formulovat výkladové pasáže, které významně spoluurčují celkový i dílčí pohled na autorský jazyk.¹⁶ Na uvedeném Mattauschově schématu jsou tři kategorie výkladového autorského slovníku; každá z nich představuje paradigmatický typ určený mírou redukce metajazykových informací. Do levého sloupce Mattausch umístil slovníky, které redukují jmenovaný druh informace v oblasti rozsahu lemmat, jejich sémantického výkladu a citací. Prostřední sloupec, ve kterém se nacházejí slovníky, jež mají blízko k typu thesauru, tvoří práce, jejichž komentující aparát je částečně redukován z hlediska citací (doložení hesel v kontextu) a výkladu. V pravém sloupci jsou potom slovníky, které Mattausch *de facto* již považuje za thesaury, čemuž odpovídá jejich povaha z hlediska

16 Srov. Umbach 1986, s. 162.

rozsahu excerpovaného materiálu, který představuje kompletní autorské lexikum, dále způsobu jeho doložení v rámci kontextu a současně i charakter výkladového aparátu.

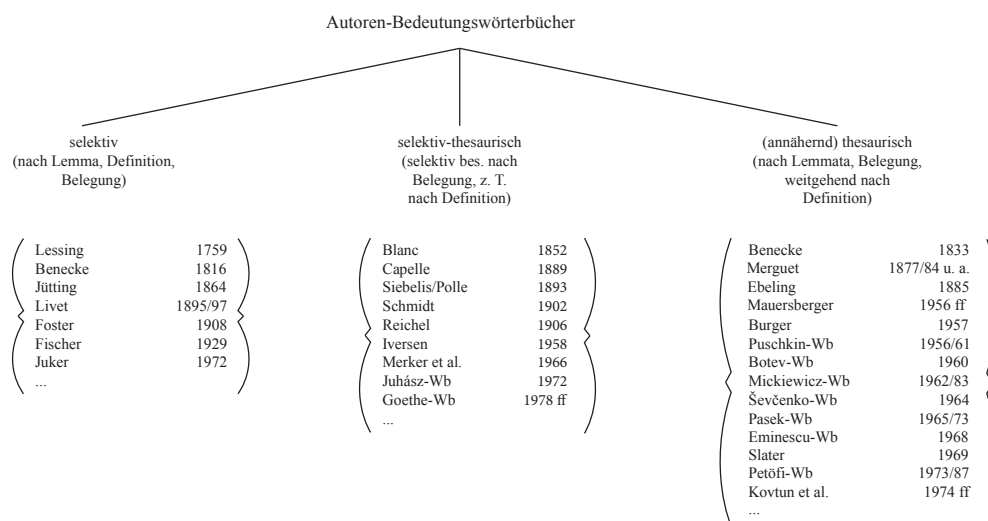


Schéma 2: Taxonomie výkladových autorských slovníků podle J. Mattausche, která vychází z rozsahu a způsobu zpracování slovesného materiálu. Schéma přebíráme ze studie J. Mattausche *Das Autoren-Bedeutungswörterbuch* (Mattausch 1990, s. 1550).

Hovoří-li Mattausch o autorském slovníku, rozlišuje mezi pojmy autorský slovník (v širším slova smyslu) a autorský thesaurus: „Thesaurische Wörterbuchformen begnügen sich nicht wie andere lexikographische Formen mit einem so oder so verkürzten, präparierten Wortbestand und häufig auch nicht mit einer engen Bedeutungsauswahl, sie bilden vielmehr – wenn auch im Rahmen eines begrenzten Textkorpus – die ganze lebendige Sprachwirklichkeit ab.“¹⁷ (Mattausch 1990, s. 1553) Další členění autorských slovníků podle Mattausche závisí na tom, zdali se jedná o jednojazyčný či vícejazyčný slovník, popř. se odvíjí od historické perspektivy excerpovaného díla. Mezi další způsoby, o kterých Mattausch hovoří, náleží takové koncepce slovníku, jež se řídí abecedním řazením nebo řazením dle specifických kritérií, např. na základě synonym. Současně však uvádí, že komplexní realizaci autorského slovníku představuje takový typ, jenž obsahuje zejména komentáře k významu slov a jehož reprezentativní podobou je zmíněný thesaurus (Ibidem, s. 1552). S tím jsou úzce spojeny i další požadavky a možnosti autorského slovníku, který může uvádět informace o stylistice, poetice, komunikačně-pragmatických aspektech, ale rovněž údaje o historickém kontextu, způsobu použití výrazu

17 [„Slovník typu thesaurus se podobně jako jiné lexikografické formy nespokojí s tak či onak redukovanou, preparovanou slovní zásobou a často rovněž ani s omezeným významovým výkladem. Tento typ slovníku odráží mnohem více – i když je tomu tak v rámci ohraničeného textového korpusu – totiž celou živou jazykovou skutečnost.“]

v odlišném funkčním kontextu apod. Je rovněž možné, jak Mattausch dodává, sledovat literární aspekty jako motivy, témata či toposy (Ibidem, s. 1554–1555). Neméně důležitou oblast autorského slovníku tvoří podle Mattausche vlastní jména, která mohou být jednoduše inventarizována nebo naopak podrobněji komentována (Ibidem, s. 1556). Mattausch do této oblasti řadí jména postav a literárních míst stejně jako mytologická a historická jména.¹⁸

Při koncipování autorského slovníku platí všeobecné zásady, které v souvislosti s obecnou lexikografií formuloval rovněž Franz Josef Hausmann ve studii *Wörterbuchtypologie* (1989): „Die Frage: Wozu dient das Wörterbuch? führt zu einer Funktionstypologie. Die Frage: Wie sieht das Wörterbuch aus? führt zu einer phänomenologischen Typologie des Wörterbuchs.“¹⁹ (Hausmann 1989, s. 970) Hausmann tímto rozumí členění slovníku dle praktických a teoretických (lexikografických) aspektů. Mezi první náleží oblast, na kterou je slovník zaměřen. Podle Hausmanna se obecně jedná o slovníky orientované na určitou problematiku, např. slovníky vymezené pro konkrétní obor nebo slovníky používané při praktickém psaní nebo systematické výuce (Ibidem, s. 970). Fenomenologická typologie, kterou Hausmann uvádí jako druhé kritérium při tvorbě slovníků, představuje způsob, jakým je zpracováno a uspořádáno lexikum z teoretického hlediska. Jedná se o strukturaci lexikální databáze dle formálně-teoretických aspektů, mezi které náleží například segmentace dle slabik, kořenů, afixů, zkratk nebo slovních druhů apod. (Ibidem, s. 971). Hausmannova klasifikace autorských slovníků vychází ze základních distinktivních rysů, které na jedné straně definuje jazyk a množina textů, na straně druhé lexikografická kritéria. Dle jazykové materie Hausmann rozlišuje slovníky jednojazyčné či vícejazyčné – tyto se dále mohou dělit podle toho, zdali jsou zaměřeny obecně či specializovaně (*allgemeines einsprachiges/zweisprachiges Wörterbuch; einsprachiges/zweisprachiges Spezialwörterbuch*);²⁰ podle lexikografických kritérií pak na takové typy slovníků, které se soustředí na vymezenou jazykovou oblast, např. argot (Ibidem, s. 972). Ačkoli Hausmann hovoří obecně o problematice slovníků, platí tyto zásady také pro slovníky autorské, které mohou být koncipovány podle základních kritérií, jež Hausmann jmenuje. Vždy je ovšem nezbytné funkčně propojit několik hledisek; řečeno Hausmannovými slovy, je třeba se ptát, k čemu má slovník sloužit a jakým způsobem je vhodné jej sestavit. V případě autorských slovníků samozřejmě rovněž platí, že vedle

18 Tyto aspekty však z autorského slovníku ještě nečiní plnohodnotný nástroj pro literárněvědnou práci. Stále se totiž jedná (i v případě jmenovaných literárněvědných kategorií) o katalogizaci, jakkoli mohou být takové seznamy užitečné zejména z hlediska evidence. Funkčním nástrojem se stává pojetí slovníku, které je zpracované s ohledem na konkrétní analytické a metodologické použití, jež má svoji relevanci (oporu) právě v kontextu literárněvědného myšlení.

19 [„Otázka, k čemu má slovník sloužit, směřuje k volbě typu slovníku. Otázka, jak má slovník vypadat, směřuje k fenomenologické typologii slovníku.“]

20 Viz schéma in Hausmann 1989, s. 973.

uvedených požadavků je nezbytné kriticky reflektovat stav a situaci slovesného materiálu (viz pozn. 3).

Jestliže jsme dosud pojednávali o výkladovém slovníku, popř. thesauru jako o jednom z možných typů autorského slovníku, je v tuto chvíli třeba uvést i jiné příklady. Na poli autorských slovníků se lze tradičně nejčastěji setkat s tzv. glosářem, indexem nebo konkordančním slovníkem (viz Roelcke 1994, s. 2). S rozvojem a uplatněním strojového a korpusového zpracování jazyka vznikají také frekvenční či na korpusu založené slovníky. Označení těchto základních typů bývá mnohdy přímo součástí názvu. Christa Dill v článku *Lexika zu einzelnen Schriftstellern* (1959) rozlišuje kromě již zde jmenovaných tzv. registr (*Register*), vokabulář (*Vokabular*), semaziologický autorský slovník (*semaziologische Wörterbuch zu einzelnen Dichtern*) nebo slovník citací (*Zitaten-Lexikon*).²¹ Indexem Dill nazývá takový typ slovníku, který nabízí pouze základní informace o slovníkovém hesle spolu s odkazy k jednotlivým tvarům slova užitým v konkrétních případech. Současně zdůrazňuje zejména jeho praktické přednosti spočívající v rychlém vyhledávání (Dill 1959, s. 340). K typu tzv. indexu řadí Dill takový slovník, který zpracovává pouze výběrový materiál autorského lexika (Ibidem, s. 340). Z hlediska rozsahu excerpovaného materiálu a povahy lexikografického aparátu, který tento typ předpokládá, bychom mohli hovořit o redukované formě autorského slovníku. Podobně jako Wiegand (viz dále) nebo Mattausch nepovažuje ani Dill daný typ za ideální podobu autorského slovníku, ale chápe jej jako přechodový stav směrem k výkladovému typu (Ibidem, s. 340). Kromě typologie Dill rovněž hovoří o metodologii, v rámci které definuje sedm základních předpokladů pro relevantní sestavení autorského slovníku. V centru lexikografovy práce směřující ke konkrétní formě takového slovníku mají podle Dill figurovat následující zásady:

1. Cíl a účel autorských slovníků (*Ziel und Zweck der Dichter-Wörterbücher*).
2. Textové východisko (*Textgrundlage*).
3. Rozsah excerpované slovní zásoby (*Umfang des heranzuziehenden Wortbestandes*).
4. Metoda (*Arbeitstechnik*).
5. Způsob, jakým bude heslo zpracováno a v jakém rozsahu (*Artikelaufbau und -umfang*).
6. Způsob uspořádání hesel a lemmatizace (*Anordnung der Artikel und Lemmatisierung*).
7. Sestavení zvláštních slovníků, seznamů k jednotlivým dílům (*Sondernwörterbücher zu einzelnen Werken*).²²

21 Viz Dill 1959, s. 340–342.

22 Ibidem, s. 344.

Uvedené položky jsou stále aktuální a platné, a to i v případě, jedná-li se o moderní frekvenční autorský slovník sestavený pomocí počítačové technologie. Slovníky obecně na základě své typologie a určení disponují různou šíří lexikografických informací a jejich sestavení musí vždy předcházet řada úkonů, o kterých Dill pojednává. Tyto úkony se dále specifikují podle toho, o jaký druh slovníku se jedná. Zcela zvláštní požadavky v oblasti přípravy a realizace bude vyžadovat speciálně koncipovaný autorský slovník. Jedinečná kritéria však vždy vyrůstají na fundamentu primárních nároků, které stojí u základu každého slovníku.

Zajímavý pohled nejen na typologii autorských slovníků, jejich funkci a smysl, ale i významný přínos v oblasti metalexikografie nabízí Herbert Ernst Wiegand. Na rozdíl od Mattausche, který zevrubně pojednává zejména problematiku výkladových autorských slovníků, jejich typologii a charakteristiku, jakož i dílčí aspekty, vychází Wiegand důrazněji z epistemologických základů. To ovšem ještě nutně neznamená, že u Mattausche shledáváme jejich nedostatek. Jak jsme mohli sledovat v předchozím výkladu, Mattausch akcentuje spíše formální typologii autorských slovníků, zatímco Wiegand zdůrazňuje dvě svým způsobem odlišné koncepce, na kterých demonstuje „konflikt“ dvou základních metodologických bází.

Hlavní póly možné realizace autorských slovníků podle Wieganda představuje výkladový slovník (*Bedeutungswörterbuch*) a typ slovníku, který se soustředí na formální kritéria spojená s popisem jazyka (*Formwörterbuch*). Rámcově se tedy jedná o diferenci mezi hlediskem sémantickým a formálním, resp. mezi důrazem, jenž je v lexikografii kladen na významové aspekty lexikálních jednotek, s čímž souvisí určitý typ slovníku, a zejména přítomnost výkladového aparátu, a požadavkem tyto jednotky formálně kategorizovat. Ačkoli Wiegand ve své studii *Bedeutungswörterbücher oder sogenannte Indices in der Autorenlexikographie? Die Eröffnung einer Kontroverse* (1986) jednoznačně upozorňuje na význam výkladových slovníků na úkor druhého typu (*Formwörterbuch*), nemusí mezi oběma póly panovat přísná a zcela neprostupná hranice. Dichotomii mezi formou a významem nelze myslet izolovaně, neboť oba póly (nejen jazykového znaku) tvoří jednotu. Wiegand tuto skutečnost sice explicitně nepopírá, avšak svým opakovaným důrazem na význam výkladového slovníku poněkud upozaďuje druhý typ, který zejména dnes – kdy máme k dispozici nejen moderní nástroje kvantitativní analýzy jazyka, ale vyvíjí a prohlubuje se i poznání na poli teorie, která předkládá speciální a cíleně formulované matematicko-statistické vzorce a výpočty zaměřené na řadu kvantitativních, ale i kvalitativních aspektů jazyka či textu – má své nepopiratelné přednosti.²³ Wiegandovo stanovisko v oblasti tzv. autorské lexikografie, jemuž budeme nyní věnovat pozornost, je zajímavé

23 Zejména z důvodů metodologických je samozřejmě možné se zaměřit pouze na vybraný typ či model slovníku.

také z jiného úhlu pohledu. Wiegandovy názory, které mají v kontextu lexikografie zcela jednoznačně svoji váhu, jsou na druhé straně formulovány na jakémsi pomyslném časovém předělu, který představuje postupné zapojování prvních moderních jazykových korpusů, se kterými se vyvíjejí i nástroje přibližující (nejen) jazykové bádání objektivním a exaktním cílům, a zejména výsledkům.²⁴ Korpusové možnosti se dnes pochopitelně uplatnily i v případě autorských slovníků. Jejich potenciál se však zdá být mnohem širší a významnější než představa pouhé strojové, počítačové katalogizace.²⁵

Dříve nežli však o této problematice detailněji pojednáme, je nezbytné se seznámit s opačnými, popřípadě alternativními koncepcemi a postoji stejně jako s otázkou vývoje autorské lexikografie. Pro názornost dalšího výkladu je nezbytné na tomto místě prezentovat Wiegandovo schéma, které klasifikuje výše zmíněné obecné typy autorských slovníků (*Bedeutungswörterbuch* a *Formwörterbuch*) do přehledného a systematického zobrazení.

24 Srov.: „Možnost správy hromadných jazykových dat a práce s nimi na počítači vedou ve svých důsledcích nejen k nesmírnému zrychlení a usnadnění lingvistovy práce, ale i k jejímu dotud nevídanému zkvalitnění. Data takto získaná, která lze průběžně snadno modernizovat a doplňovat, tak představují dnes absolutně nejbohatší a nejrealističtější **zdroj poznání** jazyka vůbec. Možnostmi kvalitativními i kvantitativními tento zdroj vysoko předčí všechny lingvistovy pracně budované kartotéky a archivy minulosti a nabízí tak proti dosavadní a tradiční individuální lingvistově **introspekci** [zvýraznil R. Z.] a elicitaci informace od rodilých mluvčích. Proto také je jazykový korpus předpokladem ke skutečné revoluci v práci s jazykem a i proto se zcela právem o posledním desetiletí tohoto století mluví také jako o dekádě korpusové lingvistiky.

Mluví-li se v přírodních a dalších vědách zcela samozřejmě o **základním výzkumu**, na který navazuje veškerý další výzkum a aplikace, pak v jazykovědě se právě takovým zdrojem a předpokladem základního výzkumu stává elektronický korpus. Docenění jeho prvotního významu vede pak i k pochopení toho, že jeho budoucí studium a široké využívání znamená skutečně novou epochu v lingvistice, nový způsob výzkumu jak co do kvality a povahy dosahovaných výsledků (zahrnující do značné míry poprvé i ve vědě obecně nutnou možnost verifikace [...]), tak ovšem i podoby a povahy metodologie práce s ním; své důsledky bude mít i pro filozofii přístupu k jazyku vůbec.“ (Čermák 1995, s. 119–120)

25 Jak na jiném místě konstatuje František Čermák (2010, s. 244), „**frekvence** slova je inherentní vlastností každého lexému, a do ní se zrcadlí více dalších aspektů.“ Kvantitativní analýzu je třeba chápat jako komplementární součást analýzy kvalitativní.

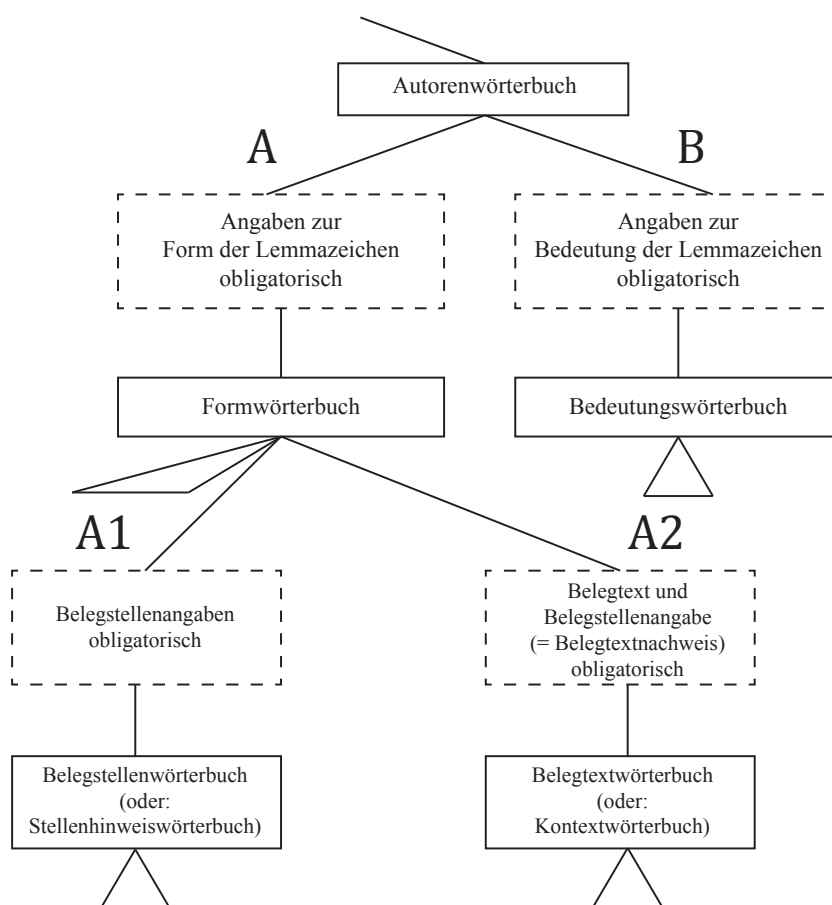


Schéma 3: Wiegandova klasifikace základních typů autorských slovníků založená na dichotomii mezi výkladovým typem a slovníkem, který je zaměřen na formální aspekty autorského slovníku. Schéma přebíráme ze studie H. E. Wieganda *Bedeutungswörterbücher oder so genannte Indices in der Autorenlexikographie? Die Eröffnung einer Kontroverse* (Wiegand 1986, s. 164).

(Pozn.: Velká písmena **A**, **B**, **A1**, **A2** do schématu doplnil autor této publikace. Původní Wiegandovo schéma tyto znaky neobsahuje.)

Ačkoli Wiegand konstatuje, že autorské slovníky náležejí obecně pod lexikografii, v rámci které tvoří specifickou oblast, současně uvádí, že je třeba rozlišovat mezi autorským slovníkem v pravém slova smyslu a typem slovníku tzv. indexu (*indices*) či konkordančním slovníkem (Wiegand 1986, s. 163). Rovněž nesouhlasí s tím, aby se typ slovníku, jež souhrnně označuje pojmem *Formwörterbuch*, nazýval autorským slovníkem, neboť jak index, tak například konkordanční autorský slovník, jenž do této kategorie náleží, naplňují podle Wieganda myšlenku autorského slovníku jen částečně, a proto takové jejich označení může vyvolat u uživatele mylná očekávání (Ibidem, s. 163–164). Co tím má Wiegand na mysli, vyplývá z jeho podrobnější charakteristiky obou základních typů. Pro oba je samozřejmě výchozí pozicí lemma (lexém). Slovníkové heslo (jak později doložíme na

konkrétních příkladech) ovšem může být strukturováno s ohledem na sémantiku, nebo naopak s důrazem na formální aspekty. Tato dichotomie stojí u základu dělení autorských slovníků na výše zmíněné dva typy. Dodejme, že také v případě slovníků primárně orientovaných na formální oblasti lexikálních jazykových jednotek (*Formwörterbücher*) nemusí docházet k eliminaci sémantického potenciálu; pouze je zde sémantika reflektována odlišným metodologickým způsobem, než s jakým se tradičně setkáváme u výkladových slovníků.

Jak je patrné z Wiegandova schématu (Schéma 3), typ *Formwörterbuch* se může realizovat v několika dalších verzích. Wiegand ve schématu demonstruje dvě možné varianty: první případ (námi označený jako **A1**) představuje slovník, jenž obsahuje výčtovou databázi obligatorních lemmat, zatímco v druhém případě (námi označený jako **A2**) se jedná kromě výčtu lemmat i o jejich kontextualizaci, jež je většinou realizována konkordancí.²⁶ V žádném z obou případů ovšem nedochází ke komentování významu či k zapojení jiného výkladového aparátu tak, jak je tomu u výkladového slovníku (*Bedeutungswörterbuch*). Wiegand typ slovníku, který se soustředí na formální rysy lexikálních jednotek, komentuje následovně: „Für alle Formwörterbücher gilt daher das Prinzip des semantischen Kommentars nicht; dies unterscheidet sie von den Autorenbedeutungswörterbüchern. Dies heißt zugleich, daß in allen Formwörterbüchern – sieht man von den Wörterbucheinleitungen und einzelnen Ausnahmefällen ab – die Sprache des Lexikographen als Kommentarsprache wörterbuchinternen Bedeutungsinterpretamente nicht gebraucht wird.“²⁷ (Ibidem, s. 165) Oproti tomu výkladový autorský slovník takové komentáře nezbytně obsahuje, což Wiegand pokládá za zcela zásadní znak autorského slovníku ve vlastním slova smyslu. Komentáře a informace o významové stránce lexikálních jednotek totiž nejenže vysvětlují způsob sémantického použití jednotky uvnitř díla, ale kromě toho pomáhají pochopit a ozřejmit starší vývojové fáze jazyka (Ibidem, s. 166).

Wiegand sice akceptuje skutečnost, že například zejména s rozvojem moderních technologií dochází k počítačovému zpracování autorského korpusu (Ibidem, s. 167), ale současně v této tendenci spatřuje úskalí, které je dáno absencí celé řady komentářů primárně se vztahujících k významové oblasti lexika. Výsledkem této selekce je podle Wieganda přílišná lexikografická nivelizace autorského jazyka způsobená jeho transpozicí na pouhé formální atributy. Z takto uspořádané lexikální zásoby se ztrácí především vlastní

26 Je pochopitelně možné hovořit o dalších typologicky spřízněných slovnících, např. frekvenčním slovníku, nebo slovníku založeném na korpusové bázi apod. Z hlediska Wiegandova členění náleží Čepův slovník do kategorie *Formwörterbuch*.

27 [„Pro všechny autorské slovníky akcentující formální princip uspořádání platí, že neobsahují komentáře významu; toto je odlišuje od výkladových autorských slovníků. To znamená, že ve všech slovnících tohoto typu – nehledě na jejich prolog a jednotlivé výjimky – se nevyskytuje jazyk lexikografů coby komentáře ve formě významových interpretamentů vycházejících ze slovníku.“]

sémantika, pragmatika, popř. širší historicko-kulturní významové specifikace. Nepřítomnost interpretační složky však pro Wieganda ještě nutně nemusí znamenat, že tyto typy slovníků nemohou být částečnou oporou pro lingvistiku a literární vědu.²⁸

Zhruba od 60. let 20. století dochází k systematickému aplikování matematických postupů v lingvistice a literární vědě.²⁹ S tím, jak se začíná uplatňovat moderní technologie (počítače) v humanitních oborech (zde v lingvistice), dochází postupně ve výzkumu k proměně nároků a cílů. Tento posun s sebou přináší i pozměněnou metodologii. V době, kdy Wiegand svůj článek publikoval, začínají se výrazněji prosazovat tendence zpracovávat autorský korpus pomocí strojové techniky, což Wiegand ve svém článku náležitě reflektuje a komentuje: „In der germanistischen Autorenlexikographie – eine Ausnahme macht nur die Goethe-Lexikographie – ist die Lage bei den Bedeutungswörterbüchern ziemlich miserabel. Bei den Formwörterbüchern dagegen hat es – dank des Computereinsatzes – große Fortschritte gegeben. Man gewinnt den Eindruck: die Erarbeitung von Bedeutungswörterbüchern ist eine **Last**, die computerunterstützte Zusammenstellung von Formwörterbüchern ist eine **Lust** für Germanisten.“³⁰ (Ibidem, s. 167–168) V závěru svého článku se Wiegand soustředí na povahu obou základních typů autorského slovníku. Je přesvědčen, že právě formalizované typy slovníků představují zvláštní a samostatnou větev v lexikografii (Ibidem, s. 168). Ačkoli jak Wiegand dále upozorňuje, není v zásadě principiálním odpůrcem počítačové lingvistiky, nepokládá ji za rovnocennou vůči tradičnímu způsobu lexikografické práce, jejímž výsledkem je výkladový autorský slovník.³¹ Takové slovníky, které se soustředí zejména na formální

28 „In allen Formwörterbüchern werden die Benutzungsmöglichkeiten viel zu pauschal angegeben und z. T. einfach behauptet. Es liegt mir durchaus fern, nun meinerseits zu behaupten, es gäbe keine sinnvollen sprach- und literaturwissenschaftlichen Benutzungsmöglichkeiten.“ (Wiegand 1986, s. 167)

[„Ve všech slovnících orientovaných na formální kritéria jsou jejich možnosti příliš paušalizovány a leckdy pouze jednoduše konstatovány. Je mi zcela cizí přesvědčení, že neexistují žádné smysluplné možnosti jejich užití v lingvistice a literární vědě.“]

29 Jedním z prvních výrazných příkladů užití matematických postupů v české lingvistice je společná práce Jaroslava Jelínka, Josefa V. Bečky a Marie Těšitelové *Frekvence slov, slovních druhů a tvarů v českém jazyce* (1961). Slovy Jaromíra Běliče se jedná o „první zachycení a zpracování frekvence slov, slovních druhů a tvarů v češtině, pro něž se započaly přípravné excerpcce materiálu již za války a jež bylo v podstatě ukončeno skoro před devíti lety“ (Bělič 1961, s. 5).

30 [„V německé lexikografii – výjimku tvoří pouze ta oblast lexikografie, která se zabývá Goethem – je situace kolem výkladových slovníků celkově žalostná. Naopak slovníky orientující se na formální aspekty autorského korpusu zaznamenávají díky počítačové technologii velký rozvoj. Z toho lze nabýt dojmu, že zpracování výkladových slovníků je **zátěž**, zatímco sestavení druhého typu slovníku pomocí počítače je pro germanisty **potěšením**.“]

31 Wiegand svůj pohled shrnuje do dvou základních bodů, jimiž předjímá své závěrečné jednoznačně kritické stanovisko vůči slovníkům založeným výhradně na zpracování formálních rysů (autorského) lexika: „(i) Ich bin durchaus der Auffassung, daß die Zusammenstellung eines Formwörterbuches eine selbständige wissenschaftliche Leistung ist.

(ii) Ich habe nichts gegen Computer, auch nichts gegen den Einsatz von Computern in der Lexikographie; im Gegenteil: ich sehe darin eine große Chance für die Sprachlexikographie und leite auf diesem Gebiet zusammen mit Peter Hellwig selbst ein Forschungsprojekt (COLEX).“ (Wiegand 1986, s. 168)

[„(i) Jsem proti názoru, že sestavení formalizovaných typů slovníků je samostatným vědeckým výkonem.

či kvantitativní výsledky a které jsou výstupem počítačově zpracovaného autorského korpusu, mají podle Wieganda sloužit pouze jako podpůrný nástroj, jenž pomáhá při sestavení výkladového slovníku, a nikoli představovat finální stádium lexikografické činnosti. Wiegand své stanovisko formuluje zcela jednoznačně: „Auf dem Wege zu Bedeutungswörterbüchern sind Formwörterbücher als Zwischenprodukte der Sprachlexikographie, als Stadien der Datenaufbereitung einzuordnen, und das gilt sogar für die Langue-Lexikographie, z. B. für Sprachstadien-Wörterbücher. Wird dieser Aspekt vergessen oder tritt er vollständig in den Hintergrund – und es gibt sehr deutliche Anzeichen dafür, daß dies mehr und mehr geschieht – dann tritt die Verslebständigung der Formwörterbücher ein. Lexikographische Zwischenprodukte werden zu Endprodukten umstilisiert. Es gilt das Motto: Sortieren statt kommentieren. [...] Computer eine benutzerabhängige Maschine ist, die im Wesentlichen Symbole segmentiert, sortiert und zählt, weiter nichts! Nein, am Computer liegt es nicht. Es liegt an den Interessen derer, die ihn einsetzen. [...] Formwörterbüchern sind keine Schlüssel zum Sinn, sie vermitteln keine Inhalte von der Vergangenheit in die Zukunft. Wenn sich ihre Herstellung verselbstständigt, werden die zentralen Aufgaben der Autorenlexikographie gerade verfehlt. Was als eine der Voraussetzungen für die Erhaltung der lexikalischen Verständigungsbasis und auch für die Sprachgeschichtsforschung gebraucht wird, sind Bedeutungswörterbücher. [...] Als Motto sollte gelten: Segmentieren, Numerieren und Sortieren vor allem, um zu Kommentieren!“³² (Ibidem, s. 168, 169) Příčinou zjevné kontroverze, kterou Wiegand staví mezi oba základní typy autorského slovníku a kterou mimo jiné demonstrativně vyjadřuje i formálně blízkými, avšak obsahově vzdálenými pojmy,³³ je především absence sémantického aspektu, který jako jediný je podle Wieganda schopen reflektovat

(ii) Nemám nic proti počítačům a také nic proti použití počítače v lexikografii; naopak v tom spatřuji velkou šanci pro lexikografii a v rámci této oblasti vedu spolu s Petrem Hellwigem výzkumný projekt (COLEX).“]

32 [„Cestou k výkladovým slovníkům jsou tyto slovníky (Formwörterbücher – pozn. R. Z.) meziprodukty lexikografie, představují stádium datové přípravy, což platí stejně tak i pro klasickou lexikografii, např. pro slovníky zpracovávající vývoj jazyka. Zapomene-li se na tento aspekt nebo ustoupí-li zcela do pozadí – a existují velmi zřetelné signály toho, že se to děje stále častěji –, dochází k osamostatňování slovníků ‚formálního‘ typu. Meziprodukty lexikografické práce jsou vydávány za výsledný produkt. Platí zde motto: třídění namísto komentování. (...) Počítač je stroj závislý na uživateli, který v podstatě segmentuje symboly, třídí a počítá, nic víc! Ne, v počítači to nespočívá. Spočívá to v zaměření těch, již ho využívají. (...) Slovníky orientované na formální popis lexika nejsou klíčem k významu, nezprostředkovávají žádný obsah jako kontakt minulosti s budoucností. Pokud dochází k osamostatnění jejich realizace, vede to k opomíjení klíčových úkolů autorské lexikografie. Výkladové slovníky jsou tím, co je potřebné jako předpoklad pro zachování základního porozumění lexiku a rovněž jeho vývoji. (...) Jako motto musí tedy platit: segmentovat, počítat a třídít především proto, aby jeho výsledkem bylo komentování.“] K tomuto Wiegandovu stanovisku je třeba poznamenat, že je formulováno – jak bylo uvedeno – koncem 80. a počátkem 90. let 20. století, kdy se formují i první elektronické jazykové korpusy. Nutno však podotknout, že v této fázi nebyly dostupné současné sofistikované analytické nástroje ve formě matematicko-statistických výpočtů a algoritmů, které poskytují takové kvantitativní výsledky, s nimiž lze relevantně pracovat i v rámci literárněvědného kontextu, aniž by se nutně jednalo pouze o „prosté“ počítání a segmentování autorského korpusu.

33 Jedná se o zvýrazněná slova v předchozí citaci, která tuto zjevnou povahu mají v němčině.

pouze výkladový typ slovníku. Právě tento rys je pro Wieganda určující, současně je také hlavním důvodem, který stojí v centru lexikografické práce. Své argumenty opírá Wiegand rovněž o povahu excerpovaného materiálu (zejména historických textů), který bez komentáře sémantiky lexikálních prvků ztrácí smysl.

Uváděná stanoviska jsou pro další výklad velice důležitá, neboť odrážejí dvojí základní metodologickou orientaci v otázce budování autorských slovníků, která současně představuje dvojí epistemologické hledisko v přístupu nejen k samotnému jazyku, ale *de facto* i k literárnímu textu/dílu, které je zpracováváno lexikografickou formou. První stanovisko, které zde bylo pojednáno, chápe sémantiku lexikálních prvků jako stěžejní úkol autorského slovníku. V tuto chvíli není důležitá otázka, zda by měl slovník disponovat metajazykovým aparátem soustředěným mimo jiné na významovou rovinu lexika, ale spíše to, jakým způsobem bude sémantika metodologicky pojednána. Pro pojetí významu v literárním uměleckém textu/díle se samozřejmě nabízí celá řada metodologických přístupů, z nichž některé mohou být dokonce vzájemně problematické (strukturálně-sémiotická analýza versus fenomenologie, archetypální kritika apod.). Na druhé straně, pokud Wiegand konstatuje, že výkladový slovník je schopen zaručit určitý transfer významu mezi minulostí a budoucností, je na místě se ptát, nakolik je tento transfer věcí uměleckého textu a nakolik odráží strategie rozumění, jež v dobovém a kulturním kontextu převládaly ve chvíli, kdy byl autorský slovník koncipován a realizován. Na straně druhé by výkladový slovník mohl fungovat jako zdroj poznání dobového způsobu rozumění a být takto využíván například v rámci kognitivních přístupů v literární vědě. Toto je ovšem již aspekt, kterým se dostáváme nad rámec Wiegandova uvažování a jeho zaměření. Wiegand naopak pokládá výkladovou podobu autorského slovníku, a zejména jeho epistemologickou bázi nikoli za produkt relativismu, ale za reprezentativní počín na poli lexikografické práce. Uvedené epistemologické hledisko vytváří dichotomii mezi dvojím základním typem autorského slovníku: je dialogem mezi hermeneutickou tendencí, jejímž produktem jsou výkladové typy slovníků, a strukturální orientací produkující formálně koncipované slovníky.

Tento dvojí aspekt je zcela zásadní. Podíváme-li se znovu na Wiegandovo schéma (Schéma 3), vidíme zde oblast, která předchází vlastní realizaci, tj. konkrétnímu typu autorského slovníku. Tato oblast je ve schématu vyznačena přerušovanou čarou (obdélníkem) a představuje záměr, koncepci, jež zakládá vlastní metodologické paradigma (**A** nebo **B**; **A1** nebo **A2**). První dva takové rámce představují nejen metodologický způsob, který směřuje k sestavení dvou základních typů autorského slovníku (*Formwörterbuch/Bedeutungswörterbuch*), ale reprezentují *de facto* epistemologickou volbu, jejímž instrumentem se posléze stává příslušná metodologie a výsledkem daný typ slovníku. Jestliže Wiegand chápe výkladový slovník jako konečnou fázi lexikografické práce

a slovník typu *Formwörterbuch* pouze za meziprodukty (*Zwischenprodukte*), potom z koncepčního hlediska, které akceptuje konkrétní badatelský záměr a přirozeně i příslušný cíl, se situace může jevit poněkud odlišně. Jinými slovy, například na počátku frekvenčního autorského slovníku stojí právě a jedině takové odborné, vědecké rozhodnutí, které usiluje interpretovat autorský korpus na základě možností, jež lexikografii poskytuje statistika a matematické zpracování jazykové materie (dnes přirozeně za pomoci počítačové technologie a korpusu). Ačkoli zpracování, jehož výsledkem jsou především formalizovaná data, stojí v opozici k výkladovému slovníku, přesto představuje vlastní – a v tomto má Wiegand zcela jistě pravdu – způsob, jehož hodnota vyplývá ze zvolené metodologie, která klade důraz na objektivizaci strukturních vztahů a poměrů v lexikonu. Současně to rovněž zcela nevylučuje sémantický rozměr, neboť také v tomto případě se pracuje se sémantikou lexikálních jednotek a zejména s jejich kontexty (kolokacemi). Je ovšem třeba si vždy uvědomit, co lze od takového přístupu očekávat, a zbytečně jeho relevanci nezkreslovat neadekvátními nároky. Cílem formální analýzy autorského korpusu je úsilí o co největší možnou míru objektivizace výsledků, s čímž logicky (v rámci metodologického postupu) souvisí i absence explicitního výkladového komentáře, tj. metajazykového interpretamentu implementovaného do slovníku. Sémantika je v případě formálních analýz, které dnes znamenají především lexikostatistický výzkum na bázi moderního jazykového korpusu, „derivována“ z imanentních strukturních zákonitostí jazykového systému, které jsou detekovány jeho kvantifikací.³⁴ Kvantitativní taxonomie, jež je přirozenou součástí moderní lexikografie, tedy nemusí nutně představovat pouze mezistupeň k výkladovému slovníku, ale může spolehlivě fungovat jako samostatný analytický proces, který ve výsledku pochopitelně vyžaduje náležitou interpretaci. Jak se pokusíme ukázat v případě Čepova korpusu a slovníku, svým způsobem nabízí uvedený postup finální materiál (ve smyslu formálního lexikografického zpracování autorského jazyka), který poskytuje specifické výstupní hodnoty. Předností takto uzpůsobeného materiálu je především přístup k jeho objektivní povaze a současně vysoký stupeň formálního zpracování, který vyplývá z aplikace příslušných matematicko-statistických měr. Již v tuto chvíli je tedy možné konstatovat, že naše stanovisko je poněkud odlišné od Wiegandova, čili že autorské slovníky, jež jsou zaměřené na formální analýzu jazyka (lexika), nepokládáme nutně za mezistupně, za jakési nehotové produkty, ale za účelové a funkční nástroje sloužící metodologicky vyhraněnému výzkumu.

Z hlediska typologie a vymezení autorských slovníků, jakož i z hlediska odborných nároků na lexikografickou práci je třeba Wiegandovy argumenty ovšem plně

34 V korpusové lingvistice tuto oblast představuje především zkoumání soustředěné na lexikální souvýskyty, kolokace, kontexty, frazeologii, idiomatiku apod.

respektovat. Názorový rozdíl, který vzniká mezi našima Wiegandovým stanoviskem ve věci tzv. *Formwörterbuch*, neznamená popření smyslu výkladových autorských slovníků. V názoru na oba základní typy autorských slovníků na rozdíl od Wieganda neakceptujeme jeho dichotomickou hierarchii. Na její místo dosazujeme funkční dichotomii vyplývající z metodologického a problémového zaměření konkrétní výzkumné činnosti.

V úvodu práce jsme konstatovali, že první autorské slovníky vznikaly z potřeby systematizovat lexikum významných autorů světové literatury. Německý lexikograf Thorsten Roelcke však poukazuje i na další důležité aspekty, které jsou s autorskou lexikografií úzce spojeny. Ve studii *Individual Sprache und Autorenlexikographie: Einführung* (1994) reflektuje metodologický vývoj týkající se především lingvistického uvažování o jazyce, jehož výsledkem je zejména funkčně stratifikační pojetí jazyka, které jej umožnilo chápat jako vnitřně diferencovanou komplexitu. Literární jazyk potom tvoří jednu z podstatných a současně relativně autonomních oblastí příslušného národního jazyka, jehož výzkum přispívá k poznání jeho jmenované komplexity. Jedním z prvních metodologických předpokladů, který podle Roelcka směřuje k formulování potřeby autorských slovníků ve 20. století, bylo pochopení literárního jazyka jako specifické autorské řeči, kterou je nezbytné zkoumat odděleně od jiných jazykových projevů (Roelcke 1994, s. 3). V další fázi se tato oblast specifikuje s ohledem na její formální kritéria, čili za autorskou řeč se považuje takový jazykový soubor (korpus), který je definován především písemným projevem autora (Ibidem, s. 4) a následně reflektován jako imanentní jazykový, resp. sémiotický systém (Ibidem, s. 6). Zjištění, že jazyk je vnitřně diferencován do jednotlivých funkčně-sémantických oblastí, které vykazují jistou systémovou autonomii, mělo za následek to, že se pozornost v oblasti lexikografie začala věnovat rovněž autorskému lexiku jako zvláštní součásti obecného jazykového systému. Výsledkem je konstatování imanence nejen celkového jazykového systému, ale i jeho jednotlivých částí, mezi které náleží oblasti idiolektu nebo individuálního řečového jednání, jež jsou součástí komplexního systémového celku.³⁵ Tento posun v uvažování představuje autorskou řeč jako ucelený a relativně ohraničený jazykový systém, který jako takový lze podrobovat

35 Jako jeden z dokladů svého tvrzení o směřování lingvistického uvažování k pojetí jazyka jako stratifikovaného systému Roelcke cituje část studie Alfreda Gerstenkorna (1971, s. 155) *Sprache – Kompetenz – Lekt: Vorschläge zur Terminologie*, která vyšla v časopise *Zeitschrift für deutsche Sprache* 27: „Jedes sprachliche System – Gemeinsprache, Einzelsprache mit mehreren Dialekten, Dialekt usw. – heiße >Lekt<. Ein Lekt sei der allgemeinste oder Oberbegriff für alle sprachlichen Systeme [...]. Idiolekte sind keine Teilmengen, sondern Lektelemente, aus denen sich jeder Lekt zusammensetzt.“ Hier werden Idiolekte als die kleinsten sprachlichen Systeme aufgefaßt, welche die Elemente größerer sprachlicher Systeme, d.h. Einzelsprachen und deren Varietäten, bilden.“ (Roelcke 1994, s. 6)

[„Každý jazykový systém – obecný jazyk, konkrétní jazyk s vícero dialektů, dialekt atd. – se nazývá >lekt<. Lekt je nejobecnější pojem či zastřešující pojem pro všechny jazykové systémy (...). Idiolekt není podmnožinou, ale částí lektu, ze kterého se každý lekt skládá.“ Zde jsou idiolekty pojímány jako nejmenší jazykové systémy, které tvoří součást větších jazykových systémů, tzn. konkrétního jazyka a jeho variant.“]

vlastní analýze a zkoumání, jež v první řadě přináší výsledky o daném systému či útvaru národního jazyka (Ibidem, s. 8). Důležité je především to, že autorský řečový projev nelze nikterak redukovat. Jestliže tedy zpočátku autorské slovníky sloužily jako zdroj poznání literárního jazyka, který byl chápán jako reprezentativní ukázka (vzorek) národního jazyka a jeho vyspělosti, tentokrát mají sloužit především pro poznání jedné z důležitých (a relativně dobře mapovatelných) systémových oblastí národního jazyka. Tato skutečnost rovněž dovoluje další specifikaci autorské lexikografie, jež reflektuje systémovou „uzavřenost“ autorského (literárního) jazyka, který se tak jeví jako „das Sprachsystem, insbesondere das System des Wortschatzes eines Autors, welches durch Selektion, Motivation sowie Historizität ausgezeichnet ist“³⁶ (Ibidem, s. 8). Obecně pojetí systému v lexikální rovině předpokládá vedle sémantiky i formální kritéria (slovotvorná), popř. kritéria pragmatická, etymologická a jiná. Autorský slovník je především díky své konkrétní koncepci schopen poskytnout přehled o uspořádání a vztazích v rámci vymezeného lexikonu, a tak umožnit pohled na jedinečné *langue*, jež formuje a utváří národní jazyk jako jeho součást. Autorský lexikon se tak stává podstatným tématem, lexikografická práce jej zpřístupňuje v systematické a strukturované podobě a umožňuje jeho další zkoumání.

S ohledem na vývoj kritického uvažování v oblasti metalexikografie Roelcke své definice dále zpřesňuje, což se projevuje v sebereflexivním zdůraznění vlastního metalexikografického diskurzu. Zcela konkrétně to znamená, že autorský slovník coby produkt lexikografické práce nelze chápat jako vlastní projev autorské řeči (není s ní identický), ale jako způsob její reprezentace a svého druhu interpretace; autorský slovník není verbálním projevem autora, ale jeho obrazem *sub specie* dané disciplíny a její metodologie. Navíc jak Roelcke uvádí, každý takový obraz je determinován i pragmatickými aspekty, zejména ohledem na cílovou skupinu uživatelů slovníku.³⁷ A konečně autorský slovník podle Roelcka nelze chápat jako náhradu samostatného interpretačního výkonu, ale jako pomocný nástroj této činnosti (Ibidem, s. 14).

36 [„jazykový systém, obzvláště lexikální systém autora, který charakterizuje selekce, motivace a také historicitá“]

37 Jelikož se jedná o finální definici, která završuje Roelckův přehled o vývoji metalexikografie v souvislosti s lingvistickými koncepcemi rámcových oblastí národního jazyka, jež mají svoji relevanci pro autorskou lexikografii, uvádíme ji zde v úplném znění: „Eine Individualsprache ist (aus Sicht der Wörterbuchbenutzung im Rahmen einer metalexikographischen Konzeption von selektiven Autoreninterpretationswörterbüchern) derjenige Teil der sprachlichen Produktionskompetenz eines Autors, welcher als Sprachsystem, insbesondere als Wortschatzsystem rekonstruiert wird, um so die sprachliche Rezeptionskompetenz eines Lesers bzw. Wörterbuchbenutzers hinsichtlich der Selektion, Motivation und Historizität dieses System zu ergänzen.“ (Roelcke 1994, s. 15)

[„Individuální řečový projev je (z hlediska užití slovníku v rámci metalexikografické koncepce výběrových interpretačních autorských slovníků) ta z hlediska díla relevantní část jazykového projevu autora, která je rekonstruována jako jazykový systém, zejména jako systém slovní zásoby tak, aby doplňoval jazykovou recepční kompetenci čtenáře či uživatele slovníku s ohledem na selekci, motivaci a historicitu tohoto systému.“]

Další důležitá rovina, která se promítá do úvah kolem autorské řeči a která se zde již v rámci citací objevila, se týká toho, co Roelcke pojmenovává jako selekci, motivaci a historicitu. Dosavadní oblast zkoumání individuálních (autorských) řečových projevů je tímto vedle významu rozšířena o další specifikující určení, která dovolují interpretovat sledovaný materiál z hlediska výběru, způsobu tvoření a historické kontextualizace jazykových jednotek, které jsou v rámci dané oblasti, tj. jazykové materie autorské řeči, významotvorné (Ibidem, s. 13). Selekcce, motivace a historicitu náleží dle Roelcka mezi základní kritéria lexikografické interpretace.³⁸ Současně jsou to právě tyto základní znaky, které definují literární jazyk jako do značné míry autonomní systém.

Roelcke ve své studii uvádí celkem osm (opěrných) definic, které odrážejí proměnu v přístupu k jedinečným řečovým projevům (autorské řeči). Tento vývoj Roelcke zaznamenává na pozadí úvah o národním jazyce jako obecném celku a jeho modifikacích, což se týká zejména idiolektů. V určité fázi úvah nad problematikou autorské řeči se mělo za to, že tato tvoří součást širšího idiolektu národního jazyka (Ibidem, s. 4). Teprve postupně docházelo k chápání autorské řeči jako samotné oblasti, které začala být věnována systematická pozornost, s čímž úzce souvisí i problematika autorské lexikografie. To ovšem ještě neznamená, že samotná teoretická reflexe představuje výslednou fázi, na jejímž konci stojí autorské slovníky jako bezproblémový finální produkt. Chápání národního jazyka jako strukturně a funkčně rozvrstvené materie, která má svá specifika mimo jiné rovněž v oblasti pragmatiky, sice představuje zásadní moment na cestě k moderně pojatým autorským slovníkům, avšak sám o sobě tento fakt nestačí k úplnému zvládnutí problémů, se kterými se lze v případě koncipování autorského slovníku setkat. Především je zde nezbytná důkladná znalost celého excerpovaného materiálu (textologické hledisko), včetně celé řady dalších informací týkajících se například autorovy biografie, historického kontextu (kulturního, politického, ekonomického a jiného) a samozřejmě dobrá obeznámenost s vývojovými fázemi jazyka a literatury ve všech jejich aspektech. Neméně důležité jsou i doplňující zdroje, které mohou být opět různé povahy, od spisovatelovy korespondence či deníků po dobové ohlasy v tisku. Dříve nežli na některé z těchto aspektů detailněji poukážeme, konstatujme, že situace, o které jsme právě pojednali, posunuje pojetí autorských slovníků od prostých manuálů, resp. kánonů sloužících za vzor správného a reprezentativního stylu a jazyka obecně směrem ke zdrojům hlubšího a intenzivnějšího poznání funkčně-sémanticky specifikované jazykové oblasti. Úzce s tím souvisí také aktuální snaha, aby tato oblast byla popisována co možná

38 Setkáme se s nimi mimo jiné i ve studii J. Mattausche *Textlexikographische Aspekte im Autorenwörterbuch (am Beispiel des Goethe-Wörterbuchs)* – viz dále.

nejobjektivněji a nejexplicitněji, tzn. na základě relevantních informací nezkrslujících interpretaci materiálu.

Snaha reflektovat literární jazyk jako svébytný jazykový prostor sice znamená významný a důležitý posun v autorské lexikografii, který se odráží v mnoha konceptech těchto prací, avšak stále ještě nelze hovořit o autorském slovníku, který by byl zaměřen primárně na literárněvědnou sféru. V popředí jsou jednoznačně lingvistické zájmy, což signalizuje i pojem *autorský jazyk*, který zastupuje dnes relevantnější označení *jazyk literárního textu/díla*.³⁹

Zajímavý příklad budování velkého autorského slovníku uvádí a komentuje ve své studii *Individualsprache und Gemeinsprache Bemerkungen zum Goethe-Wörterbuch* (1986) Horst Umbach, který hovoří o koncepci Goethova slovníku, jejímž autorem je Wolfgang Schadewaldt. Umbach zmiňuje Schadewaldtův postup, jenž zprvu spočíval ve výběru základních tematických okruhů Goethova díla: „Als die erste und wichtigste Aufgabe des Goethe-Wörterbuchs sah er [Schadewaldt – pozn. R. Z.] die Darstellung der von ihm so benannten ‚Grund- und Wesenwörter‘. Dem lag die Auffassung zugrunde (und hier wird das klassizistische Konzept deutlich und damit auch die Differenz zu der vorherrschenden historisch-genetischen Methode der Germanisten), daß es bei Goethe zwar Wandlungen gegeben habe, aber daß er ‚doch im Grunde immer der gleiche geblieben‘ sei und daß gewisse Grundelemente des Fühlens, Erlebens, Sehens, Denkens und Gestaltens als Konstanten auch in diesen ‚Grund- und Wesenwörtern‘ greifbar würden. Dazu zählte er etwa: *Gott, Liebe, Leben, Geist, Natur; Genießen, Gegenwart, Augenblick, Gesundheit*.

Hier war, das macht die Reihe wohl deutlich, mehr im Spiel als nur Goethesemantik. Wenn man, um nur das letzte Wort aufzugreifen, 1946 *Gesundheit* als einen Goetheschen Wesensbegriff konstatovala, dann ließ sich die zentrale, den physischen Bereich weit übergreifende Bedeutung des Wortes bei Goethe gewiß belegen, aber ebenso sicher hatte die Hervorhebung einen aktuellen Bezug. Der wird wohl auch deutlich in der nächsten von Schadewaldt aufgeführten Wortgruppe: *Schuld, Unruhe, Vernichtung, Hoffnung, Wiederherstellung*.“⁴⁰ (Umbach 1986, s. 163) Umbach ukazuje, jaké komplikace při

39 Označení autorský jazyk chápe jazyk literárního díla za verbální projev autora. Dnes jsou takovému pojetí blízké některé přístupy v teorii literatury, které v určitých případech (er-forma) odmítají pojem vyprávěče jako kvality odlišné od autora; na místo dichotomie autor – vyprávěč nastupuje pojetí autor – autor s ohledem na zvláštní řečovou aktivitu, jež se projevuje jako vyprávění v er-formě (viz Walsch 2007) [Ke kritice Walshovy koncepce viz Kubíček 2007].

40 [„Za první a skutečný úkol pokládal (Schadewaldt – pozn. R. Z.) představení základní slovní zásoby. Zastával názor (a zde se zřetelně ukazuje klasicistický koncept a tím i jeho odlišnost od historicko-genetické metody germanistů), že Goethe sice procházel proměnami, ale ‚přesto v zásadě zůstával stejný‘, a je tedy možné vystopovat určité základní prvky citění, prožívání, vidění, myšlení a tvoření, které tvoří konstanty také v těchto ‚hlavních a základních slovech‘. K těm počítal např. slova *bůh, láska, život, duch, příroda, požitky, přítomnost, okamžik, zdraví*.

Tato řada zřetelně ukazuje, že zde bylo ve hře cosi víc než jen sémantika. Pokud bylo jako jeden ze základních pojmů Goethova slovníku v roce 1946 uvedeno slovo *zdraví* – v případě posledního slova z řady –, potom byl sice doložen hlav-

abstraktním (tematickým) zobecnění autorského korpusu, jež má vést k reprezentativnosti autorského jazyka, dokážou způsobit například dobový kontext, předchozí interpretační tradice nebo subjektivní činitele, které determinují lexikografovo uvažování. V tomto případě se jedná o výběr tematických nebo klíčových slov, která dle Schadewaldta reprezentují centrum lexikálně-sémantické vrstvy Goethova jazyka. Nicméně ve skutečnosti – jakkoli bychom měli v úmyslu zpochybňovat jejich význam v Goethově lexikonu – se jedná o apriorní selekci, která poněkud zastírá a znemožňuje nezaújatý přístup k autorovu lexiku. Umbach se ve své práci *de facto* dotýká podobných problémů, o kterých hovoří Mattausch v souvislosti s vymežováním a určováním specifických oblastí jazykových projevů (*Individualsprache*), konkrétně tzv. autorské řeči. Důležitým kritériem se tu stává především metodologická vyhraněnost pracovního postupu, která se snaží zabránit nechtěným subjektivním intervencím interpreta-lexikografa. Jak je vidět, rozhodující úloha nastává nejen při samotném vytváření komentářů slovníkových hesel (interpretamentů), ale již v základní koncepci, rozvržení a plánu vlastního slovníku, jehož primárním kritériem musí být především objektivita při snaze zaznamenat jazykový stav a situaci, což platí u všech typů slovníku a na všech jeho úrovních. Z tohoto hlediska je tedy nutné důsledně odlišit práci lexikografa od literárního interpreta. To však ještě neznamená, že jsou tyto pozice přísně kontradiktorní. Ve výsledku ovšem musí být respektována metodologická i epistemologická autonomie každé z nich. Totéž platí i v případě autorských slovníků a způsobu, jakým jsou koncipovány a sestaveny. Ať již je metodologická volba, jež stojí u základu slovníku, jakákoli, nelze ji, resp. samotný slovník chápat jako substituci literárněvědného analytického či interpretačního výkonu.

Jak již bylo konstatováno, autorské slovníky sloužily mimo jiné k poznání vývojových stadií daného jazyka, neboť literární díla díky své verbální fixaci představovala spolehlivý zdroj excerpcí. Tento názor ovšem nelze aplikovat bez omezení a nekriticky, neboť v případě autorského lexika se jedná o zvláštní segment národního jazyka, který tento nereflektuje v dostatečné úplnosti a reprezentativnosti. Zejména poslední pojem (reprezentativnost) má v uvedeném kontextu zvláštní význam. Tvůrci starších autorských slovníků z 19. století chápali jazyk (významného) spisovatele jako natolik reprezentativní, že bylo vhodné, ba dokonce praktické jej systematicky poznávat a zpracovávat, neboť byl vnímán jako obraz kultivovanosti a vyspělosti národního jazyka.⁴¹ Pojem reprezentativnosti

ní, fyzickou oblast dalece převyšující význam, ale tato volba má rovněž svůj aktuální zřetel. Ten je také zřejmý v dalším Schadewaldtově výběru slovní skupiny: *vina, neklid, záhuba, naděje, obnova.*"]

41 V této souvislosti stojí za připomenutí, že v české jazykovědě se v první polovině 20. století objevila koncepce tzv. dobrého autora. Jejím smyslem bylo vytvořit dostatečně reprezentativní obraz národního jazyka, který by se zakládal na vyspělé literární kultuře.

Podobná situace měla platit i v souvislosti s Goethovým slovníkem, jenž neměl být pouze slovníkem jednoho autora, ale reprezentantem kultivované němčiny a celé jedné vývojové etapy jazyka, jak konstatuje Umbach (Umbach 1986, s. 164).

se ovšem proměnil spolu s vývojem myšlení o jazyce, zejména v souvislosti s chápáním jazyka jako strukturně stratifikovaného systému. Reprezentativnost se přesunula z roviny *oblasti* na *materii*, tzn. na samotný jazyk s důrazem na jeho funkční vyváženost,⁴² přičemž *oblast* dále zůstává důležitým činitelem, ovšem již nikoli dominantním.⁴³ *Oblast* užití nahradil pojem *systém*, čímž se proměnil i význam pojmu reprezentativnost.

O tom, že literární jazyk nemůže být spolehlivým (reprezentativním) obrazem celého národního jazyka, podává jasný doklad již zmíněná studie H. Umbacha, který jako jednu ze zvláštností při koncipování Goethova slovníku jmenuje rozdíl mezi juvenilními a pozdějšími texty. Podobně Umbach hovoří o vlivu různých oblastí zájmu, kterým se Goethe v průběhu života věnoval (např. optika, botanika). Nemalý problém podle Umbacha způsobovalo Goethovo zvláštní zacházení se slovy spočívající v slovní hře (Ibidem, s. 172), kompozitech (Ibidem, s. 168), individuálních pojmenováních apod., které jsou jedinečnou a nedílnou součástí Goethova jazyka a jeho systému. V neposlední řadě je nezbytné si uvědomit, že literární jazyk disponuje specifickými výrazovými prostředky a lexikem, jež náleží právě do oblasti literárního jazyka. Z hlediska reprezentativnosti jako pojmu chápaného systémově je tento jazyk reprezentantem (indexem) své třídy; s odkazem na Jurije M. Lotmana je reprezentantem sekundárního modelujícího systému, a v konečném důsledku samozřejmě i celku národního jazyka s tím, že tento na své rovině disponuje vlastními imanentními pravidly a strukturací. Tak jako nelze na základě jednoho či několika autorů vystavět dostatečně plastický obraz národního jazyka, nelze tak spolehlivě učinit ani v případě literárního jazyka, který rovněž disponuje množstvím odlišných situací a oblastí.⁴⁴ Ne vždy musí platit, že literární jazyk je natolik svébytný, že se zcela evidentně odlišuje od běžné pragmaticko-komunikační formy jazyka, jako je tomu zejména u poezie či avantgardních próz v obecném slova smyslu. Naopak u realistických próz se výrazové prostředky stejně jako jejich soubor (text) mohou jevit takřka identickými s běžným jazykovým projevem. Nicméně i v tomto případě platí, že literární jazyk je vždy systémem prostředků a vztahů *sui generis*, a jako takové je třeba je zkoumat. Jednoznačně to dokládá i strukturní charakter korpusu Čepových próz, který vykazuje jiné

42 Zejména v dnešní době, kdy jsou běžně k dispozici moderní jazykové korpusy, se k uvedenému požadavku připojuje také důraz na dostatečně velký vzorek jazykového materiálu.

43 V současnosti je toho dokladem budování rozsáhlých korpusů národních jazyků, které zahrnují texty z různých oblastí jazyka, tedy nikoli pouze texty umělecké literatury, ale veškeré žánrově dostupné texty. Navíc, jak ukazuje i situace Českého národního korpusu (ČNK), excerpují se nejen fixované texty, ale budují se i korpusy mluveného jazyka, např. Pražský mluvený korpus, Brněnský mluvený korpus nebo Olomoucký mluvený korpus. Zvláštní kapitolou jsou diachronní korpusy, v případě ČNK se jedná o projekt Diakorpus, nebo srovnávací korpusy, např. InterCorp a jiné (podrobněji viz WWW: <<http://ucnk.ff.cuni.cz/struktura.php>>).

44 Relativně ideální situací by v tomto případě bylo vybudovat korpus na bázi dostupných literárních textů určitého období, nejlépe pak několik takových korpusů, které by bylo možné vzájemně konfrontovat. Současné možnosti jazykových korpusů i analytických nástrojů vycházejí této myšlenky vstříc.

strukturní vztahy a zákonitosti, než je tomu u reprezentativního lexikonu obecného jazyka.⁴⁵ V žádném případě tedy nelze úspěšně redukovat obraz národního jazyka na jazyk literárních děl, jakkoli může platit, že právě v těchto textech dosahuje jazyková úroveň vysoké estetické hodnoty.

V souvislosti s přípravou Goethova slovníku hovoří o problematice autorského jazyka na konci 50. let 20. století také Manfred Gräfe, který výše uvedené skutečnosti reflektuje ve studii *Goethe und die Probleme der neueren deutschen Lexikographie* (1959). Gräfe zmiňuje především to, že Goethův jazyk nelze bezvýhradně pokládat za reprezentativní obraz němčiny 18. století, neboť se jedná o specifické autorovo zacházení s jazykovým materiálem, stejně jako nelze nekriticky interpretovat významy některých básnickových slov a slovních spojení prostou sémantickou analogií k současnému jazykovému úzu (Gräfe 1959, s. 304n). Takový kritický přístup, jenž respektuje stratifikační rámce národního jazyka, dokáže nejen funkčně vydělit a klasifikovat jeho jednotlivé oblasti, ale především přispět k vyšší míře porozumění složitému jazykovému systému, a to nejen z hlediska celku národního jazyka v rámci určitého vývojového stadia, ale také s ohledem na vývojové zákonitosti jeho autonomních oblastí, v tomto případě literárního jazyka.

Do teoretických debat nejen o reprezentativnosti slovníků výrazně přispěl rozvoj moderní technologie, zejména možnost počítačově zpracovávat a analyzovat přirozený jazyk. Mění se tím nejen způsob vlastní práce, zejména excerpce jazykových dat, která již nemusí být závislá na manuální katalogizaci, ale i důsledky, které korpusově orientovaná lingvistika nabízí pro interpretaci jazykové materie a jejího systému. Základní přínos korpusové lingvistiky pro oblast lexikografie spočívá především v možnosti shromažďovat a pořádat velké množství jazykových dat a současně využívat statistických algoritmů při jejich analýze. Cesta k počítačově zpracovaným korpusům znamená nejen usnadnění práce, ale představuje zejména důležitou cestu k objektivizaci získaných a analyzovaných hodnot. Lexikostatistika je propojením specifické metodologie aplikované na strukturní systém jazyka. V rámci takové analýzy lze v první řadě objektivně detekovat obecná strukturní pravidla a zákonitosti vybraných jazykových celků či segmentů, které podléhají matematicko-statistické interpretaci. Konkrétně lze například sledovat, jakým způsobem jsou určité výrazy v korpusu kvantitativně, ale i kvalitativně distribuovány, v jakých kontextových pozicích (lexikální kontexty, větné či textové kontexty, nebo dokonce kontexty ve speciálně vybraných korpusech) nejčastěji figurují a v jakých se naopak ocitají na periférii apod. Pro autorskou lexikografii znamená využití počítačové lingvistiky především možnost pracovat se statistickými hodnotami, které se dnes nabízejí širokému využití, jež nemusí výlučně směřovat do lingvistiky. Týká se to jednoduchých frekvenčních

45 Tuto skutečnost reflektujeme ve třetím oddíle na příkladu frekvenční distribuce tematických slov a jejich průběhů.

statistik lexikálních (nebo jiných) jednotek přes výpočty relativních či kumulativních frekvencí až po složitější definice pro výpočet bohatství slovníku, jeho tematické oblasti, deskriptivity či aktivity textu⁴⁶ nebo pro analýzu kontextových vztahů.⁴⁷

Jednou z nesporných předností korpusového přístupu je možnost vytvářet rozsáhlé a různě strukturované korpusové databáze, které po dalším zpracování poskytují bohatý materiál informací a údajů. U nás je takovou největší databází Český národní korpus (ČNK), který spravuje Ústav Českého národního korpusu v Praze. ČNK je vnitřně strukturován například z hlediska žánrové příslušnosti textů, formy (mluvenost/psanost) či jejich historicity atd.⁴⁸ Lze říci, že čím bohatší bude tato vnitřní struktura národního korpusu a čím rozsáhlejší bude jeho materiál, tím reprezentativnější bude samotný korpus, přičemž přesnější a objektivnější bude i pohled na národní jazyk. V tom spočívá další z nesporných výhod elektronického korpusu: v možnosti jeho neustálého rozšiřování a flexibility vzhledem k neustále se vyvíjející a proměňující se jazykové situaci.

Výsledky korpusové a kvantitativní lingvistiky se rovněž nabízejí jako vhodná platforma pro zpracování autorských slovníků, a to především na frekvenční bázi. Velkou výhodou je kromě jmenované přesnosti i možnost praktického srovnávání různě definovaných množin jazykových dat, např. mezi různými autorskými slovníky nebo mezi autorskými slovníky a národním korpusem, nebo schopnost aplikovat lexikostatistickou metodologii a její specifické vlastnosti tam, kde se tradičně používají jiné metody. Posledně jmenovaná problematika tvoří teoretický rámec této práce, v jehož kontextu hodláme reflektovat vztah mezi kvantitativně-korpusovými postupy a strukturálně-sémiotickými pozicemi v literární vědě.⁴⁹

Výběrově zaměřený výklad o autorské metalexikografii může být dále rozšiřován.⁵⁰ Abychom nezůstali pouze na rovině teoretického uvažování, v následující části uvedeme některé příklady autorských slovníků, které demonstrují podoby, v jakých mohou autorské slovníky existovat. I zde však platí, že výčet, který bude následovat, nemůže být ani zdaleka úplný. Soustředili jsme se pouze na některé významnější práce s tím, že jsme se pokusili ukázat především jejich rozdílnou typologii.

46 Viz např. Čech – Popescu – Altmann 2014.

47 Viz Cvrček 2013.

48 Srov. stať M. Šulce *Tematická reprezentativnost korpusů* (2001).

49 Této problematice je věnován druhý a třetí oddíl.

50 Obecně o typologii slovníků (v tomto případě ovšem nikoli výhradně autorských) pojednávají lexikografické práce, u nás mezi mnohými např. publikace *Manuál lexikografie* (1995). Problematikou lexikografie obecně se v této práci nezabýváme.

1.3 Vybrané příklady autorských slovníků (stručný historický přehled)

Výchozí bázi každého autorského slovníku je jazyková oblast vymezená rozsahem literární tvorby daného autora. Jak konstatuje Josef Mattausch ve studii *Kommunikativ-pragmatische Sprachbetrachtung und Lexikographie: Zur Spezifik des Autorenwörterbuchs* (1982), v případě autorských slovníků nespočívá rozhodující kritérium v poznání jazykového systému jako takového, ale v pochopení a zmapování zvláštní jazykové situace uvnitř oblasti vymezené literární tvorbou daného autora.⁵¹ Metodologické hledisko, které lze uplatnit v souvislosti se zpracováním autorského slovníku, závisí na několika oblastech, které se ovšem odvíjejí od následujících základních skutečností:

- a) První kritérium vychází z **povahy materiálu**, kterým je literární jazyk. Rámcově lze v této souvislosti konstatovat, že autorský slovník může respektovat například *historicitu materiálu (diachronii) nebo jeho strukturní aspekty (synchronii)*, popř. obojí.
- b) Druhé kritérium je založeno na vytčeném **cíli a záměru**, čemu (a komu) je autorský slovník určen. Lze uvažovat obecné autorské slovníky, tj. takové, které k jednotlivým heslům podávají výklad a současně uvádějí zdrojový výskyt hesla. Na druhé straně je možné koncipovat specializovaný autorský slovník, který bude zaměřen na konkrétní (dílčí) oblast (např. vymezená tvorba) a adresován především užší zájmové skupině (lingvistům, literárním vědcům). To předpokládá i způsob zpracování slovníku.
- c) Poslední kritérium určuje **metoda lexikografické práce**. V současnosti sem náleží především využití moderních postupů a technologií, které rozšiřují a specifikují pohled na jazyk. Na mysl máme především strojově (počítačem) zpracovaná data. V určité opozici potom stojí výkladový typ slovníku, který samozřejmě může využít výsledků kvantitativní

51 „Das Autorenwörterbuch gehört zur Kategorie des *speziellen, beschränkten* Wörterbuchs (nach Zgusta ‚restricted dictionary‘ – im Gegensatz zum Typ des ‚general dictionary‘, dessen Objektbereich die allgemeine Nationalsprache oder auch die nationale Literatur- und Standardsprache ist.).

Sein Gegenstand ist die Sprache eines bestimmten (wesentlichen, repräsentativen) Autors, dokumentiert in seiner schriftlichen Hinterlassenschaft, seinem Gesamtschaffen oder einzelnen Werken und Werkgruppen.“ (Mattausch 1982, s. 305)

[„Autorské slovníky náležejí do kategorie speciálních, ohraničených slovníků (podle Zgusty ‚restricted dictionary‘ – na rozdíl od typu ‚general dictionary‘, jehož předmětem zájmu je obecně národní jazyk nebo také jazyk národní literatury a jazykový standard).

Jeich zájmovou oblastí je řeč určitých (známých, reprezentativních) autorů dokumentovaná v jejich písemné pozůstatosti, souborném dílu nebo v jednotlivých pracích a souborech děl.“]

Z těchto důvodů má obzvláštní důležitost sémantika. Jak ostatně dále praví Mattausch: „Wörterbücher stehen und fallen mit der Qualität ihrer Bedeutungserklärungen.“ (Ibidem, s. 307)

[„Slovníky stojí a padají s kvalitou svého významového aparátu.“].

a korpusové analýzy (např. kontextové analýzy), avšak jeho těžištěm je především výkladový aparát.

Mattausch mezi možnými uplatněními autorského slovníku uvádí například jeho využití pro poznání vývoje básnického jazyka (Mattausch 1990, s. 1553), který lze zkoumat v diachronním i synchronním kontextu. To ovšem již předpokládá určitý způsob, kterým bude autorský slovník zpracován, resp. důraz buď na formální, či na významovou stránku jazyka a zaměření slovníku. Jedná se *de facto* o výsledek řady rozhodnutí, jejichž základem jsou zde výše uvedené předpoklady, které určují koncepci autorského slovníku.

Podle Thorstena Roelcka je možné při zpracování autorského slovníku postupovat dvěma základními způsoby. První zakládá *de facto* koncepci thesauru, která spočívá na co možná úplné excerpci a interpretaci autorského jazyka. Druhým, který je pro Roelcka „schůdnější“, je výběr konkrétního problému, jemuž je podřízena výsledná (redukováná) podoba autorského slovníku (Roelcke 1994, s. 1–2). Následující výklad bude zaměřen na konkrétní typy autorských slovníků.

1.3.1 Cizojazyčné autorské slovníky

Systematicky se autorské slovníky začaly zpracovávat zejména od 19. století.⁵² Jejich smyslem bylo v první řadě reprezentovat kulturní vyspělost národního jazyka prostřednictvím významného literárního představitele. Nejstarší autorské slovníky však zpracovávaly lexikální materii antických autorů.⁵³ T. Roelcke v již zmíněné studii konstatuje, že autorské slovníky v předvědeckém období sloužily zejména jako příručky pro poznání jazyka spisovatele; se stejným názorem jsme se setkali rovněž u Christy Dill. K takovým pracím náleží například latinsky psaný *Lexicon Aeschyleum* (1830), *Lexicon Homericum* (1885), *Lexicon Taciteum* (1830) a řada dalších.⁵⁴ Velkou tradici mají autorské slovníky zejména v německém a anglickém prostředí. J. Mattausch konstatuje, že v německé lexikografii dochází k výraznějšímu rozvoji autorské lexikografie především po druhé světové válce. Do té doby podle Mattausche prakticky neexistoval německý slovník, který by obsáhl

52 Mattausch dokonce uvádí, že v Německu již v roce 1724 existoval základ k autorskému slovníku Martina Luthera (Mattausch 1990, s. 1557).

53 Viz rovněž Králík 1961, s. 239.

54 Srov. Marie Těšitelová: „Pokusy o zjišťování slovní zásoby u některých autorů, i když ne vždy naprosto uvědomělé a úplně, jsou staršího data. Ve světové literatuře k nim můžeme počítat indexy k dílům klasických spisovatelů, např. k básnickému dílu J. Milтона (z r. 1857–59 a z r. 1867), k Homérově *Iliadě* (z r. 1875), k dílu Dantovu (z r. 1880) apod. Hlavním úkolem indexů obvykle bývá zjistit charakteristické rysy slovníku některého autora a ukázat, v čem navazuje na své předchůdce, popřípadě v jakém vztahu je k současníkům nebo co do něho přejala jeho následovníci.“ (Těšitelová 1955, s. 297)

kompletní autorský korpus. Slovník Ludwiga Gottfrieda Blanca,⁵⁵ který se o to poprvé pokusil, se orientuje na slovní zásobu Dantovy *Božské komedie* (Mattausch 1990, s. 1557). Ovšem ani v tomto případě se nejedná o kompletní výkladový slovník typu thesauru, jak Mattausch uvádí, ale o typ výběrového thesauru (*selektiv-thesaurisch*). Základní díla moderní autorské lexikografie⁵⁶ se objevují teprve ve druhé polovině 20. století. Náleží k nim v první řadě velký projekt německých lexikografů, který je spojen s excerpcí a systematizací Goethova díla, dále slovník Puškinův, Gorkého apod. Reflexe této situace, která zejména v německé lexikografii rezonovala v souvislosti s velkým Goethovým slovníkem (*Goethe-Wörterbuch*), zanechala v českém prostředí pouze skromnější odezvu. Dokladem je drobný článek Oldřicha Králíka *Autorské slovníky* publikovaný v roce 1961 v časopise *Slovo a slovesnost*, ve kterém autor k dané situaci poznamenává: „Typ, který se ustaluje v poválečných letech, usiluje postihnout v úplnosti jazykový odkaz národních klasiků a tím zvládnout závažnou část z vývoje národního jazyka. V SSSR vychází slovník Puškinův, v Polsku se chystá zvláště vyčerpávající slovník Mickiewiczův, na Ukrajině Ševčenkův, v Maďarsku Petöfiho a Juhásze, v Bulharsku vychází slovník Botevův, v Rumunsku se připravuje slovník Emineskův, v Německu se pokoušejí zvládnout obrovité úkoly slovníku Goethova.“ (Králík 1961, s. 239) S rozvojem velkých slovníkových projektů souvisí i metodologická debata, která směřuje k základnímu usměrnění a vymezení lexikografické práce; na straně druhé, jak dodává Mattausch, chybí doposud k dané problematice výraznější mezinárodní diskuze (Mattausch 1990, s. 1558).

1.3.1.1 Konkordanční slovníky

V roce 1875 vydal Edwin Abbott autorský konkordanční slovník k dílu Alexandra Popea pod názvem *A Concordance to the Works of Alexander Pope*. Již z názvu vyplývá, že se nejedná o výkladový slovník, ale o konkordanční typ slovníku, což znamená, že ve slovníku jsou zaznamenány výrazy (slova) spolu s nejužším kontextem. V tomto případě se nejedná ani o typ lemmatického slovníku, ve kterém by byly jednotlivé tvary uspořádány pod společné lemma, ale o výčet jednotlivých gramatických slovních forem a jejich kontextu. Tato strategie je zvolena proto, aby bylo zřejmé, nejen v jaké pozici se sledovaný výraz objevuje, ale i jaké má významové okolí, tedy v jaké sémantické platnosti jej autor použil. Slovníkové informace ovšem nevyplývají z připojených lexikografických komentářů

55 Mattausch hovoří o Blancově slovníku, který vyšel v roce 1852 v Lipsku pod názvem *Vocabolario Dantesco ou dictionnaire critique et raisonné de la Divine Comédie de Dante Alighieri* a vztahuje se k Dantově *Božské komedii* (Mattausch 1990, s. 1557).

56 Pojmem *autorská lexikografie* se snažíme odlišit oblast lexikografie specializující se na autorské slovníky od obecné lexikografie. Jedná se o pojem (kalk) přenesený z němčiny (*Autorenlexikographie*), v níž existuje řada prací na toto téma.

(interpretamentů), které jmenovaný Abbottův slovník neobsahuje, ale poskytují jej vlastní konkordanční řádky, resp. kontext, který je přístupný interpretaci *ex post*. Úkolem Abbottova slovníku, jak ostatně vyplývá i z autorova úvodního komentáře, je především demonstrovat autorský jazyk, jeho gramatiku a umělecké prvky, v tomto případě zejména rým. Abbottovou snahou je vytvořit na základě zpřístupnění Popeova autorského slovníku jakousi reprezentativní pomůcku pro uživatele anglického jazyka. Z toho důvodu také Abbott nevytváří výkladový, ale konkordanční slovník, který je schopen heslovitě uvádět konkrétní příklady užití jednotlivých výrazů. Slovník jasně dokládá, že impulzem k jeho zpracování byl nejen Popeův význam pro klasickou literaturu, ale rovněž vědomí reprezentativnosti jeho jazyka, který může prostřednictvím lexikografického zpracování posloužit za vhodný příklad kultivovaného zacházení s národním jazykem; v přímé souvislosti s uvedenými charakteristikami je rovněž didaktické zaměření slovníku.⁵⁷

Abbor----	{ <i>rhythmes</i> }	more-----S. iii. 65
air-----	{ <i>with</i> }	atmosphere----D. iv. 423
„-----		star-----R.L.i. 107
appear'd--		reward-----D. ii. 25
awake----		speak-----D. iv. 609
Baal-----		call-----„ „ 93

Ukázka 1: Doklad nepravidelných nebo neobvyklých rýmových skupin, který předchází vlastnímu slovníku. Údaje v pravém sloupci odkazují k zdrojovému textu a straně, např. písmeno D označuje Popeovo satirické dílo *The Dunciad*, označení iv. potom uvádí, že se jedná o čtvrtou knihu, číslo na konci uvádí paginaci. Přejato z Abbottova slovníku *A Concordance to the Works of Alexander Pope* (Abbott 1875, s. XVII).

57 „Every student of Modern English, and indeed every Englishman who wishes to use English words correctly, will find help from knowing how Pope used them. [...] Pope's English is not only correct, it is also, as Dryden's is, modern. There is no substantial difference between it and the English of the present day, except that Pope is more exact than most modern authors in the use of words. [...] To what extent French influence and court influence may have modified the language, or the introduction of the rhyming drama may have simplified the idiom by the exclusion of the old Elizabethan Periodic structure, this is not the place to consider, but whatever the causes may have been, the fact is certain, that we may glance down page after page of this Concordance without finding a line or phrase that might not have been written in the nineteenth century.

It is Pope's modernness, as well as correctness, that makes him so valuable a model for the student of modern English.“ (Abbott 1875, s. III, IV–V)

[„Každému studentu moderní angličtiny a samozřejmě každému Angličanovi, kteří hodlají správně užívat anglická slova, pomůže, bude-li vědět, jak jich užíval Pope. [...] Popeova angličtina je nejen správná, ale také – podobně jako Drydenova – moderní. Mezi současnou a Popeovou angličtinou není zásadní rozdíl, Pope je pouze exaktnější v užití slov než většina moderních autorů. [...] Na tomto místě se nebudeme zamýšlet nad tím, do jaké míry mohl jazyk ovlivnit působení francouzštiny nebo dvora nebo jak mohlo zavedení dramatu v rýmech zjednodušit způsob vyjadřování tím, že nahradilo starou alžbětinskou periodickou strukturu, ovšem ať již bylo příčinou cokoliv, je jisté, že můžeme projít tuto konkordanci (autor má na mysli Popeův slovník – pozn R. Z.) stránku po stránce a nenajdeme v ní řádku ani obrat, které nemohly být napsány v devatenáctém století. Díky své modernosti a přesnosti je Pope cenným vzorem pro studenty moderní angličtiny.“]

Rebel.

A R. to the very king he loves *M.E.* i. 203
 Bind r. Wit, and double chain on chain *D.* iv. 158
 Still r. nature holds out half my heart *E.A.* 26
 The r. Knave who dares his prince engage *R.L.* iii. 59
 But who, weak r-s, more advance her cause *D.* iv. 86
 Aspiring to be Angels, Men r. *E.M.* i. 128
 What serves one will, when many wills r. *E.M.* iii. 274
 Say, does thy blood r., thy bosom move *S.* iii. 55

Ukázka 2: Konkordance. Přejato z Abbottova slovníku *A Concordance to the Works of Alexander Pope* (ibidem, s. 251).

Na stejném principu jsou založeny také další konkordanční slovníky. Náleží k nim i slovník Daniela Barrona Brightwella *A Concordance to the Entire Works of Alfred Tennyson* (1869). Jak je zřejmé již z názvu, Brightwell zpracoval kompletní autorův korpus. Motivace k sestavení tak obsáhlého díla, jaké představuje tento slovník, je podobná jako v přechozím případě. Je jí snaha zpracovat tvorbu jednoho z významných reprezentantů anglické kultury a formou autorského konkordančního slovníku představit jazykovou a stylovou úroveň dobové angličtiny. Brightwell si byl vědom příliš velkého rozsahu své práce a komplikací, které by mohly pro uživatele nastat v případě dohledávání pojmů a slov, proto se pokusil zjednodušit systém tím, že před samotný slovník umístil tabulku rozsáhlejších básní, u kterých je vždy uveden odkaz na soubor, ve kterém se příslušná báseň (text) vyskytuje. V případě veršů je vždy vyznačen jejich rozsah a paginace. Kromě toho slovník obsahuje seznam excerpovaných textů a odkazy na soubory, ve kterých se tyto vyskytují, spolu s uvedením roku jejich vzniku. Určitá nevýhoda takto sestavených prací spočívá v tom, že při zpětném dohledávání širšího kontextu je uživatel odkázán vždy k určitému konkrétnímu vydání autorova díla, z něhož lexikografové pořizovali excerpta.

K Tennysonovi se takřka o půl století později vrátil Arthur Ernest Baker prací *A Concordance to the Poetical and Dramatic Works of Alfred, Lord Tennyson* (1914).⁵⁸ Jak sám v předmluvě uvádí, podnětem ke zpracování slovníku byly praktické požadavky čtenářů, zejména studentů anglické literatury (Baker 1967, s. VII). Na rozdíl od Brightwella, jehož Baker v úvodu nezmiňuje, nezpracovává celý autorský korpus, ale pouze jeho část, tedy poezii a dramatické texty. Ke struktuře svého slovníku poznamenává, že kromě toho, že je sestaven abecedně, uvádí pod samostatná hesla rozdílné gramatické i sémantické tvary slov, která jinak přísluší pod společné lemma: „The Concordance has been arranged in strict alphabetical sequence; the different sense or grammatical functions of a word are

⁵⁸ Celý název této práce zní: *A Concordance to the Poetical and Dramatic Works of Alfred, Lord Tennyson, including the Poems contained in the "Life of Alfred, Lord Tennyson" and the "Suppressed Poems", 1830–1868*. Názvy uvedené v uvozovkách jsou vydání Tennysonova díla, ze kterých Baker čerpal.

frequently distinguished under separate headings; the dialect words are paraphrased; all proper names are included, and occasionally some indications has been added of their identity.⁵⁹ (Ibidem, s. VII) Každé heslo tak obsahuje výčet identických tvarů, kdy je ke každému z nich na řádku připojen odkaz k názvu díla a paginace, která odkazuje ke zdrojovému vydání.

Kindness	<i>I could trust Your k.</i>	<i>To the Queen</i>	20
	looking ancient k on thy pain.	<i>Locksley Hall</i>	85
	I think your k breaks me down;	<i>Enoch Arden</i>	318
	money can be repaid; Not k		
	such as yours. ⁴	<i>Enoch Arden</i>	321
	Soul-stricken at their k to him,	<i>Aylmer's Field</i>	525
	summer of his faded love, Or ordeal by k;	"	561
	more in k than in love,	<i>Merlin and V.</i>	907
	your brother's love, And your good father's k. ⁴	<i>Lancelot and E.</i>	945
	he wrote 'Their k,' and he wrote no more	<i>To Marq. Of Dufferin</i>	36
Kindred (adj.)	But branches current yet in k veins. ⁴	<i>Princess ii</i>	245
	To black and brown on k brows.	<i>In Mem. lxxix</i>	16
	To-day they count as k souls;	" xcix	19
	But each has pleased a k eye,	" c	17
Kindred (s)	Grate her harsh k in the grass:	<i>Princess iv</i>	125
	Thy k with the great of old.	<i>In Mem. lxxix</i>	8
	craft of k and the Godless hosts Of heathen	<i>Guinevere</i>	427

Ukázka 3: Každý tvar má vlastní heslo. Převzato z Bakerova slovníku *A Concordance to the Poetical and Dramatic Works of Alfred, Lord Tennyson* (Ibidem, s. 358).

Konkordanční slovníky byly sestavovány také pro celou řadu dalších významných autorů, což je případ publikace *A New and Complete Concordance or Verbal Index to Words, Phrases and Passages in the Dramatic Works of Shakespeare with a Supplementary Concordance to the Poems* (1889),⁶⁰ jejímž autorem je John Bartlett. Autor na slovníku pracoval od roku 1876 a za podklad mu mimo jiné posloužila kompletní edice Shakespearova díla z roku 1875. Ve slovníku se Bartlett snaží zaznamenat kromě tradičních výrazů i takové, které jsou pro autora typické. Z toho důvodu uvádí jako samostatná hesla také

59 [„Konkordanční slovník je uspořádán do přesné abecední řady; rozdílné významy nebo gramatické funkce slov jsou často odlišeny samostatným heslem; slova, která náleží k dialektům, jsou parafrázována; všechna vlastní jména jsou ve slovníku zahrnuta a občas je jejich určení dodán nějaký příznak.“]

60 Označením „nový“ v názvu díla autor zřejmě reaguje na starší Shakespearův slovník, který vyšel v roce 1845 v Londýně a jehož autorkou je Mary Cowden Clarke. Celý název její práce zní *The Complete Concordance to Shakespeare being a Verbal Index to all the Passages in the Dramatic Works of the Poet*. Clarke vycházela z první úplné edice Shakespearova díla, kterou v osmi svazcích pod názvem *The Pictorial Edition of the Works of Shakespeare* vydal Charles Knight, editor a shakespearovský badatel. Bartlett v předmluvě svého slovníku konstatuje, že jeho nová práce vychází z modernější edice Shakespearova díla. Doslova uvádí, že základem mu byla „Globe Edition of Shakespeare“ z roku 1875; jedná se o šestisvazkovou publikaci *A Collection of the Plays, Romances, Novels, Poems and Histories Employed by Shakespeare in the Composition of his Works*, jejímž editorem je William Hazlitt. Bartlett se vůči staršímu Knightovu vydání vymezuje zejména po stránce rozsahu, úplnosti, a proto je pro své potřeby doplňuje o některé mladší ediční úpravy.

složené výrazy, resp. slovní konstrukty (lexikální syntagmata) sestavené ze dvou i více jinak samostatných slov. V těchto případech se vlastně nejedná o klasická kompozita, ale o seskupení (syntagmatiku) slov, která jsou pro autora příznaková. Skutečnost, že Bartlett tato syntagmata chápe jako jeden pojem, z nich v rámci autorského slovníku činí samostatné lexikální formy, které disponují specifickým koherentním významem.⁶¹ Na případě Bartlettova slovníku pozorujeme, že se autor snaží reflektovat i určité lexikální konstrukce, které chápe z hlediska Shakespearova díla jako symptomatická. Projdeme-li výskyty těchto konstrukcí (většinou se jedná o dvouslovná spojení), zjistíme, že jejich frekvence je minimální, tzn. jeden nebo dva výskyty na rozdíl od ustálenějších výrazů; např. spojení *bloody hand* je ve slovníku zastoupeno celkem sedmkrát, zatímco struktura *bloody hour* nebo *bloody house*, *bloody kind*, *bloody king* či *bloody knife* pouze jednou. Je ovšem třeba konstatovat, že v případě těchto lexikálních souvýskytů se nejedná o prostý souvýskyt, který by bylo možné detekovat na základě kolokační frekvence, ale o ustálenou (pevnou) lexikální strukturu, jež je v rámci Shakespearova jazyka výjimečná (odtud i její nízká frekvence). Tento příklad ukazuje, že problematika frazémů může být studována rovněž v rámci relativně uzavřeného systému literárního díla, ve kterém jsou dané konstrukce schopny realizovat specifickou funkci a význam. Na rozdíl od výše uvedených prací Bartlett nezaznamenává veškeré kontexty určitého slova, ale pouze ty, které chápe oproti běžnému užití v jazyce jako symptomatické.

Blessed heavens.	O blessed heavens!	<i>Coriolanus</i>	iv 2	20
Blessed hour.	This is the period of my ambition: O this blessed hour!			
		<i>Mer. Wives</i>	iii 3	48
Blessed labour.	A blessed labour, my most sovereign liege			
		<i>Richard III.</i>	ii 1	52
Blessed land.	And that the people of this blessed land May not be			
	punish 'd	<i>3 Hen. VI.</i>	iv 6	21
Blessed lottery.	Octavia is A blessed lottery to him			
		<i>Ant. and Cleo.</i>	ii 2	248

Ukázka 4: Heslo *blessed* je ve slovníku zastoupeno jednatřicetkrát, avšak veškeré kombinace s tímto slovem, jež vytvářejí příznakovou lexikální strukturu, se objevují pouze jednou. Převzato z Bartlettova slovníku *A New and Complete Concordance or Verbal Index to Words, Phrases and Passages in the Dramatic Works of Shakespeare with a Supplementary Concordance to the Poems* (Bartlett 1937, s. 128).

61 Martina Kastnerová ve své publikaci *Shakespeare a teorie interpretace – hledání adekvátního interpretačního přístupu* (2011) uvádí, že Shakespearovo dílo bylo pro lexikografické účely vytěžováno již od 18. století: „Za výborného znalce Shakespearova díla byl považován [...] Samuel Johnson (1709–1784). Johnson se věnoval vlastní literární tvorbě, novinářské činnosti, ale též lexikologii a literární historii. Pracoval mj. i na vydání velkého dvousvazkového *Slovníku anglického jazyka*, přičemž některá slova v něm bývají doplněna i ilustrujícími citáty. Právě pro *Slovník* shromáždil tisíce citací ze Shakespeara, neboť již před tímto projektem navrhuje další, nové vydání Shakespeara, které prozatím však bylo nakladatelem zavrženo.“ (Kastnerová 2011, s. 25)

V roce 1917 připravila skupina badatelů z USA autorský slovník Johna Keatse pod názvem *A Concordance to the Poems of John Keats*.⁶² Důvody, které vedly k jeho sestavení, jsou podobné jako u všech předchozích případů, tj. vedle samotného významu autora jde o jeho osobitý způsob zacházení s jazykem. Autoři se pokoušejí dekodovat Keatsův básnický jazyk, který – jak sami podotýkají – byl v minulosti opředen spíše „esencialistickou“ interpretací.⁶³ Konkordanční slovník má dle Broughtonových slov systematicky utřídit jazykový materiál autora tak, aby bylo možné lépe a přehledně zkoumat atributy tohoto jazyka, tj. funkční užití výrazů, jejich význam a kontextuální strukturu (viz Broughton 1963, s. 8). Jedná se opět o standardní konkordanční slovník, jak jsme se s jeho podobou mohli setkat v předešlých případech. V úvodu Broughton pojednává nejen o Keatsově významu v rámci dobového kulturně-historického kontextu, ale současně popisuje zvláštnosti jeho básnického jazyka.⁶⁴ Mimo jiné uvádí, že Keats užívá vlastních netradičních kombinací slov či slovních forem zakončených atypickými sufixy. Slovník zaznamenává tato pro Keatsův jazyk příznaková slova z hlediska morfematické skladby či slovtvorného způsobu⁶⁵ (viz Ibidem, s. 13–14).

Již v tuto chvíli lze učinit některá zobecnění na základě dosavadního pozorování daného typu autorského slovníku. Hlavní smysl konkordančního autorského slovníku spočívá především v úsilí systematicky demonstrovat autorský lexikon spolu s minimálním, avšak funkčním kontextem. Absence výkladového aparátu (komentářů) vyplývá

62 Autory slovníku jsou Dane Lewis Baldwin, Leslie Nathan Broughton, Laura Cooper Evans, John William Hebel, Benjamin F. Stelter a Mary Rebecca Thayer.

63 Explicitně to reflektují následující Broughtonova slova: „We know also that by nature he possessed a rare and delicate sensitiveness to impression. In the words of his friend, the artist Haydon, ‘He was in his glory in the fields. The humming of a bee, the sight of a flower, the glitter of the sun, seemed to make his nature tremble; then his eyes flashed, his cheek glowed, his mouth quivered.’ All of this we may not understand, nor can we measure and explain all the powers and workings of genius; but a knowledge of the facts makes the poetry of Keats more comprehensible.” (Broughton 1963, s. VI–VII) [„Víme také, že byl od přírody obdařen vzácnou a jemnou citlivostí vůči dojmům. Slovy jeho přítele umělce Haydona: ‚V polích byl ve svém živlu. Bzuket včel, pohled na květinu, třpyt slunce se zdály rozechvívát jeho povahu; oči se mu rozzářily, tváře rozjasnily, ústa rozechvěla.‘ Tomu všemu nemusíme rozumět, ani nemůžeme změřit a vysvětlit veškerou moc a sílu génia; avšak znalost faktů činí Keatsovu poezii srozumitelnější.“]

64 „Although Keat’s vocabulary does not, perhaps, contain a large number of root words, his poetry, nevertheless, shows variety in phrase and in combinations of words.” (Ibidem, s. VIII) [„Ačkoli Keatsův slovník možná neobsahuje velké množství slovních kořenů, jeho poezie vykazuje rozmanitost frází a kombinací jednotlivých slov.“]

65 Z toho, co Broughton uvádí, vyplývá, že se nejedná o standardní anglická kompozita, ale o zcela příznakové složeniny, které mají relevantní platnost coby funkce pojmenování právě v oblasti Keatsova básnického jazyka. Právě systematicky zpracovaný autorský slovník dokáže takové zvláštnosti nejen utřídit, ale současně poskytnout fundovaný základ pro detailnější studium podobných jevů. Broughton to ostatně jasně konstatuje v úvodu: „Through the Concordance it is now easier to obtain an accurate conception of the size of Keat’s vocabulary in his poems, its variety, the extent of its excesses and mannerisms, and especially its sources.” (Ibidem, s. 8) [„Díky této konkordanci je nyní snazší vytvořit si přesný obraz toho, jak rozsáhlý je slovník Keatsových básní, jaká je jeho rozmanitost, jaký je rozsah jeho idiosynkrazí a manýrismů a zejména z jakých pochází zdrojů.“]

z typu autorského slovníku, který v první řadě operuje s výčtovostí lexikálních jevů a jejich dokumentací. Proto jej lze velmi dobře využívat i pro didaktické účely, jak jsme se o tom mohli u některých příkladů přesvědčit. Tomu odpovídá možnost rychlého vyhledávání, a především konkrétní doklady užití daného výrazu, ze kterých vyplývá jeho funkčně-sémantická charakteristika.

Konkordanční slovníky mohou být i vícejazyčné, jako je tomu v případě autorského slovníku *A Concordance of the Latin, Greek and Italian Poems of John Milton* (1923), jehož autorem je Lane Cooper.⁶⁶ Takto zaměřený slovník dokáže shromáždit a utřídit materiál různých jazykových proveniencí a vytvořit základ pro komparativní studium slovních či kolokačních souvislostí. V případě Cooperovy práce se jedná o Miltonovy básně psané v latině, řečtině a italštině. Rovněž zde, jak uvádí Cooper, byla pro sestavení slovníku hlavní motivací jeho praktická stránka a spolu s ní i potřeba systematizovat lexikální korpus autora na takové úrovni a komplexnosti, aby mohl být využíván pro hlubší a systematictější odborné studium.⁶⁷

Od druhé poloviny 20. století se začínají v lexikografii čím dál častěji využívat počítačově zpracovaná data. Využití elektroniky má v tomto ohledu značné přednosti, a i přes některé námitky (viz stanovisko H. E. Wieganda) se tato oblast stává významnou a důležitou součástí moderní lexikografie, která nachází uplatnění také v rámci autorských slovníků. V roce 1971 vyšel v Anglii konkordanční slovník zaměřený na vybranou tvorbu George Bernarda Shawa, který pod názvem *A Concordance to the Plays and Prefaces of Bernard Shaw* sestavil Earl Dean Bevan. Jedná se o počítačově zpracovaný slovník, jenž využívá v té době nově dostupnou metodu děrných štítků. Každé slovo bylo definováno na speciálním štítku sérií vyražených znaků (atributů) uspořádaných do matic a následně

66 Kromě tohoto slovníku vydal v roce 1916 slovník věnovaný římskému básníkovi Quintu Horatiu Flaccovi pod názvem *The Concordance to the Works of Horace*.

67 „I trust it will be helpful to students who might otherwise, with the aid only of Bradshaw's *Concordance* or Lockwood's *Index*, confine themselves to Milton's poems in English. One can hardly think of a Miltonic investigation which should not take the Latin verse into account. Henceforth, the study of the poet on the side of his verse, if not of his prose, may be carried out as fully and exactly as the study of Dante, or the Bible, or Homer and the ancient poets generally.“ (Cooper 1923, s. IX)

[„Věřím, že pomůže studentům, kteří by se jinak – pouze na základě Bradshawovy *Konkordance* nebo *Indexu* Lockwoodovy – omezili na Miltonovu poezii v angličtině. Lze těžko pomyslet na miltonovské bádání, které by nebralo v úvahu latinské verše. Od této chvíle může být studium jeho poezie, když už ne jeho prózy, po stránce verše realizováno s takovou úplností a přesností, jako je tomu při studiu Danta, bible, Homéra a vůbec antických autorů.“]

V citaci Cooper odkazuje na dva slovníky. Jedná se o práci pod názvem *A Concordance to the Poetical Works of John Milton* (1894), autorem je John Bradshaw, který sestavil *Anglickou antologii od Chaucera po Tennysona*. Druhý slovník se jmenuje *Lexicon to the English Poetical Works of John Milton*, připravila jej Laura Emma Lockwood. Miltonově tvorbě (podobně jako u jiných autorů a jejich děl) je věnováno více slovníků. V roce 1867 vyšel rovněž miltonovský slovník *A Complete Concordance to the Poetical Works of John Milton*, jehož autorem je Charles Dexter Cleveland. Cleveland v předmluvě upozorňuje, že v předchozím miltonovském slovníku, který si vzal jako podklad ke své práci (jedná se o publikaci Henryho Johna Todda *Verbal Index to the Poetry of Milton* z roku 1809), našel 3362 chyb (Cleveland 1867, s. V).

zpracovaných počítačem.⁶⁸ Tato metoda umožnila mechanickou automatizaci slov a jejich kontextu spolu se základním statistickým měřením, které v tomto případě představovala tzv. absolutní frekvence. Výhodou počítačem analyzovaných dat je zejména jejich přesnost, rychlost a variabilita spočívající v možnosti měnit kritéria (podmínky) vyhledávání, řazení a třídění. Strojové zpracování jazykové materie samozřejmě závisí nejen na vyspělosti počítačové technologie, ale rovněž na teoretické úrovni kvantitativní lingvistiky, jejímž úkolem je mimo jiné formulovat funkční a vzhledem k povaze zkoumaných jevů relevantní algoritmy. Je proto zřejmé, že způsoby automatického zpracování jazyka se v důsledku aktuálního vývoje moderních počítačových technologií a teoretických metod radikálně liší (co do možnosti kvantifikace) od předchozích stadií.⁶⁹

CAPT I SD	(217)	pin in it, he wears a suit of clean white linen	ACCEPTABLE	in color, if not in cut, to the moorish mind, the
JOAN 6	(143)	all the duties which have made my sex specially	ACCEPTABLE	in heaven, have taken up the sword, even to the
SUPR HANDBOK	(208)	normal life of his forefathers, this view is very	ACCEPTABLE	to Englishmen, who always lean sincerely to
SUPR HANDBOK	(179)	race afoot (the hive and the anthill being as	ACCEPTABLE	to her as utopia), the demonstrations of
HART PREFACE	(31)	the conventions of the theatre can be easily	ACCEPTABLE	, or the purpose of the drama obvious to him.
MTH4 II	(189)	for the reception of a colonizing party would be	ACCEPTABLE	, they are at your disposal. And when I say a

Ukázka 5: Strojový výčet konkordancí se zkratkami zdrojů. Přejato z Bevanova slovníku *A Concordance to the Plays and Prefaces of Bernard Shaw* (Bevan 1971, s. 38).

Děrnostítkové metody je využito v případě slovníku *A Concordance to Beowulf* (1969),⁷⁰ který edičně připravil Jess Balsor Bessinger, Jr., a naprogramoval Phillip Hagar Smith, Jr. Jak ostatně uvádí P. H. Smith, nový konkordanční slovník na rozdíl od Cookovy práce vyniká přesností právě díky počítačovému zpracování korpusových hodnot a rovněž díky svému většímu rozsahu (Bessinger – Smith 1969, s. IX, X). To, co jsme konstatovali výše, tj. že počítačově interpretovaná data vykazují vyšší potenciál, zejména co se týče možností statistické systematizace a analýzy, se ukazuje také na příkladu tohoto slovníku. Kromě konkordančního seznamu, který je vytvořen metodou KWIC (Key Word in Context), obsahuje abecední řazení hesel dle klesající frekvence. Slovník disponuje

68 Děrnostítkovou metodou byl zpracován i český autorský slovník věnovaný Petru Bezručovi. Děrný štítek v tomto případě tvořilo 80 sloupců a 8 řádků; sloupce byly navíc segmentovány do skupin po deseti. Takto bylo vytvořeno pole (sít) o velikosti 640 bodů, kde každý z nich mohl nést vlastní informaci specifikovanou konkrétní číselnou souřadnicí.

69 O současných možnostech a směrech v kvantitativně orientovaných literárněvědných studiích pojednáme v závěru publikace.

70 Již v roce 1911 vznikl stejnojmenný slovník, jehož autorem je profesor anglické literatury a jazyka na univerzitě v Yale Albert Stanburrough Cook. Jedná se o konkordanční typ, který ovšem vypouští některá slova, zde konkrétně číslovky, předložky a zájmena (viz Cook 1911, s. 3).

třemi způsoby řazení hesel: abecedním, konkordančním a frekvenčním, přičemž druhý typ v sobě kombinuje oba zbývající.⁷¹

a)

ABROTEN (2)

ðæ t hine seo brimwylf abroten hæfde. 1599
ond hi hyne ða begen abroten hæfdon, 2707

b)

1599	ða ðæs monige gewearð ðæ t hine seo brimwylf	a b r o t e n	hæfde. Ða com non dægæs. Næs ofgeafon hwate
2707	(ferh ellen wræc), ond hi hyne ða begen	a b r o t e n	hæfdon, sibæðelingas. Swylc sceolde secg wesan,
187	ne wenan, white gewendan: wel bið ðæm ðe mot	æ f t e r d e a ð d æ g e	drihten secean ond to fæder fæðmum
885	cynnes sweordum gesæged. Sigemunde gesprong	æ f t e r d e a ð d æ g e	dom unlytel, syððan wiges heard wyrn

Ukázka 6: Strojové konkordance s uvedením paginace a absolutní frekvencí výskytu slova. První část ukázky je zkrácenou konkordancí s frekvencí uvedenou v závorce. Přejato z Bessingerova a Smithova slovníku *A Concordance to Beowulf* (Bessinger – Smith, 1969, s. 1, 285).

Na uvedených příkladech, které v žádném případě nepodávají vyčerpávající přehled,⁷² jsme mohli sledovat, čím se konkordanční typ slovníku vyznačuje. Jeho základní bázi tvoří soupis všech slov v autorském korpusu spolu s definovaným kontextem, který slouží především k heslovitému přehledu o způsobu zapojení výrazu do kontextu, ze kterého by měly být identifikovatelné základní funkčně-sémantické vlastnosti daného lexika. Kvůli snazšímu hledání jsou konkordanční slovníky řazeny abecedně a sestaveny buďto formou hesla-lemmatu, nebo úplného výčtu všech tvarů slova uvedených vždy na samostatném konkordančním řádku. Důmyslnější systematizaci v klasifikaci hesel umožnila zejména počítačová technologie. Jak jsme mohli sledovat, konkordanční slovník může excerpovat celý autorský korpus nebo pouze jeho vybranou část. Kromě svého odborného zaměření sloužily autorské konkordance didaktickým účelům, a to jako ukázky vhodného (především sémantického) užití slov v syntaktických konstrukcích.

71 První typ je založen na abecedním řazení hesel, u kterých je v závorce připojen údaj o frekvenci. V rámci hesla je potom vyznačen minimální kontext, ve kterém se výraz vyskytuje. Druhý typ se zakládá na abecedním řazení všech výrazů excerpovaného korpusu, jejichž četnost je dána variantností jejich kontextuálního zapojení. V daném případě se jedná rovněž o rozšířenější kontext, který umožňuje přesnější sémanticko-funkční specifikaci výrazu. Poslední způsob řazení, který je ve slovníku použit, se odvíjí od klesající frekvence. Tento seznam tvoří slova tříděná do frekvenčních skupin.

72 Na Cornellově univerzitě, která se stala jedním z center pro sestavování autorských slovníků, kromě výše jmenovaného Bessingerova slovníku vyšly slovníky věnované básnické tvorbě Matthewa Arnolda, W. B. Yeatse, Emily Dickinsonové, Williama Blakea, Jeana Racina nebo slovník zaměřený na Byronova *Dona Juana*. Kromě seznamu, který je uveden v Bessingerově slovníku, odkazuje na řadu významných autorských slovníků rovněž zde citovaný J. Mattausch v bibliografii své studie *Das Autoren-Bedeutungswörterbuch*.

1.3.1.2 Výkladové slovníky (v redukované podobě)

Také v případě výkladových slovníků se lze setkat s jejich různými podobami. Je možné se zaměřit na vybranou oblast autorova díla nebo zpracovat dílo v celém rozsahu, jak o tom hovoří například J. Mattausch. V této kapitole pojednáme o typologicky odlišných autorských slovnících, než o jakých jsme hovořili doposud.

Název, který jsme pro tuto kapitolu zvolili, není terminologickým označením,⁷³ ale pracovním vymezením skupiny slovníků, které disponují omezeným komentářem hesla nebo zpracovávají pouze část autorského materiálu, např. skupinu děl nebo dokonce dílo jediné, popř. kombinují oba přístupy. V zahraničním prostředí se v tomto případě často setkáváme s označením *glossary*. Na rozdíl od konkordančního slovníku již tento typ obsahuje ke každému z uvedených hesel alespoň stručný komentář. Nejedná se sice o plnohodnotný thesaurus, neboť komentáře nepodávají natolik komplexní informace, jako je tomu u thesauru (viz následující kapitolu), ovšem na rozdíl od konkordančního slovníku jsou jednotlivá hesla opatřena kratším výkladovým aparátem, který stručně a přehledně komentuje význam a užití výrazu v autorském korpusu. Jedná se většinou o sémantickou klasifikaci pojmu v závislosti na jeho kontextu. Součástí těchto slovníků jsou samozřejmě i doklady užití daného výrazu, které vždy nemusejí reprezentovat úplné konkordanční výčty, ale například pouze redukované doklady typologicky příznakových kontextových struktur.

Podobně jako v případě konkordančních autorských slovníků začal i tento typ vznikat již v 19. století, jak dosvědčuje práce Alexandra Dycea (a Harolda Littledalea⁷⁴) věnovaná W. Shakespearovi; její název zní *A Glossary to the Works of William Shakespeare* (1880). Dyce se v předmluvě dotýká jednoho z typických problémů, které souvisí se způsobem realizace komentářů.⁷⁵ Jelikož se jedná o beletrii, dochází velice často k tomu, že slov je užito v modifikovaném či přeneseném významu. Jak ukážeme zejména na příkladu Goethova slovníku, je nezbytné, aby v případě těchto komentářů platily takové odborné zásady, které dokážou eliminovat nežádoucí subjektivně-interpretační hledisko. Kromě toho může ovšem dojít i k celé řadě jiných sporných či problémových jevů. Komplikovanost a zejména náročnost, s jakou se autor takového slovníku musí vypořádat,

73 Používá se obvykle pouze označení výkladový slovník, který se standardně dělí na jednojazyčný a vícejazyčný.

74 Vycházíme z prvního reprintu (1970) vydání z roku 1902, na kterém se podílel H. Littledale, profesor anglického jazyka a literatury na univerzitě v Jižním Walesu.

75 „Among the difficulties incident to a glossarist not the least is that of determining the nicer shades of meaning in which many words are used; and very probably some philologists may think that I have occasionally made distinctions where none in fact exist, and sometimes confounded what ought to have been kept distinct.” (Dyce 1970, s. 5)

[„Mezi obtížemi, které náleží k práci glosátora, patří mimo jiné určování jemnějších významových odstínů v užití jednotlivých slov. Někteří filologové si nejspíš mohou myslet, že jsem v některých případech stanovil rozdíly tam, kde ve skutečnosti žádné neexistují, zatímco jindy spojil to, co by mělo být odlišeno.“]

dosvědčují i slova H. Littledalea, jenž v souvislosti se slovníkem Alexandera Schmidta, jehož celý název zní *Shakespeare-Lexicon: Vollständiger englischer Sprachschatz mit allen Wörtern, Wendungen und Satzbildungen in den Werken des Dichters* (1874), konstatuje v předmluvě (prvního) reprintu Dyceova slovníku: „It is no detracting from the value of Schmidt’s monumental work to say that it gives sometimes too much and sometimes too little. Its very fulness of reference in the case of words and phrases that need no commentary is not seldom a positive hindrance to the seeker for light, and on the other hand the comprehensiveness of Schmidt’s plan does not admit of the adequate presentation of opposite points of view or detailed illustration of specially difficult or interesting matters.“⁷⁶ (Littledale 1970, s. VII) Především se tedy jedná o zásadu věcné charakteristiky tvaru a významu lexikální jednotky spolu s jeho kontextovým určením, jež se váže nejen k oblasti textu, ale současně k dalším rámcům, mezi které náleží dobově-historický zřetel nebo aspekty pragmatické. Povaha a charakter komentářů rovněž závisí na vývojovém stupni příslušných vědeckých disciplín (literárněteoretických a historických, ale současně i lingvistických, textologických a řady dalších). Na rozdíl od konkordančního slovníku zde dochází k rozšíření formální části o komentované oddíly; důležitou úlohu tak má především metajazyk a metodologická strategie. Dyceův slovník dle Littledaleových slov představuje jakýsi přechodový stupeň mezi konkordančním typem a výkladovým slovníkem (Ibidem, s. VII). Jak můžeme sledovat na struktuře hesla Dyceova – Littledaleova slovníku, jedná se většinou o vysvětlení rámcového významu příslušného slovního tvaru (platí především pro apelativa) vzhledem k jeho bezprostřednímu kontextu, tedy nikoli o jeho obsáhlejší charakteristiku, která by zahrnovala i další aspekty (pragmatické, historické, etymologické apod.), jako například v *Goethe-Wörterbuch* (viz dále). Jiná je však situace u vlastních jmen, kde může být komentář podrobnější, zejména pak v případech, kdy je významově funkční ozřejmit například jejich genezi (viz Ukázkou 7).

76 [„Není snižováním hodnoty Schmidtova monumentálního díla, pokud konstatujeme, že někdy poskytuje příliš mnoho a jindy příliš málo. Obsírnost jeho odkazů v případech slov a frází, které nepotřebují další komentář, je nezdávkou překážkou tomu, kdo sám tápe ve tmě, a na druhou stranu komplexnost Schmidtova plánu nepřipouští adekvátní uvedení rozdílných úhlů pohledu nebo detailní ilustraci obzvláště obtížných nebo zajímavých otázek.“]

Acheron, M.N.D. iii. 2. 357; Tim. iv. 3. 44; Mcb. iii. 5. 15:
 „The infernal river, supposed by Sh. to be a burning
 lake“ (SCHMIDT): Fairfax, in his translation of the *Gerusalemme* (published before *Macbeth* was produced), has made
 Ismeno frequent „the shores of *Acheron*,“ without any
 warrant from Tasso;

„A Christian once, Macone he now adores,
 Nor could he quite his wonted faith forsake,
 But in his wicked arts both oft implores
 Helpe from the Lord and aide from Pluto blake;
He, from deepe caues by Acherons darke shores
 (Where circles vaine and spels he vs'd to make),
 T'aduse his king in these extremes is come;
 Achitophell so counsell'd Absalome.“

B. ii, st. 2.

The original has merely

„Ed or dalle spelonche, ove lontano
 Dal volgo esercitar suol l'arti ignote,
 Vien,“ &c.

For instances how loosely the name *Acheron* is used by
 our early poets, see, in Sylvester's *Du Bartas*, ed. 1641,
The Second Day of the First Week, p. 15, *The Vocation*,
 pp. 149, 155, and *The Fathers*, p. 162; also Hubert's
Edward the Second, p. 161, ed. 1629.

Ukázka 7: Na rozdíl od Schmidtova slovníku (viz Ukázka 8) je zde podrobněji komentován způsob, jakým Shakespeare v *Macbethovi* použil slovo *Acheron*. S odkazem na Schmidtův výklad (viz zahrnutou citaci části Schmidtova hesla) se zde konstatuje, že Shakespeare užil slova *Acheron* ve smyslu podsvětní řeky s tím, že předpokládal, že význam souvisí s pekelným hořícím jezerem. Jak je ovšem uvedeno v Dyce – Littledaleově výkladu, tento význam byl iniciován Fairfaxovým poněkud volným překladem (z roku 1600) části Tassova *Osvobozeného Jeruzaléma*, resp. jeho užitím slova *Acheron* v místě, kde se u Tassa nevyskytuje. Zde jsou srovnávány obě tyto pasáže, tj. Fairfaxův překlad a originál Tassova textu. Na tomto příkladě rovněž vidíme to, co měl Littledale na mysli, když Schmidtovu slovníku vytýkal místo jeho přílišnou strohost ve výkladu (srov. Ukázku 7 a 8). Přejato z Dyce – Littledaleova slovníku *A Glossary to the Works of William Shakespeare* (Dyce – Littledale 1970, s. 3–4).

Acheron, the infernal river, supposed by Sh. to
 be a burning lake: Mids. III, 2, 357. Tit. IV, 3, 44.
 Mcb. III, 5, 15. (cf. H4B II, 4, 170. H6B I, 4, 42.
 Lr. III, 6, 8).

Ukázka 8: Přejato ze Schmidtova slovníku *Shakespeare-Lexicon. A Complete Dictionary of all the English Words, Phrases and Constructions in the Works of the Poet* (Schmidt – Sarrazin 1902, s. 12).

Shakespearovský materiál se podobně jako u jiných významných autorů ukázal jako velmi cenný zdroj pro jazykové a literární studium. V roce 1910 vyšel další shakespearovský komentovaný slovník pod názvem *A New Shakespearean Dictionary*⁷⁷ autora Richarda Johna Cunliffea. Jeho hlavním záměrem, jak autor v předmluvě konstatuje, bylo vytvořit seznam slov, která jsou pro Shakespeara příznaková z hlediska funkčního i sémantického

77 Slovo „nový“ v názvu odkazuje ke Schmidtovu a Bartlettovu slovníku.

užití.⁷⁸ Jak rovněž Cunliffe dodává, slovník neposkytuje úplný seznam všech slovních jednotek Shakespearova díla, ale zejména ty formy, které z aktuálního hlediska představují vlivem historického vývoje jazyka odlišný způsob užití oproti aktuální jazykové situaci. Na konkrétních příkladech Cunliffe demonstuje, jak se mění interpretace významů, které na první pohled mohou působit jednoznačně a nekomplikovaně. Zaměřením je slovník podle jeho autora určen především těm, kdo usilují o hlubší a odborné poznání Shakespearovy slovní zásoby, zejména o přesné definování její sémantiky (Cunliffe 1910, s. VII).

ADVERTISEMENT

(1) Notification, information: *This advertisement is five days old* – I H₄ III₂ 172.

(2) Admonition, warning, precept: *My griefs cry louder than advertisement* – Ado V₁ 32. *An advertisement to a proper maid ... to take heed of ...* – All's IV₃ 240. *Yet doth he give us bold advertisement That ... we should on* – I H₄ IV₁ 36.

Ukázka 9: Uvedena je dvojí rozdílná významová charakteristika lexika a současně je doložena jeho kontextem. Přejato z Cunliffého slovníku *A New Shakespearean Dictionary* (ibidem, s. 5).

Dalším příkladem lexikografické práce, kterou bychom mohli zařadit do námi sledované kategorie autorských slovníků s redukovanou podobou a kterou zde pouze zmíníme, může být jiný shakespeareovský slovník *A Shakespeare Glossary* (1911), jehož autorem je Charles Talbot Onions. Blízkým typem slovníku z hlediska způsobu zpracování hesla je například *A Chaucer Glossary* (1979), na kterém se podíleli Norman Davis, Douglas Gray, Patricia Ingham a Anne Wallace-Hadrill. Podobně jako u Cunliffého slovníku ani zde nejde o kompletní excerpci celého Chaucerova lexikonu, ale o výběr takových jevů, které jsou svým způsobem příznakové ve způsobu jejich užití. Konkrétně se jedná o zaměření na lexikum, které z pohledu současného jazyka (angličtiny) představuje méně časté či již zaniklé jevy, a současně o doklady specifického básnického zacházení s jazykem, jež je typické pro daného autora.⁷⁹ Takto speciálně koncipované slovníky jsou obvykle zaměřeny

78 „I have endeavoured to exclude all words and senses of words which are still in good literary use, and, except where there is obscurity, all senses which are merely contextual and do not represent an authenticated usage.“ (Cunliffe 1910, s. V)
[„Snažil jsem se vyloučit všechna slova a jejich významy, které se stále běžně vyskytují v literatuře, a – s výjimkou případů, které by mohly být nesrozumitelné – rovněž takové významy, které jsou čistě kontextového charakteru a nerepresentují autentické užití jazyka.“]

79 „The primary aim of this Glossary is to explain the meanings of words and phrases in Chaucer's works which are used in ways unfamiliar in modern English, and to refer them to a number (necessarily limited) of typical instances. It also illustrates Chaucer's use of many expressions not necessarily unfamiliar but characteristic of his language. [...] The Glossary does not give a complete account of Chaucer's vocabulary, because that is available in the Concordance compiled

na užší skupinu, kterou v první řadě tvoří studenti a vědečtí pracovníci; ostatně v některých pracích je to výslovně uvedeno. Chaucerův slovník je navíc doplněn úvodním pojednáním o básnickém jazyce s detailnějším rozpracováním jazykové situace v rámci jednotlivých slovních druhů; pozornost je tu věnována stavbě a struktuře slovních druhů, např. pro autora typickým způsobům slovního zakončení. Slovníková hesla pak kromě krátké charakteristiky významu, resp. uvedení analogického tvaru v současném jazyce obsahují i stručnou etymologickou charakteristiku formou (zdrojového) odkazu k diachronnímu stupni příslušné jazykové proveniencie (např. stará angličtina, střední latina, stará němčina, severní dialekty staré francouzštiny apod.).

Příkladem autorského slovníku, který zpracovává jediné dílo, je *Wörterbuch zu Goethes Werther* (1966) Erny Merker.⁸⁰ Jedná se o slovník excerpující slovní zásobu Goethovy epistolární novely *Utrpení mladého Werthera* (1774).

absterben *langsam dahinsterben* ich könnte mich vertrauen, wenn so ein paar Bäume in meinem Hofe stünden, und einer davon stürbe vor Alter ab 98; *übtr auf das Leben selbst* dem Unglücklichen, dessen Leben unter einer schleichenden Krankheit unaufhaltsam allmählich abstirbt 49 ich möchte gleich einem edlen Waffenträger das Schwert ziehen und meinen Fürsten von der zückenden Quaal des langsam absterbenden Lebens auf einmal befreien 101 *oss*

[...]

acht 1 *Zahlwort a. Geschwister* 19; *als Uhrzeit elliptisch* zauderte bis Acht [B a.] 131; *ifst Wdg für: eine Woche seit a. Tagen* 78B 115uö 2 *Zahlzeichen* 8, *subst, als Tanzfigur* als wir uns scheiden mußten die grosse Achte zu machen 26

[...]

Albert *neben der heute übl Flexion findet sich in A u B vorwiegend die alte Flexion: Gen Albertens, Dat u Akk Alberten; ebs bei den übrigen Eigennamen* aus dem Munde... Albertens 116 A den Frieden zwischen Alberten und seiner Frau 116 A Oftbened ich Alberten 62 uö

Ukázka 10: Přejato ze slovníku Erny Merker *Wörterbuch zu Goethes Werther* (Merker 1966, s. 5, 6, 8).

by J. S. P. Tatlock and A. G. Kennedy (the Carnegie Institute of Washington, 1927, reprinted Gloucester, Mass., Peter Smith, 1963).“ (Davis – Ingham – Wallace-Hadrill 1979, s. VII)

[„Hlavním záměrem tohoto slovníku je vyložit význam slov a frází z Chaucerových děl, které jsou užity takovým způsobem, který je pro moderní angličtinu neznámý, a odkázat na typické příklady (jejichž počet je nutně omezený). Slovník rovněž ilustruje Chaucerovo užití mnoha výrazů, které nejsou nutně neznámé, ale jsou charakteristické pro jeho jazyk. [...] Slovník nepodává úplný výčet Chaucerova lexikonu, protože ten je dostupný v podobě konkordančního slovníku sestaveného J. S. P. Tatlockem a A. G. Kennedym (Carnegie Institut ve Washingtonu, 1927, přetisk Gloucester, Mass., Peter Smith, 1963).“]

80 Kromě uvedené autorky se na slovníku podíleli Johanna Graefe a Fritz Merbach.

Základním metodologickým postupem strukturování hesla se zde stalo významové kritérium. Na ukázce můžeme sledovat sémantické specifikace jednotlivých slov. Každé heslo je určeno svým kontextem. V případě slovesa *odumírat* (absterben) se jedná o použití ve spojení s životným a neživotným subjektem. První případ je dokladován v rámci užití výrazu ve spojení slovesa a substantiva *strom*, přičemž ve výkladové poznámce je význam určen souslovím *pomalu odumírat*. V českém překladu zní citovaná pasáž z Goethova *Werthera* následovně: „Já, jenž bych se usoužil, kdybych měl pár takových stromů na svém dvorku a jeden by mi věkem ztrouchnivěl, já se musím na tohleto dívat!“⁸¹ Druhý způsob, jakým je sloveso použito v textu, je pomocí interpretamentu definován jako „přenesení na vlastní život“ (*übertragen auf das Leben selbst*); v překladu: „A můžeš chtít na nešťastníkovi, jenž nezadržitelně pozvolna umírá na zákeřnou nemoc, aby rázem, bodnutím dýkou učinil konec svým útrapám?“⁸² V podobném významu je slovesa užito i v posledním případě, který je v hesle uveden a jehož překlad je: „Ó příteli, jako šlechtný zbrojnoš bych tasil meč a osvobodil svého knížete rázem od hrozných muk ponenáhlého odumírání; a za osvobozeným polobohem bych poslal vlastní duši.“⁸³ Podobně je tomu u dalších hesel; u číslovky *osm* (acht) je definován způsob užití výrazu, a to jednou jako prosté číslovky, jindy jako elipsy vyjadřující počet hodin a v jiném případě taneční figury. Posledním heslem, které zde uvádíme, je vlastní jméno *Albert*. Následující komentář ukazuje další možnou povahu výkladového aparátu spočívající v objasnění historické gramatiky: v obou excerpovaných verzích (*A* z roku 1774 a *B* z roku 1786) je jméno v genitivu i akuzativu zapsáno způsobem, který je odlišný od současné němčiny. Jak je uvedeno v komentáři, jedná se o starou flexi (*alte Flexion*), konkrétně o starou koncovku.⁸⁴

Slovník sice obsahuje bohatý a různorodý metalexikografický komentář k excerpovanému slovesnému dílu, avšak jedná se o takový typ, který Mattausch nazývá výkladovým slovníkem jednoho díla (*Werkwörterbuch*). V případě Goethova díla existuje také autorský thesaurus, o kterém stručně pojednáme v následující kapitole. Na uvedených příkladech hesel lze dokumentovat, že v případě výkladových autorských slovníků plní lexikografický komentář širší spektrum funkcí, které zahrnuje nejen významovou interpretaci, ale například také vysvětlení dobového způsobu užití či zvláštnosti historické gramatiky, způsob psaní apod. Na tomto slovníku je zjevná důslednost, s jakou se k výkladu hesel přistupuje, což se vyznačuje zejména důkladnou obeznámeností s různými verzemi původního textu.

81 Kurzivou je zvýrazněna část, která je citována ve slovníku. Český překlad Erika A. Saudka (Goethe 1964, s. 73).

82 Kurzivou je zvýrazněna část, která je citována ve slovníku. Český překlad Erika A. Saudka (Ibidem, s. 41).

83 Kurzivou je zvýrazněna část, která je citována ve slovníku. Český překlad Erika A. Saudka (Ibidem, s. 74).

84 V současné němčině by jméno *Albert* mělo v genitivu koncovku -s, tzn. *Alberts* (popřípadě bez koncovky), zatímco v dativu a akuzativu by zůstalo bez koncovky, tzn. stejně jako v nominativu by zde byl tvar *Albert*.

Jiným výkladovým slovníkem, který excerpuje nikoli celé autorské dílo, ale jeho vybranou část, v tomto případě pouze české texty Jana Amose Komenského, je slovník Huberta Rösela *Wörterbuch zu den tschechischen Schriften des J. A. Comenius* (1983). U nás o něm podrobněji pojednal například Zdeněk Tyl, který práci vytkl řadu nedůsledností ve věci excerpce a struktury hesla nebo též na rovině významové analýzy (viz Tyl 1985). Oproti slovníku Erny Merker má Röselova práce výrazněji redukované komentáře, které se většinou omezují na překlad hesla do němčiny, popř. na synonyma v současné češtině. V určitých případech může být z hlediska aktuálnosti či relevance způsob výkladu dokonce sporný (viz poznámka k Ukázce 11). Tyto příklady ukazují na možné komplikace, které mohou vzniknout při volbě konkrétního způsobu realizace interpretamentu. Na uvedených novočeských ekvivalentech je patrné, že jsou vystaveny přirozenému zastarávání, nebo dokonce že uvádějí neadekvátní ekvivalent. Na rozdíl od předchozího slovníku se tu vlastně nejedná o aplikaci metajazyka, neboť namísto interpretace (viz uvedená hesla slovníku E. Merker) je v tomto případě uveden překlad nebo historický ekvivalent.

frej m milostky, zálety Buhlerei, Freien, Lieblei, Liebeshandel
nedopouštět žádná neslušnosti ... v ... frejích (PP 439, 11); frašky básně freje (LS 230, 23); Místo čistoty ... freje, oplzlosti (D 24, 33); tanců, frejů, lichev ... hleděli (L 55, 30)

[...]

frejř m nápadník, záletník Freier, Liebhaber, Buhler
Frejř když obdrží aby mu [panna] zasubena byla, bude ženichem (DJO 48, 22); shromáždím frejře své (M 462, 13)

Ukázka 11: U hesla *frej* je použito jako synonymum vedle slova *zálety* i zcela atypické a běžně neužívané *milostky*. U hesla *frejř* jsou zase použita synonyma, která například z hlediska současné jazykové (lexikální) situace ustupují frekventovanějšímu *milanec* [nápadník = 406 Af, resp. 3,34 Rf (i.p.m.); záletník = 41 Af, resp. 0,34 Rf (i.p.m.); milanec = 2855 Af, resp. 23,47 Rf (i.p.m.)].⁸⁵ Přejato z Röselova slovníku *Wörterbuch zu den tschechischen Schriften des J. A. Comenius* (Rösel 1983, s. 85).

Uvedené slovníky E. Merker a H. Rösela představují dvě rozdílné strategie v pojetí autorského výkladového slovníku v redukované podobě. Současně na těchto příkladech vidíme, že to, co jsme označili pojmem *redukováný*, se může vztahovat k různým oblastem slovníku; v prvním případě se jedná o redukci v rozsahu excerpovaného materiálu, druhý

85 Český národní korpus – SYN2010. Ústav Českého národního korpusu FF UK, Praha 2010. Dostupné z WWW: <<http://www.korpus.cz>>.

případ představuje redukci v koncepci výkladového aparátu, který závisí na celkovém pojetí slovníku.⁸⁶

1.3.1.3 Výkladové slovníky (v neredukované podobě)

Autorské slovníky pochopitelně nevznikaly pouze v angloamerickém prostředí.⁸⁷ V tuto chvíli bychom se rádi věnovali několika zahraničním slovníkům, které rovněž představují typ výkladového slovníku, avšak na rozdíl od předchozího disponují rozsáhlejším a podrobnějším interpretačním aparátem, který se nesoustředí pouze na základní sémantickou charakteristiku lexikální jednotky, ale podává širší informace například z oblasti pragmatiky, etymologie nebo historicko-kulturního kontextu. Zároveň se jedná o takové práce, jež zpracovávají kompletní autorský korpus. Z hlediska rozsahu materiálu i náročnosti lexikografického zpracování není možné, aby byly realizovány individuálně, jak jsme se s tím mohli setkat například u slovníku E. Merker, ale vyžadují spolupráci a koordinaci týmu badatelů a odborníků. Realizace tohoto typu slovníku náleží k nejnáročnějším.

Od druhé poloviny 20. století vznikly v Polsku čtyři reprezentativní autorské slovníky, které zahrnujeme do oblasti velkých výkladových slovníků, tzv. thesaurů. Prvním z nich (řazeno chronologicky) je *Słownik języka Adama Mickiewicza* (1962). Sestavila jej skupina badatelů pod vedením Konrada Górskeho. V tomto případě se již setkáváme se slovníkem, jehož těžiště spočívá nejen v excerpci celého autorského lexikonu, ale současně v důkladném komentování a vysvětlení významů a funkcí slovních jednotek s ohledem na jejich kontext.

Jak uvádějí autoři předmluvy, historie tohoto slovníku sahá do konce 19. století (Górski 1962, s. VII), ovšem s jeho realizací se započalo až po druhé světové válce. Koncepce slovníku se během let měnila; zejména docházelo k debatě nad tím, jakým způsobem budou excerpovaná data reflektována v samotném slovníku. Konkrétně se jednalo o to, zdali budou vyznačeny veškeré kontexty daného hesla (lemmatu), nebo pouze kontexty vybrané, a to na základě dostatečné reprezentativnosti pojednávaného jevu.

86 „Daher reifte in mir bald der Entschluß, die bis heute bekannten und zugänglichen tschechischen Werke des Comenius zuverlässig zu exzerpieren, die Lemmata zu definieren und zu übersetzen und jedes Stichwort mit einem Beleg genau zu beschreiben. Das Postulat war also: Vollständigkeit des tschechischen Wortschatzes des Comenius zu erreichen.“ (Rösel 1983, s. V)

[„Proto ve mně brzy uzrálo rozhodnutí spolehlivě exerpovat do dnešní doby známé a přístupné české dílo J. A. Komenského, definovat a přeložit lemmata a každé heslo vysvětlit na jeho dokladech. Postulátem bylo: dosáhnout úplné slovní zásoby českých děl J. A. Komenského.“]

87 O situaci v jiných zemích (Francie, Itálie, Rusko atd.), o kterých v této publikaci není pojednáno, podrobně referuje například Christa Dill (1959, s. 342–344).

Ve výsledku obsahuje slovník veškeré citace, a to bez další segmentace, která by oddělila výskyt jazykové jednotky v próze, poezii a v jiných textových oblastech. *Słownik języka Adama Mickiewicza* excerpuje veškerý dostupný textový materiál autora, tzn. kromě textů uměleckých i Mickiewiczovy poznámky, zápisky atd. (Ibidem, s. X, XI). Jedná se tedy o jednotlivý korpus složený z různorodého textového materiálu, který není ve slovníku dále vnitřně členěn; slovy autorů předmluvy: „W obrębie hasła nie będzie podziału na słownictwo dzieł artystycznych i słownictwo pism prozaicznych. [...] Każdy wyraz hasłowy będzie ilustrowany wszystkimi wypadkami jego użycia.“⁸⁸ (Ibidem, s. XI)

Způsob, jakým jsou zpracována samotná hesla, je založen na komentování a interpretaci sémanticko-funkčního charakteru lexika. Zde ovšem platí nezbytný předpoklad, že tato část hesla musí být koncipována s ohledem na zdrojový materiál, což se odráží nejen v kvalitě komentáře, ale i v jeho rozsahu, který – pokud to nezbytně nevyžaduje situace (viz dále, příklad z Goethova thesauru) – by měl být minimální, stručný a věcný s důrazem na hlavní a v základě slovníku koncepčně definované významy a způsoby užití daného lexika.

AKTOR (7) rz. m. ◇ lp. B. aktora (1), lm. M. aktory (2), aktorowie (2), C. aktorom (1), B. aktorów (1) ◇

1. *Artysta sceniczny* (5): Inni rycerze Litewscy i Polscy momentalnie występują na scenie i nikną, nie iak aktorowie dramatu, ale raczej iak cienie czarnozięzkiey latarki Jag 133–5. [s porcelany wydęte osoby] jakoby aktory na scenie, | Zdawały się przedstawiać jakoweś zdarzenie; PT XII 43–4. Odyniec patrzył na zwierzę jak na aktora L₂ 429.

Przenośnie (2): Wszyscy odeszli; widze i aktory. D₃ *Przegląd* 364. publiczność potępiła sędziów jako aktorów niezręcznych O duchu 149–50.

2. *Bohater utworu literackiego* (2): Aktorowie jego są ludzie zwyczajni Karp 75. Prawda, że [...] Karpiński [...] bardzo często niepotrzebnie uświęcone zwyczajem imiona, ale teraz ledwie nie śmieszne, Filonów i Menalków aktorom swoim nadawał. Karp 87–91.

Ukázka 12: Heslo kromě základního tvaru (lemmatu) obsahuje jednotlivé gramatické formy, které se v korpusu vyskytují. U každého tvaru je uvedena jeho základní (absolutní) frekvence. V případě, že je slova použito v různých významových modifikacích, je každá skupina klasifikována ve zvláštní subkategorii. Přejato z práce *Słownik języka Adama Mickiewicza* (Ibidem, s. 37).

Jak je z uvedené ukázky zřejmé, k lemmatu *herec* (aktor) je v závorce uvedena jeho absolutní frekvence (*Af*), která vyjadřuje celkový počet všech lexikálních tvarů v celém korpusu. Následuje výčet všech gramatických tvarů spolu s jejich frekvencí a gramatickou specifikací; u lemmatu *herec* je tedy uvedeno, že se jedná o substantivum a u jednotlivých

88 „V rámci hesla se nerealizuje členění na slovník uměleckých a neuměleckých děl. [...] Každé heslo bude ilustrováno veškerými příklady jeho užití.“

tvář je určen pád. Následně je heslo opatřeno dvěma základními skupinami. Každá z nich představuje sémantickou kategorii, do které lexikum náleží. Jinými slovy se jedná o způsob jeho rámcového významového užití v korpusu. V daném případě je výrazu *herec* (a jeho tvarů) užito ve standardním významu „scénický umělec“ a současně (kromě významu přeneseného, tj. v tomto případě metonymického⁸⁹) ve významu „hlavní hrdina (postava) literárního díla“. Dominantním strukturním činitelem v uspořádání hesla je sémantická klasifikace, které je podřízena oblast gramatická a frekvenční (která ji doplňuje) spolu s implicitní informací týkající se žánrového či druhového rozlišení textu (proza vs. verš), jak je ostatně patrné i na uvedené ukázce, kde jsou kombinovány příklady z primárně neliterárních (ve smyslu ne-uměleckých) a literárních textů.⁹⁰ Součástí slovníku je také chronologické uspořádání všech textů A. Mickiewicze, což přispívá nejen k jejich přehledné databázi, ale k rámcové představě o využití a distribuci jednotlivých lexikálních jednotek. Tato oblast, na rozdíl od jinak koncipovaných autorských slovníků, však není s vlastním slovníkovým korpusem systematictěji propojena.

Druhou publikací, kterou zde stručně představíme, je *Słownik języka Jana Chryzostoma Paska* (1965), na jehož realizaci se podílel kolektiv odborníků z oddělení historie polského jazyka na Polské akademii věd (Polska Akademia Nauk)⁹¹ pod vedením Haliny Konecznej. Odbornými lektory byli například Zenon Klemensiewicz nebo Stefan Hrabec, kteří jsou rovněž spojeni s výše jmenovaným slovníkem A. Mickiewicze.⁹² Vedle mapování lexikální situace si autoři slovníku J. Ch. Paska kladli za cíl především přispět k poznání dobového jazykového stadia a jeho vývojové situace,⁹³ což v uvedeném případě představuje lexikum a jazykový systém 17. století. Ačkoli je Paskův jazyk, resp. jazyk literárního díla přirozenou součástí daného národního jazyka, je třeba mít stále na paměti, že se jedná o jeho dílčí oblast, která je ve výsledku definována řadou různých kritérií, kromě jiných např. i funkčně-sémantickými a pragmatickými aspekty. Jinými slovy,

89 Metonymický vztah v tomto případě vzniká mezi slovy *soudce* a *herec*; jedná se o tuto pasáž: „publiczność potępiła sędziów jako aktorów niezręcznych“ [„publikum zatratilo soudce jako nešikovné herce“].

90 První ukázka pochází z textu nazvaného *Uwagi nad Jagelonidą* (Jag), druhá je z veršovaného eposu *Pan Tadeusz* (PT) a třetí je z korespondence z let 1832–1844, kde Mickiewicz hovoří o polském básníku Františkovi Karpińském (Karp).

91 Jedná se o následující autory: Danuta Lankiewicz, Barbara Rykiel, Krystyna Siekierska, Teresa Sokołowska, Kazimierz Zelazko.

92 Stefan Hrabec lektoroval rovněž *Słownik języka Adama Mickiewicza*. Kromě jiných se na slovníku jako recenzent a konzultant podílel i významný český bohemista a rusista Jaroslav Bauer.

93 „Historia tego, co bywa metaforycznie nazywane życiem wyrazów, to historia sposobów ich używania, posługiwania się nimi przez ludzi, o tym zaś możemy się dowiadywać badając język autorów i zestawiając uzyskiwane w ten sposób wiadomości. Znaczy to, że podbudową wszelkich słowników historycznych powinny być możliwości najdokładniejszej i wyczerpującej opracowania słownictwa poszczególnych autorów.“ (Doroszewski 1965, s. VI–VII)

[„Historie toho, co bývá metaforicky nazýváno životem slov, je historií způsobů jejich užívání a zacházení, o tom se lze dozvědět na základě výzkumu autorského jazyka v porovnávání získaných vědomostí. Znamená to, že základem veškerých historických slovníků by mělo být nejdůkladnější a vyčerpávající zpracování slovní zásoby jednotlivých autorů.“]

v rámci stratifikačně rozvrstveného národního jazyka se musejí kriticky reflektovat i jednotlivé jeho oblasti, což v případě umělecké literatury představuje obezřetný a kritický přístup při aplikaci závěrů získaných z analýzy literárních uměleckých textů na povahu celostního jazykového systému.⁹⁴ Nemusí totiž vždy platit, že literární jazyk je spolehlivým reprezentantem národního jazyka. Literární jazyk je jazykem svého druhu. Mezi jednotlivými funkčními oblastmi jazyka sice platí korelační vztahy, ty jsou ovšem dynamicko-funkční povahy. Jak jsme sledovali u Roelcka (viz Roelcke 1994), dílčí idiolekty autor jednoduše nepokládá za pouhou formální součást celku národního jazyka, ale chápe je jako samostatné, do značné míry autonomní jazykové oblasti, které strukturně formují útvary vyššího řádu. Totéž platí i pro umělecký jazyk. Proto je značně problematické, chápeme-li jednu z oblastí národního jazyka, v tomto případě literární jazyk, jako reprezentativního zástupce komplexní jazykové situace dané doby. Do jaké míry by mohl být uvedený postup zavádějící, ukazují například texty Vladislava Vančury, jejichž lexikum i syntax jsou z estetických důvodů záměrně archaické. Analogická situace může nastat i v opačném případě, kdy autor cíleně užívá neobvyklých výrazů (neologismů) nebo příznakově deformuje syntax či vyšší textové struktury.⁹⁵ Konečně lze říci, že literární jazyk, a to i tehdy, pokud se na první pohled zdá být bezpříznakový, ve skutečnosti nikdy bezpříznakový není. Vyvozovat z něj generalizované závěry týkající se národního jazyka jako celku může být mylné, a to zejména proto, že – jak prokázala především moderní literární věda – je součástí imanentního vývojového systému, který i když netvoří izolovanou množinu, vyvíjí se po specifických historických trajektoriích. To ovšem není případ jmenovaného slovníku.⁹⁶ V dané souvislosti si lze rovněž uvědomit, co pro lexicografii (a v důsledku nejen pro ni) znamená moderní korpusová lingvistika a strojové

94 Zde je například nutné počítat s důkladnou znalostí vývoje literárních forem, se kterými úzce souvisí další strukturně-sémantické aspekty, např. otázka významu, lexikálních, syntaktických syntagmat apod.

95 V případě starších textů jsou to ovšem většinou pouze tyto texty, ze kterých lze čerpat poznání o jazykové situaci dané doby. Je rovněž třeba konstatovat, že funkční (žánrová) divergence ve starších fázích literárního vývoje nebyla natolik stratifikovaná, jako je tomu dnes, a proto i užití jazykových prostředků bylo nivelizovanější a jednotnější, což na druhé straně opět neznámá, že lze bezvýhradně nahrazovat poznání celku národního jazyka jazykovou situací konkrétních uměleckých textů.

96 „Oddając do druku *Słownik języka Jana Chryzostoma Paska* chcielibyśmy jeszcze raz podkreślić, że jest on jednocześniej reprezentatywny dla jednej z odmian języka ogólnego wieku XVII, uzupełnia zatem w znacznej mierze dotychczasowe zbiory leksykograficzne pochodzące z nieco wcześniejszego okresu, przede wszystkim *Słownik Knapiusza*.” (Skorupka – Szlifersztejnowa 1965, s. XXI)

[„Vrátíme-li se k vydání *Slovníku jazyka Jana Chryzostoma Paska*, rádi bychom ještě jednou zdůraznili, že se současně jedná o reprezentativní slovník jedné z verzí národního jazyka 17. století, který doplňuje do značné míry předchozí lexicografické sbírky pocházející z poněkud staršího období, především *Slovník Knapského*.”] Jedná se o trojjazyčný slovník Řehoře Knapského *Thesaurus polonolatinograecus seu promptuarium linguae Latinae et Graecae*, který vznikl v roce 1621 v krakovské tiskárně Francisca Caesaria.

zpracování jazykové materie,⁹⁷ které nabízejí nejen možnost funkční segmentace národního jazyka, ale i prostředky a nástroje k objektivnímu zkoumání a srovnávání jeho částí.

DEPUTACYJA (3) rz ż 1. «wydelegowanie kogo dla załatwienia jakiejś sprawy publicznej; delegacja; poselstwo» (2): Wysyłaia tam tedy mnie na Deputacyią a nay bardziey respektęm Ięzyka Łacinskiego bo tam lada chłop polacinie mowi a po Niemiecku rzadko kto. 62v–63r. Ia tez na owey Deputacyey byłem y nigdy się nieproznowało. 68 r.
2. «sprawowanie funkcji deputowanego, posła, posłowanie» (1): Margrabia Pinchowski w Lublinie na Deputacyey umarl. 278v–279r. (*łac.* deputatio)

formy: lp. B. *deputacyią* 62 v; Ms. *deputacyjej* 68r, 279r.

Ukázka 13: Heslo *deputace* je uvedeno ve dvou významových kontextech. V prvním z nich je pojmem *deputace* označeno poselství, zatímco ve druhém případě se jedná o výsledek zprávy (zde konkrétně jejího účinku). Přejato z práce *Słownik języka Jana Chryzostoma Paska* (Koneczna 1965, s. 135).

Způsob strukturace hesla Paskova slovníku je na první pohled analogický s předchozím slovníkem. Autoři ostatně odkazují k úzu, jenž platí pro slovník Mickiewiczův. Jediným rozdílem, který je ovšem spíše formální modifikací než zásadní metodologickou změnou, je koncové uvedení gramatických forem v nové polštině a etymologie tam, kde se jedná o výrazy, jejichž základ je cizojazyčný (latinský, francouzský apod.). Základní rozvržení opět respektuje sémanticko-funkční zřetel,⁹⁸ součástí hesla je i základní frekvence distribuovaná dle významových oblastí.

Jedním z dosud nejmladších polských autorských výkladových slovníků je *Słownik polszczyzny Jana Kochanowskiego* (1994), který pod vedením Mariana Kucaly sestavili jazykovědci z Institutu polského jazyka Polské akademie věd.⁹⁹ Podle slov M. Kucaly byla jednou z hlavních motivací snaha segmentovat z korpusu textů staré polštiny (16. století)¹⁰⁰ tu část, která náleží do oblasti individuálního autorského stylu (autorské řeči), a dále úsilí dokladovat jednotlivé výrazy dostatečným kontextem a přesnými citačními odkazy (Kucala 1994, s. V). Na první pohled se může zdát, že jde především o praktické hledisko, ale v záměru slovníku se odráží také obecnější metodologické stanovisko, které

⁹⁷ Viz pozn. 24.

⁹⁸ První způsob je užitím výrazu ve smyslu funkce osoby, jejího úkolu, poslání, kdežto druhý význam je výsledkem této činnosti, tzn. zprávou.

⁹⁹ Jedná se o následující autory: Zofia Bukowcowa, Krystyna Data, Dorota Janeczka, Krystyna Kajtochowa, Urszula Poprawska, Bożena Sieradzka.

¹⁰⁰ Kucala jmenuje *Słownik polszczyzny 16 wieku* (1966–1992).

vyjádřil ve své práci například T. Roelcke.¹⁰¹ Jedná se o směřování od celku národního jazyka k jeho stratifikačním segmentům, což v tomto případě představuje jazyk autora. Na rozdíl od předchozích dvou slovníků se zde nejedná o úplnou excerpici celé tvorby Kochanovského, ale pouze o jeho polské texty.¹⁰² Velice zajímavý byl rovněž vývoj slovníku z hlediska zpracování hesel. M. Kucala v předmluvě uvádí, že původním záměrem bylo demonstrovat jedinečný způsob autorského zacházení s jazykem, zejména pak v oblasti pojmenování, což se mělo promítnout přímo do koncepce hesla: „Chodziłoby o pokazanie tego, jak Kochanowski ukazywał dane desygnaty, jak opisywał różne zjawiska świata fizycznego i duchowego. W cytatach nie musiałyby występować wyrazy hasłowe, ale bliskie synonimy, zaimki i inne wskaźniki pojęciowe.“¹⁰³ (Ibidem, s. VI) Realizace takového typu výkladového slovníku by představovala nejen velice náročnou práci, ale současně i poměrně snadno napadnutelný výsledek; slovník by se totiž zakládal na zvolené interpretační strategii, která by organizovala celý lexikografický materiál. Uvedený postup by mohl být v první řadě kritizován z hlediska porušení objektivity. Výsledná podoba slovníku ovšem respektuje tradiční přístup, kdy jsou hesla doložena přímými citacemi, které umožňují jejich další studium především z hlediska kontextové analýzy.

ambicyja *chciwość bogactw 1, może też
zaszczytów, sławy 2;
a. jest zła 1, do kresu ludzi wiedzie 2 ~
r. -yja 1 ~ w p. 2.*

*1. IeŃli zła w oczasu wáŃzych (heretyków)
mŃza/ y proceŃsiia/|Aza lepszé łákomŃtwo/ y
tá ambiciia?| Fr 3,22,4. 2. Do tegóŃ kréfu y
ámbicia ludzi wiedzie: Wr 17.*

[...]

101 Viz studii T. Roelcka *Individualsprache und Autorenlexikographie (Einführung)*.

102 „Słownik nasz nie będzie słownikiem całej literackiej spuścizny Kochanowskiego – będzie to słownik jego dzieł polskich. Kochanowski, jak wiadomo, pisał także utwory w języku łacińskim. Trudno by jednak było łączyć w słowniku wyrazy polskie i łacińskie, cytaty polskie i łacińskie. Dla twórczości łacińskiej poety trzeba będzie zestawić słownik oddzielny, a może dodać dodatkowy tom – łaciński – do naszego słownika. Tekstów łacińskich Kochanowskiego jest o wiele mniej niż polskich, a więc o wiele mniej niż polskich jest łacińskich wyrazów i przykładów. Zmieszczą się one w jednym nie nazbyt obszernym tomie słownika.“ (Ibidem, s. V–VI)

[„Náš slovník nebude slovníkem celé Kochanovského literární pozůstalosti – bude to slovník jeho polských děl. Kochanovský, jak známo, psal také útvary v latinském jazyce. Bylo by totiž obtížné spojovat ve slovníku polské a latinské výrazy, polské a latinské citáty. Pro latinskou poezii bude třeba sestavit samostatný slovník, a může být dodatkovým oddílem – latinským – našeho slovníku. Latinských textů J. Kochanovského je mnohem méně než polských, a ještě méně než polských je latinských výrazů a citátů. Vejdou se do jednoho ne zbytečně rozsáhlého oddílu.“]

103 [„Šlo by o demonstrování toho, jak Kochanovský užíval konkrétních označení, jak popisoval různé jevy fyzického i duchovního světa. V citátech by nemusely být heslové výrazy, ale blízká synonyma, zájmena a jiné nociónální ukazatele.“]

arcybiskupi *należący do arcybiskupa 1, 2;*
 a. kamienica 1, koń 2 ~ w p. 1, 2.
 1. przyzło im sie miiąć prawie iuż przed
 kámienicą Arcybiskupią. Ap 7. 2. Książdz
 Krupłki chciał Arcybiskupá ochotnie minąć/
 á fzkápá iego ftánął/ y zátárt sie z koniem
 Arcybiskupim ták/ że sie żadnym obyczáiem
 nie dał odwieźdz/ Ap 7.

Ukázka 14: Přejato z práce *Słownik polszczyzny Jana Kochanowskiego* (Ibidem, s. 67, 76).

Jak je zřejmé z uvedené ukázky, předností slovníku je i to, že hesla jsou dokladována v transliterované podobě. Na rozdíl od předešlých dvou slovníků není v tomto případě sledována základní frekvence výrazu ani jeho etymologie. Součástí práce jsou rovněž přísloví, synonyma, básnické figury, tropy nebo rýmy (viz Ibidem, s. IX–XV). Z hlediska taxonomie, kterou provedl J. Mattausch, můžeme zde hovořit o tzv. *Werkgruppenwörterbuch*, který má účelově redukováný rozsah. Při přípravě slovníku byla využita děrnoštítková strojová metoda, která se ovšem při vlastní práci dle slov M. Kucały příliš neosvědčila (viz Ibidem, s. 7).

M. Kucała v předmluvě samozřejmě odkazuje ke zde již jmenovaným autorským slovníkům. V době, kdy byl promýšlen Kochanovského slovník, vznikal v polském prostředí jiný obdobný projekt, tentokrát věnovaný romantickému básníku Cyprianu Norwidovi. *Słownik języka Cypriana Norwida* (2009), jehož koncepce se utvářela od roku 1983 (viz Puzynina 1988, s. 1), byl realizován ve spolupráci Polské akademie věd a Varšavské univerzity pod redakčním vedením Jadwigy Puzyniny. Před vydáním vlastního slovníku vyšla v roce 1988 publikace *Słownik języka Cypriana Norwida: zeszyt próbny*, která je jakýmsi manuálem, v němž autoři vysvětlují celkovou koncepci slovníku. Mimo jiné je zde uveden vzor slovníkového hesla a podrobně je vysvětlena strategie jeho sestavení. Zaměření slovníku má podle Puzyniny „przede wszystkim pomoc w rozumieniu niejednokrotnie bardzo trudnych, „ciemnych“ tekstów Norwida. [...] Słownik ma również uwypatniać istotne cechy języka Norwida – przede wszystkim jego języka poetyckiego. Temu mają służyć dane statystyczne, podawane osobno dla poezji i prozy.“¹⁰⁴ (Ibidem, s. 2) Norwidův slovník má již elektronickou verzi, která nabízí nejen širší možnosti vyhledávání a třídění hesel, ale také základní souhrnné statistiky.¹⁰⁵ Hesla jsou opatřena základní gramatickou

104 [„především pomoci v porozumění nejednou velmi obtížným, „temným“ Norwidovým textům. [...] Slovník má také zvýraznit podstatné rysy Norwidova jazyka – především jeho jazyka básnického. Tomu mají sloužit statistická data, uváděná zvlášť pro poezii a prózu.“]

105 Dostupné z WWW: <<http://www.slownikjezykanorwida.uw.edu.pl/index.php>>

charakteristikou a údaji o frekvenci v závislosti na rámcové distribuci lexikální jednotky (užití v poezii a próze). Kromě jiného slovník zachycuje frazeologii a idiomatiku.¹⁰⁶

K jednomu z nejvýznamnějších výkladových autorských slovníků náleží německý *Goethe-Wörterbuch* (1966),¹⁰⁷ který byl zpracováván od 40. let 20. století na berlínském pracovišti Německé akademie věd. Postupně se ovšem projekt rozrostl do té míry, že na jeho realizaci pracovalo několik akademických pracovišť. V Berlíně excerpovali literární dílo, v Lipsku deníky, v Hamburku korespondenci a v Tübingen přírodovědecké spisy (Schadewaldt 1966, s. VI). První sešit slovníku tak vyšel pod společnou patronací berlínských, göttingenských a heidelberských akademie věd a jeho hlavními redaktory byli Wolfgang Schadewaldt, Werner Simon a Wilhelm Wissmann. Slovník byl od počátku prestižním projektem německé lingvistiky a lexikografie. První fáze zpracování probíhaly pomocí lístkového katalogu, který byl lemmaticky strukturován, takže každý tvar a k němu náležící ukázka byly pořádány pod své lemma. To je zřejmé i na struktuře slovníkového hesla, které zdůrazňuje významovou rovinu lexika. Kromě toho komentáře obsahují další informace týkající se způsobu (úzu), jakým Goethe slovo používal (zde například v kontextu určitého idiolektu), a doplňují jej o odkazy k oblastem jejich historického výskytu. Kromě sémantické a funkční modifikace výrazu závislé na kontextu užití ukazuje heslový komentář (viz Ukázku 15) i různé gramatické formy nebo konkurenční lexikální varianty.

106 Součástí projektu měly být i dílčí, speciálně zaměřené výstupy: „W *Zeszytach słownikowych* chcemy publikować opracowania grup tematycznych wyrazów, które uznajemy za ważne dla charakterystyki twórczości Norwida. [...] Obecnie mamy już zaplanowane zeszyt pierwszy, zawierający wybrane pojęcia etyczne (m. in. sumienie, prawda, kłamstwo, wolność, niewola) oraz zeszyt drugi dotyczący kolorystyki Norwida.“ (Ibidem, s. 3–4)

[„V *Sešitech slovníku* hodláme publikovat zpracování skupin tematických slov, která považujeme za důležitá z hlediska charakteristiky Norwidovy tvorby. [...] Celkově již máme naplánovány první sešit, obsahující vybrané etické pojmy (mj. svědomí, pravda, lež, svoboda, nesvoboda), a také druhý sešit týkající se barev u Norwida.“]

107 Rok v závorce je rokem vydání prvního sešitu. První z nich byl publikován v roce 1966 a zahrnoval hesla *A – abrufen*, druhý sešit byl vydán po roce a zahrnoval hesla *Abrufung – Ackerwand*, třetí sešit vyšel opět po roce a obsahoval hesla *Ackerwerk – Allgemeinheit* atd. Poslední desátý sešit vyšel v roce 1978 a obsahoval hesla *Ausführungsart – azurn*. Těchto prvních deset sešitů pak tvořilo první svazek celého slovníku čítající něco málo přes 1300 stran.

- Abend** Zur zeitlichen Erstreckung: gemäß süd- u westdt Sprachgebrauch läßt G den A. vielfach nur bis zum Einbruch der Dunkelheit (die für ihn ‚Nacht‘ ist, zB *achte Nachts* B2,305,8, Gfin Stolzb 20. [9.75]) sich erstrecken u bereits am Nachmittag beginnen (vergleichbar dem Gebrauch in den romanischen Sprachen), zB A-s in Duderstadt B3,203,1 ChStein 13. [12.77] A-s nach 4 Tgb 29.11.77 uö In Weimar paßt sich G, wie es scheint, dem nord- mitteldt Sprachgebrauch an: die Tagebücher der ersten Weimarer Zeit zeigen noch häufiges Vorkommen von ‚Nacht‘ neben A., während A. später dominiert (vgl zB auch ‚Jägers Nachtlied‘, in der späteren Fassung ‚Jägers Abendlied‘), ohne die alte Sprachgewohnheit ganz zu verdrängen, zB Zu Fuß bis Meiningen. Schöne Nacht. Um 9 Uhr angekommen Tgb 9.10.15 [es]waren wenige Stunden des Tags und der Nacht, wo er nicht in das Familienwesen, in die gesellige Tagesreihe mit eingriff 29,44,16 DuW 17 Die Tages-einteilung der Italiener entspricht demnach seinem eigenen Zeitsinn: In einem Lande wo man des Tags genießt, besonders aber sich des A-s freut, ist es höchst bedeutend wenn es Nacht wird...die Nacht schließt den A. und macht dem Tag ein Ende.

Ukázka 15: Mimo jiné se zde konstatuje, že dle deníkových záznamů Goethe ve svém prvním výmarském období užíval častěji slovo *Nacht* (noc) než *Abend* (večer), a to oproti výmarskému úzu. Ovšem později u Goetha převládalo slovo *Abend*, které současně používal pro označení té části dne, která se pojí se soumrakem, zatímco pokročilejší fázi večera označoval slovem *Nacht* (*achte Nachts* – osm hodin večer). Tato skutečnost, jak je uvedeno ve výkladu, se projevila i v přepsání původního názvu Goethovy básně *Jägers Nachtlied* (Lovcova noční píseň) na pozdější *Jägers Abendlied* (Lovcova večerní píseň). Pro snazší dohledávání a odkazování je každá strana rozdělena do osmi skupin po deseti řádcích. Přejato z prvního sešitu *Goethe Wörterbuch* z roku 1966 (Schadewaldt – Simon – Wissmann 1966, s. 24).

Jak lze sledovat na tomto příkladu, výklad hesla obsahuje kromě řady formálních aspektů také různé způsoby užití pojmu v jednotlivých Goethových textech a především interpretaci tohoto užití s ohledem na případnou modifikaci či genezi pojmu v autorském lexikonu. Důležitou úlohu zde tedy má změna v autorově způsobu psaní a používání slov. Goethův thesaurus zahrnuje kompletní soubor všech jeho textů, nejen tvorbu uměleckou, ale i odbornou. Jednotlivá hesla jsou detailně komentována, a to jak s ohledem na kontext, v jakém se vyskytují, tak vzhledem k autorské intenci nebo dobovému užití. Tuto skutečnost dokládá i J. Mattausch, který na konkrétních příkladech ukazuje, jak se modifikuje význam jednotlivých slov v závislosti na uvedených kritériích. Jasně se tu demonstruje metoda komentování sémantické vrstvy lexikálních jednotek, která spočívá v kritické revizi polyfunkční valence jednotky jak vzhledem k autorskému jazykovému systému, tak vzhledem k vnějšímu jazykovému systému. Důležitou úlohu při lexikografickém zpracování hesel sehrávají znalosti dobových reálií stejně jako nezbytné znalosti z různých vědních oborů. Obzvláště na příkladu Goethova slovníku se ukazuje, jak široký znalostní záběr předpokládá snaha objektivně a věcně referovat k významové stránce lexikálních jednotek a s jakou pečlivostí musí být sestaven vlastní heslový komentář.¹⁰⁸

Historii a charakteristice Goethova slovníku se ve studii *Textlexikographische Aspekte im Autorenwörterbuch – am Beispiel des Goethe-Wörterbuchs* detailně věnuje J. Mattausch.

¹⁰⁸ Viz rovněž příklady, o kterých hovoří M. Gräfe (Gräfe 1959, s. 304n).

Na tomto místě upozorníme na jeden zajímavý příklad z jeho studie. Jedná se o heslo *Encheiresis*.

Encheiresis *kunstreicher Handgriff, Verfahren; von der Natur: ihr Vermögen, (aus einzelnen Substanzen) Leben hervorzubringen* ob wir gleich gern der Natur ihre geheime E., wodurch sie Leben schafft und fördert, zugeben .. so kann der Mensch .. doch nicht von dem Versuche absteigen, das Unerforschliche so in die Enge zu treiben, bis er sich .. willig überwunden geben mag B49,210,26 Wackenroder 21.1.32 idVbdg 'E. naturae', im Munde Mephistos in boshafte-sinnverkehrender Anwendung des Ausdrucks¹ auf die wiss. Welterkenntnis iSv zergliedernde Behandlung der Natur ('naturae' = *genobj*) [Meph:] Wer will was Lebendiges erkennen und beschreiben, | Sucht erst den Geist heraus zu treiben, | Dann hat er die Theile in seiner Hand, | Fehlt leider! nur das geistige Band. | Encheiresin naturae nennt's die Chemie Faust I 1940 ~ *Urfaust* 372

¹ nach EOLippmann, GJb29,1908,164 übernommen von dem Chemiker u. Botaniker JRSpielmann (*Institutiones Chemiae*, 1763), dessen Vorlesungen an der Straßb. Med. Fak. G. besucht hatte

Cornelia Schulze

Ukázka 16: Přejato z elektronické verze *Goethe-Wörterbuch* (Dostupné z WWW: <<http://www.woerterbuchnetz.de/GWB?lemma=encheiresis>>).

Mattausch k heslu ve své práci uvádí: „Das Interpretament enthält neben der ausgehobenen Verbindung (a) den Hinweis auf den Textcharakter als Figurenrede, (b) eine Charakteristik der Aussagetendenz Mephistos (mit dem Klammerhinweis auf die gegenüber dem herkömmlichen Gebrauch veränderte Konstruktion), (c) die Angabe des speziellen Anwendungsbereichs und (d) die hier zutreffende Deutung des Ausdrucks. Eine zusätzliche Anmerkung verweist auf die Übernahme des Terminus von dem Chemiker und Botaniker J. R. Spielmann, dessen Vorlesungen an der Straßburger Medizinischen Fakultät Goethe besucht hatte, und gibt hierzu die entsprechende Literaturangabe.“¹⁰⁹ (Mattausch 1991, s. 722) Součástí komentáře (interpretamentu) je specifikace toho, v jakém kontextu a žánru se výraz u Goetha vyskytuje. První doklad je z korespondence, zatímco druhý se objevuje v Mefistofelově promluvě, kdy hovoří s žákem, který vstupuje do dialogu Mefistofela a Fausta s úmyslem najít v Mefistofelovi svého učitele. Mefistofeles s neskryvanou ironií komentuje a interpretuje marné lidské snažení usilující pochopit zákonitosti přírody. Ironie spočívá ve významovém posunutí pojmu, totiž v jeho reverzibilní konotační transformaci, kterou zakládá význam *vznikání* (*Encheiresis*) jako původně přirozený přírodní proces, jenž je nahrazen vědeckým, tedy lidským

¹⁰⁹ [„Interpretament vedle vybraného spojení (a) obsahuje odkaz na textový příznak hesla jako řečové figury, (b) charakteristiku Mefistofelova výroku (s ohledem na obvykle opačně používanou konstrukci), (c) údaj o zvláštním způsobu použití a (d) zde vhodný význam výrazu. Doplnující poznámka odkazuje na převzetí termínu od chemika a botanika J. R. Spielmann, jehož přednášky na štrasburské lékařské fakultě Goethe navštěvoval a odkud použil odpovídající literární údaj.“]

vysvětlením přírodních zákonitostí (pomocí chemie), což je z hlediska promlouvajícího (Mefistofela) projevem lidské nedokonalosti a omezeného vědění. Každá takováto snaha je totiž odsouzena pouze k povrchnímu nazírání na přírodu, které věda a racio odnímá její skutečnou podstatu.¹¹⁰ Zde můžeme pozorovat, co vše může komentář obsahovat; vedle nociónálních významů může být jeho náplní i charakteristika promluvy literárních postav stejně jako geneze pojmů. Mezi roky 2004–2009 byly zpřístupněny první čtyři oddíly Goethova slovníku rovněž elektronicky.¹¹¹

Obdobnou situaci z hlediska koncepce hesla sledujeme také u práce Christy Dill *Wörterbuch zu Goethes West-östlichem Divan* (1987). Dill se podílela i na výše uvedeném „velkém“ Goethově slovníku, což autorka explicitně uvádí v předmluvě svého slovníku: „Der Aufbau der Wörterbuchartikel folgt im wesentlichen den für das Goethe-Wörterbuch geltenden Regeln und Grundsätzen. Der Artikel kann eine Vorbemerkung haben, in der Besonderheiten der Schreibung, der grammatischen Form, der Wortbildung u. ä. vermerkt sind. Der Hauptteil ist nach Bedeutungen gegliedert.“¹¹² (Dill 1987, s. IX) Heslo má úvodní významový komentář, který jednak podává nociónální význam, jednak respektuje kontextovou sémantiku slova, včetně jeho sekundárních významů, např. metafor. Ačkoli slovník Christy Dill excerpuje pouze jedno dílo, a podle Mattauschovy taxonomie tak náleží do kategorie *Werkwörterbuch*, uvádíme jej v podkapitole výkladových slovníků s neredukovanou podobou na základě způsobu zpracování hesla, který jej metodologicky pojí s *Goethe-Wörterbuch*.

Návrh členění jednotlivých autorských slovníků v této práci není striktní taxonomií. U některých příkladů vidíme, že mohou náležet do více než jediné kategorie, což je situace nejen slovníku Ch. Dill, ale i řady jiných lexikografických prací. Důležitým kritériem je především zaměření slovníku a jeho funkce, které v souvislosti s povahou excerpuvaného materiálu mohou mít různě modifikovanou podobu.

110 Pro uvedenou pasáž jsme použili překlad Erika A. Saudka: „Cos živého poznat a popsati-li zkusíš, / dřív ducha z toho vyhnati musíš, / abys částice všechny v svých rukou měl; / jen ztratil se pohříchu duševní tmel. / Encheiresis naturae to v chemii sluje, / jež sama sebe tak bulíkuje.“ (Goethe 1979, s. 41)

111 Dostupné z WWW: <<http://gwb.uni-trier.de/de/>>

112 [„Výstavba hesla v podstatě sleduje důležitá pravidla a principy platné pro Goethův slovník. Heslo může poskytnout předběžnou charakteristiku, ve zvláštním případě je zaznamenána gramatická forma, slovotvorba apod. Hlavní oblast hesla je členěna podle významu.“]

1.3.1.4 Strojově zpracované slovníky

Od poloviny 20. století se strojová analýza jazyka prosazuje čím dál významněji, což se odrazilo i v oblasti autorské lexikografie.¹¹³ S tím souvisí důmyslnější propracování výpočetních algoritmů, které rozšiřují funkční užití zejména v oblasti statistických výpočtů. Ukázkou slovníku, který je cele zpracován strojovou metodou, může být již konkordanční slovník E. D. Bevana, který za pomoci počítače vygeneroval plnou konkordanci dramatických textů G. B. Shawa. Pokročilejší způsob využití strojové analýzy autorského lexika představuje *Word Index to James Joyce's Portrait of the Artist* (1967), který sestavil Leslie Hancock. Slovník se skládá z alfabetycky seřazených lexémů, ke kterým jsou připojeny číselné hodnoty, jež jsou výstupním formátem děroštitkové metody. Nejedná se tedy o slovník výkladový ani konkordanční v tradičním slova smyslu, ale o typ, jenž představuje rané vývojové stadium počítačem plně realizovaného autorského slovníku. Jakoli si lze obtížně představit, že s takovým slovníkem lze prakticky zacházet, představuje důležitou součást vývoje autorské lexikografie (a pochopitelně nejen jí) směřující k sofistikovanějším způsobům, jak strojově zpracovávat jazykovou materii. Hancock ostatně v předmluvě zdůraznil základní výhody strojově zpracovaného materiálu, se kterými se později setkáme v souvislosti s jazykovým korpusem (viz pozn. 24). K základním přednostem náleží především rychlost a úspornost ve vyhledávání konkrétních dotazů.¹¹⁴ Ale jak se ukáže teprve o něco později, není to pouze praktické hledisko, ale také vlastní přístup k jazyku, který proměňuje uvažování o jazyce i lingvistickou metodologii.

AGES -----	(12)	109/25
119/13	129/07	132/13
149/14	149/15	214/29
225/04	168/33	217/16
AGILE -----	(3)	090/32
160/16	240/06	
AGINCOURT -----	(1)	219/15
AGING -----	(1)	185/26

Ukázka 17: Výhodou slovníku je především spolehlivě zjiitelná frekvence jednotlivých slovních tvarů. Vlastní kontext je zde nahrazen numerickým kódem; při praktickém použití je tedy nezbytné si příslušnou pasáž vždy dohledat v pramenech. Přejato z Hancockova slovníku *Word Index to James Joyce's Portrait of the Artist* (Hancock 1967, s. 5).

113 Jedním z prvních iniciátorů zavedení počítačů do lexikografie byl podle Stephena Ramsaye italský jezuitský kněz Roberto Busa, „who in the late 1940s undertook the production of an automatically generated concordance to the works of St. Thomas Aquinas using a computer.“ (Ramsay 2008, s. 477)

[„který s pomocí počítače koncem roku 1940 automaticky vygeneroval konkordance prací sv. Tomáše Akvinského.“]

114 Kromě toho Hancock uvádí, že výhodou strojového programu, který byl původně vyvinut pro analýzu Joyceovy prózy, je možnost jeho dalšího použití při tvorbě jiných autorských slovníků (Hancock 1967, s. V).

Čtení numerického kódu se řídí následujícími pravidly: první číslo v závorce je hodnotou absolutní frekvence (*Af*), tj. všech výskytů daného tvaru v korpusu; první trojmístné číslo před lomítkem je paginace; následující dvojmístné číslo za lomítkem je určení řádku na stránce.

Jak je zřejmé ze struktury hesla, slovníky zpracované děrnoštítkovou metodou sice disponují automatizací při strojovém zpracování autorského lexika, avšak stále se jedná o takový typ výsledku, jehož povaha (na rozdíl od výkladových slovníků) spočívá v prostém statistickém konstatování základní frekvence. Navíc u Hancockova slovníku je jeho užití závislé na konkrétním vydání Joyceovy prózy, což jej poněkud limituje v jeho samostatném použití. Přesto je však právě tato práce dobrým příkladem toho, jakým směrem se strojová a kvantitativní analýza jazyka (nejen v autorské lexikografii) postupně vydala. V této fázi strojového zpracování se ještě u autorských slovníků nesetkáváme s pokročilými funkcemi statistických hodnot a výpočtů; o těch lze hovořit teprve v souvislosti s rozvojem moderní korpusové lingvistiky.

1.3.2 České autorské slovníky

V českém prostředí se diskuze o moderním autorském slovníku výrazně rozvíjí zejména od druhé poloviny 20. století. Dokládá to mimo jiné článek Oldřicha Králíka *Autorské slovníky*, který vyšel v roce 1961 v časopise *Slovo a slovesnost*. Králík zprvu stručně rekapituluje situaci v zahraničí a posléze *de facto* shrnuje to, o čem v této publikaci již byla nejednou řeč: „Typ, který se ustaluje v poválečných letech, usiluje postihnout v úplnosti jazykový odkaz národních klasiků a tím zvládnout závažnou část z vývoje národního jazyka. V SSSR vychází slovník Puškinův, v Polsku se chystá zvláště vyčerpávající slovník Mickiewiczův, na Ukrajině Ševčenkův, v Maďarsku Petöfiho a Juhásze, v Bulharsku vychází slovník Botevův, v Rumunsku se připravuje slovník Emineskův, v Německu se pokoušejí zvládnout obrovité úkoly slovníku Goethova.“ (Králík 1961, s. 239) Dále Králík konstatuje, že po stránce metodologické není u autorských slovníků „dosud ujasněn problém vztahu individuálního a obecného...“ (Ibidem, s. 239) Vedle tohoto vztahu Králík hovoří o dalších problémech, které se nezbytně pojí s každou koncepcí autorského slovníku: „Zvláštní problém je určení úseku textu tak, aby byl získán naprosto jasný kontext příslušného slova. [...] Základní problém je volba edice autora. Často se vychází z posledního akademického vydání, ale např. u Goetha bylo nutno sáhnout k nejúplnější dosud edici z minulého století, přičemž se respektovala vydavatelská zlepšení moderních edic.“ (Ibidem, s. 239) Jak vidíme, Králík zde formuluje poznatky, se kterými jsme se setkali nejen v článku Ch. Dill (Králík na ni ostatně odkazuje), ale o kterých jsme rovněž

na základě článku T. Roelcka prohlásili, že náleží k základním (obecným) metodologickým předpokladům lexikografické práce.

V kontextu české autorské lexikografie je zajímavé, že v minulosti podle Dill existoval návrh na realizaci Máchova slovníku, ke které ovšem nikdy nedošlo.¹¹⁵ Je možné se domnívat, že to byly právě tyto důvody, o kterých Ch. Dill hovoří, jež zabránily dalšímu pokračování na tomto projektu, tedy značná míra subjektivity Máchovy řeči. Na straně druhé, jak jsme viděli u řady autorských slovníků, je právě výjimečný charakter literárního jazykového projevu často důvodem pro sestavení samostatného (autorského) slovníku. V některých případech je dokonce slovník založen převážně na výběru takového lexika, které je příznakové, a nikoli obecně platné.

Máchův slovník zůstává nepochybně jednou z výzev, které stojí před českou (autorskou) lexikografií. Nacházíme především dva důvody, které by měly směřovat k realizaci takového projektu: Mácha podobně jako ostatní autoři, kterým byl ve světě věnován speciální autorský slovník, náleží k čelným reprezentantům národní literární tradice. Druhý důvod je ovšem mnohem závažnější, neboť se pojí s otázkou, která byla v minulosti vznesena ohledně Máchova autorství některých textů, mimo jiné např. prózy *Cikáni*. Těmito požadavky nemusejí být argumenty hovořící pro Máchův slovník zdaleka vyčerpány. Na tuto problematiku stručně odkážeme v závěru monografie, kde se pokusíme nastínit další možné perspektivy kvantitativně-korpusové analýzy literárních textů.

1.3.2.1 Strojově zpracované slovníky (frekvenční)

V českém prostředí existuje několik autorských, *de facto* frekvenčních slovníků,¹¹⁶ z nichž k nejstarším náleží *Konkordanční a frekvenční index k Slezským písním Petra Bezruče*, který zpracoval kolektiv pod vedením Jitky Štindlové¹¹⁷ v mechanografické laboratoři Ústavu pro jazyk český ČSAV (dnes ÚJČ AV ČR) v Praze roku 1969. Jak je zřejmé z názvu, kombinuje se zde typ indexu a frekvenčního slovníku, což má svoji přirozenou logiku. Index je schopen tvořit metodologický průsečík lexikostatistické metody a metody konkordanční. Navíc je tento způsob dán i typem strojového zpracování (děroštitkovou metodou), který jsme zaznamenali v případě Hancockova slovníku. Index (v daném případě však nikoli lemma) je reprezentantem všech identických slovních tvarů v korpusu,

115 Jedná se o jedinečný doklad takového projektu: „Das Macha-Wörterbuch, auf der Einzelexzerption der Werke Machas für das Wörterbuch der tschechischen Literatursprache basierend, steht vor dem Problem, daß der Stil Machas sehr individuell gefärbt ist.“ (Dill 1959, s. 345)

[„Máchův slovník, jenž je založen na excerpci Máchových děl pro slovník českého literárního jazyka, stojí před problémem Máchova stylu, který je příliš individuálně zabarven.“]

116 Jiný typ u nás nebyl realizován.

117 Spolupracovníky jsou: Viktor Fícek, Eva Macháčková, Olga Burdová a Míla Jiráková.

nositelem frekvence a současně je ve jmenovaném slovníku schopen subsumovat údaje o kontextové situovanosti tvaru pomocí strojového výpisu konkordance ve formátu sedmimístného čísla.

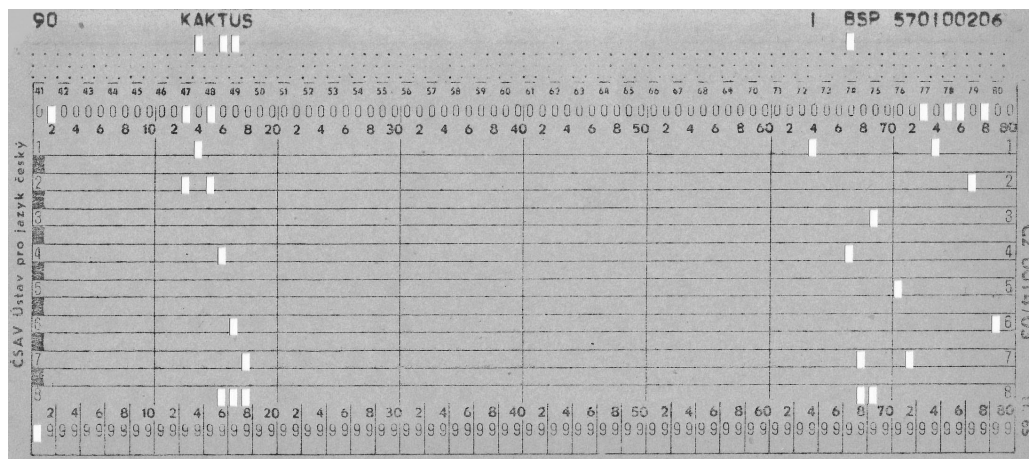
dunajské	1	4703509
duní		5208906
	2	5901906
duše		0101102
		2601505
		3902905
		5101502
		5701704
		7600403
	7	8301003

Ukázka 18: Přejato ze slovníku *Konkordanční a frekvenční index k Slezským písním Petra Bezruče* (Štindlová 1969, s. 25¹¹⁸).

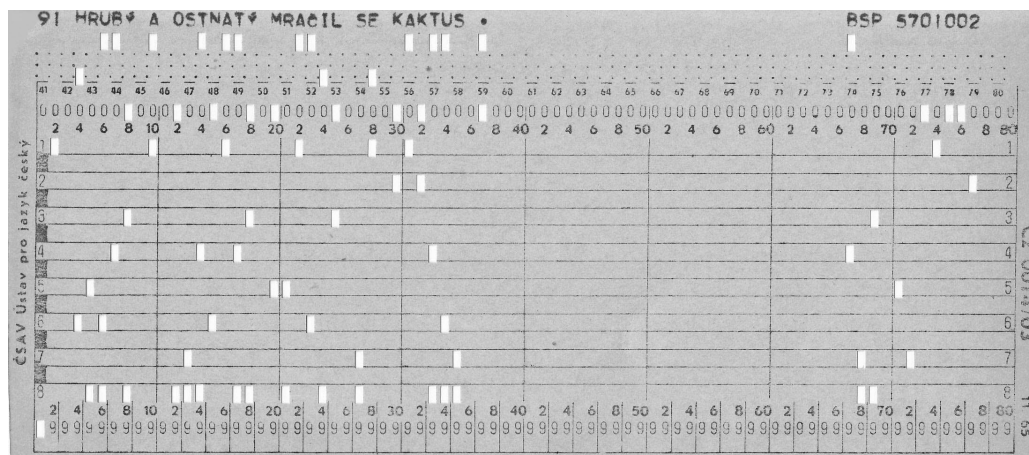
Struktura hesla je podobná situaci, s jakou jsme se setkali v Hancockově slovníku. Dalo by se říci, že se jedná o modifikovanou verzi, která respektuje povahu excerpovaného materiálu poezie. Z toho důvodu obsahuje numerický kód více relevantních informací, než je tomu v případě prozaického textu. Předsunuté číslo na začátku numerické řady indexu je opět základní frekvenční hodnotou, absolutní frekvencí (*Af*). Sedmimístná číselná řada se interpretuje následovně: „Tedy první dvě čísla označují básně od první do poslední (01–88). Další tři místa obsahují informaci o verši a poslední dvě o pozici slova ve verši.“ (Ibidem, s. 16) U námi uvedeného případu to znamená, že například tvar *duní* se objevuje v básních 52 (*Já I, II, III*) a 59 (*Horník*); v první básni se vyskytuje v 89. verši na 6. pozici, zatímco v druhé básni je to ve verši číslo 19 a rovněž na 6. pozici. Tento způsob lexikografického zpracování ukazuje na další aspekt (autorské) lexikografie, který spočívá ve zvoleném interpretačním způsobu, jenž je dán automatickým formátem. Autoři slovníku si jako výchozí zdroj vybrali vydání *Slezských písní* z roku 1957 a básně číslovali podle pořadí, v jakém jsou ve sbírce uvedeny. Oproti výkladovým slovníkům se tu zcela vytrácí sémantické hledisko. Na straně druhé je ovšem možné přesně stanovit základní statistické informace a tyto následně pořádat podle zvolených kritérií. Je možné si představit nejen abecední automatické třídění lexika, jak je tomu ve slovníku, ale například řazení dle frekvence (např. klesající) nebo seznamy lexika v jednotlivých básních (kvalitativně či opět frekvenčně uspořádané) a ty pak statisticky srovnávat; možné by bylo i detailněji statisticky analyzovat každou báseň na základě lexika, jeho celkové frekvence a pozice ve verši. Kontext je zde vlastně nahrazen číselnou řadou, avšak počítačové

118 Paginace odkazuje k oddílu Konkordanční index.

zpracování lexikonu v daném případě umožňuje dodatečný strojový výpis lexikálního okolí zjišťovaného slova.



Obrazek 1a: Děrný štítek (slovní štítek) s vyražením slova z Bezručovy básně Červený květ. Přejato ze slovníku *Konkordanční a frekvenční index k Slezským písním Petra Bezruče* (Ibidem, s. 19).



Obrazek 1b: Děrný štítek (textový štítek) s vyražením verše z Bezručovy básně Červený květ. Přejato ze slovníku *Konkordanční a frekvenční index k Slezským písním Petra Bezruče* (Ibidem, s. 18).

Frekvenčním slovníkem je také práce Petra Holmana *Frequenzwörterbuch zum lyrischen Werk von Otokar Březina* (1993). Dle slov autora, „základním posláním tohoto slovníku je podat **celkový obraz úplné slovní zásoby Březinova básnického díla**, tedy všech veršů, které byly básníkem zařazeny do definitivního vydání jeho díla, *Tajemnými dálkami* počínaje a dvěma posmrtně objevenými básněmi (*Za všechno díky*, *Návrat*) konče. Důležitým cílem slovníku je také, aby poskytl všechny číselné údaje (absolutní a relativní frekvence, distribuční koeficienty, údaje o disperzi atd.) potřebné pro další kvantitativní

výzkum Březinova jazyka, ale širě i jazyka devadesátých let a jazyka obecně“ (Holman 1988, s. 43). Holman ve jmenované studii podrobně charakterizuje jednotlivé lexikostatistické údaje, které jsou zajímavé především svým koncepčním rozvržením. Jak vidíme na Ukázce 19, je možné frekvenční autorský slovník navrhovat v různých funkčních modifikacích. Tradiční způsob představuje vyjádření frekvence lemmat, popř. slovních tvarů (tokenů), a to jak v základní, tzv. absolutní frekvenci (Af), tak v hodnotách relativní frekvence (Rf). Absolutní frekvence udává celkový výskyt všech tvarů lemmatu v korpusu, zatímco relativní frekvence vyjadřuje například procentuální zastoupení jednotky v určitém korpusu nebo se může jednat o normovanou frekvenci, která umožňuje srovnávat frekvenční údaje z nesteréjně velkých datových množin (korpusů). Relativní frekvence tedy může být vyjádřením pouhého procentuálního výskytu jednotky v určené množině nebo komparativní normalizovanou hodnotou. Vedle frekvencí lemmat či lexikálních tvarů je například možné sledovat frekvenci jednotky v kontextu, a to opět jak absolutní, tak relativní. Podobně je možné počítat s pravděpodobností výskytu určitého lexému v definovaných množinách-korpusech nebo specifikovat a analyzovat kontextové okolí lexému. Holman ve svém slovníku hovoří rovněž o disperzi, tj. míře rozptylu jednotky (lemmatu) v definované množině textů.¹¹⁹ Specifičnost kompozice Březinova slovníku spočívá v desetinném třídění; po každých deseti lemmatech, distribuovaných dle hodnot relativní frekvence do jednotlivých Březinových sbírek, je uvedeno deset statistických hodnot, které jsou sumačním vyjádřením vztahů mezi lemmaty, tzn. jsou zde součty Af , Rf , procentuální vyjádření frekvencí nebo kumulativní hodnoty Af , Rf . Takto je řešena jedna z hlavních částí slovníku, jiná potom obsahuje řazení lexikálních jednotek podle klesající Af , disperze a dalších parametrů. Na rozdíl od lineárního nepřerušovaného řazení disponuje tato strategie systematictější kategorizací a přehledností. Jinými slovy každá množina o deseti lemmatech zvláště je lexikostatisticky klasifikována jako celek, což při vyhledávání konkrétních lemmat ve slovníku umožňuje tyto dále klasifikačně zpřesnit a zpřehlednit pomocí jejich příslušnosti do příslušné desítkové množiny.

¹¹⁹ V ČNK je místo disperze zavedena tzv. průměrná redukováná frekvence (ARF – average reduced frequency), která stanovuje rozložení lexému v definovaném korpusu.

P	R	L	FR	FA DS DB	TD	SZ	VP	SCH	R	Z	D
21	21	NAD/NADE	144	178/6/74	33	11	19	37	52	26	80,96
22	22	SLUNCE	143	170/6/80	9	22	35	36	34	34	83,98
23	23	SVĚTLO	138	179/6/84	32	22	60	36	18	11	76,94
24		ZRAK	138	156/6/88	20	30	39	25	23	19	88,27
25	24	KVĚT	132	144/6/87	28	18	29	25	18	26	91,93
26	25	MY	123	160/6/47	4	20	22	38	33	43	77,04
27	26	JÁ	115	199/6/50	77	76	28	8	3	7	57,54
28	27	S/SE	109	122/6/71	24	25	15	17	27	14	89,35
29	28	I	105	118/6/52	20	18	23	16	29	12	89,00
30	29	DEN	104	130/6/76	40	27	26	16	10	11	80,14

+FR10: 1251 -FR10: 611 %FR10: 5,20 K FR: 9297 %FR: 38,62
+FA10: 1556 -FA10: 768 %FA10: 4,38 KFA: 10760 %FA: 30,34

Ukázka 19: Přejato z Holmanova slovníku *Frequenzwörterbuch zum lyrischen Werk von Otokar Březina* (Holman 1993, s. 124).

Moderními frekvenčními slovníky, které využívají širokou škálu lexikostatistických možností, jsou především *Slovník Karla Čapka* (2007) a *Slovník Bohumila Hrabala* (2009). Obě publikace vznikly na pracovišti Ústavu Českého národního korpusu pod vedením Františka Čermáka¹²⁰ a představují typ frekvenčního slovníku, který poskytuje velké množství formálních informací o jazyce autorů,¹²¹ jež se zakládají na moderních lexikostatistických metodách. Ve výsledku jsou ovšem slovníky zpracovány korpusovými metodami, což znamená, že kromě Af je možné zjišťovat i pokročilejší lexikostatistické výsledky, včetně různých kontextových relací (kolokací). Obzvláště cenné jsou shrnující interpretace jazykových vlastností autorů, které poskytují exaktní vhled do jejich charakteristiky, což se týká například slovních druhů, lexika, syntaxe, stylistiky, kolokace, frazeologie apod. Těchto údajů je využito nejen k vlastnímu popisu a analýze jednotlivých aspektů autorského jazyka, ale také ke komparaci; srovnávány jsou slovníky nejen mezi

120 Na *Slovníku Karla Čapka* se dále autorsky podíleli: Tomáš Bartoň, Renata Blatná, Václav Cvrček, Milena Hnátková, Jan Kocek, Marie Koptřivová, Michal Křen, Karel Kučera, Věra Schmiedtová, Martin Stluka, Michal Šulc, Pavel Vondříčka, Martina Waclawicová.

Na *Slovníku Bohumila Hrabala* se dále autorsky podíleli: Tomáš Bartoň, Václav Cvrček, Milena Hnátková, Jan Kocek, Marie Koptřivová, Michal Křen, Karel Kučera, Renata Novotná, Věra Schmiedtová, Martin Stluka, Dominika Šrajzerová, Michal Šulc a Martina Waclawicová.

121 Jak konstatuje František Čermák ve studii *An Author's Dictionary: The Case of Karel Čapek*, hlavním důvodem pro volbu tohoto autora byla moderní povaha jeho jazyka, která si dokázala udržet aktuálnost: „The choice fell on him not only because of his considerable reputation (having been widely translated, introducing, among other things, the word *robot* into the international vocabulary) but, primarily, because of his language, which is still, after some 60–80 years, surprisingly modern and exercising a huge influence on modern Czech language.” (Čermák 2008, s. 323)

[„Volba padla na něj nejen kvůli jeho značné pověsti (v minulosti byl široce překládán, mimo jiné přispěl slovem *robot* do mezinárodní slovní zásoby), ale především proto, že jeho jazyk je i po 60–80 letech stále překvapivě moderní a má velký vliv na moderní češtinu.“]

sebou, ale i s databázemi SYN2005, SYN2005:PUB nebo SYN2005:BEL.¹²² V obou případech navíc dochází k vnitřní segmentaci korpusu dle rámcové žánrové příslušnosti jednotlivých textů, což umožňuje srovnání mezi jeho jednotlivými částmi.

<i>Slovník Karla Čapka</i>	<i>Slovník Bohumila Hrabala</i>
próza	delší próza
	kratší próza
poezie	poezie
drama	
publicistika	publicistika
odborná literatura	
korespondence	

Tabulka 1: Žánrové dělení ve *Slovníku Karla Čapka* a *Slovníku Bohumila Hrabala*.

Toto členění je v první řadě motivováno lingvistickými (korpusovými) ohledy; jinými slovy, distribuce lemmat je zde řízena základní žánrově-druhovou příslušností textu nebo jeho rozsahem. Tato klasifikace je *de facto* logickým navazováním na obecný úzus struktury českého korpusu.¹²³ Z takto zaměřeného metodologického hlediska obě práce představují především zdroj pro lingvistické zkoumání. Snahou autorů je koncipovat autorský jazykový korpus tak, aby jej bylo možné využít pro sledování imanentních jevů každého z autorských lexikonů a jeho částí a současně umožnit přístup k širšímu funkčnímu srovnání, které by přesáhlo hranice těchto dílčích korpusů, tzn. především s celky ČNK.

¹²² Označení jednotlivých korpusů řady SYN, tj. synchronních korpusů.

¹²³ Struktura ČNK se řídí požadavkem reprezentativnosti, který v tomto případě vychází z funkčního hlediska, tj. funkční stratifikace národního jazyka (podrobněji k reprezentativnosti ČNK např. Šulc 2001; Křen 2013, s. 12–14). Konkrétně to znamená, že například v aktuálním SYN2010 jsou zastoupeny tyto skupiny textů: 40 % beletrie, 27 % odborná literatura a 33 % publicistika (Dostupné z WWW: <<http://ucnk.ff.cuni.cz/syn2010.php>>). Jiné kritérium reprezentativnosti zastává Centrum pro zpracování přirozeného jazyka, které disponuje vlastním korpusem národního jazyka. V tomto případě je reprezentativnost určena kvantitativně, tzn. rozsahem tohoto korpusu, který hromadně excerpuje především texty z internetu. Pro srovnání: SYN má aktuálně přes 2 miliardy slov, czTenTen12 přes 4 miliardy slov.

	celkem	delší próza	kratší próza	poezie	publicistika
koncertní	17	12	1	0	4
koncertovat	6	1	4	0	1
koncipovat	4	1	1	0	2
koncový+	4	3	1	0	0
končetina	2	1	0	0	1
končíci	12	5	6	1	0
• nikdy ³⁰ nekončíci					

Ukázka 20a: Abecední řazení hesel. Uvedeny jsou Af, což znesnadňuje srovnání mezi nesterjně velkými množinami. Kromě abecedního a frekvenčního řazení (viz Ukázku 20b) disponuje slovník seznamem hapaxů (slov s jediným výskytem) a zkratk nebo přehledovými statistickými tabulkami a grafy. (Pozn.: znaménko + charakterizuje kontextovou situovanost lemmatu, znaménko · vyznačuje typické kolokace lemmatu.) Přejato ze *Slovníku Bohumila Hrabala* (Čermák – Cvrček 2009, s. 138).

rank	heslo	frekvence
201	pán	803
202	křičet	803
203	vejít	801
204	vrátit	799
205	najednou	792
206	kam	785
207	jaký	783
208	docela	782

Ukázka 20b: Frekvenční řazení hesel. Přejato ze *Slovníku Bohumila Hrabala* (Ibidem, s. 520).

Kromě výše uvedených prací se ve fázi přípravy nachází autorský slovník Jaroslava Seiferta, na kterém se podílí pracovní kolektiv pod vedením Svatavy Machové. Slovník vzniká od roku 1997 na Pedagogické fakultě Univerzity Karlovy v Praze (viz Janovec 2003, s. 197), v současnosti se nachází stále ve stavu rozpracování. Některé dílčí výsledky prezentují studie S. Machové nebo Ladislava Janovce (viz bibliografii). Kromě projektů, které u nás nebyly realizovány,¹²⁴ je třeba zmínit některá dílčí zkoumání vztahující se k problematice autorského lexika a jeho frekvence, především práce Marie Těšitelové, jež jsou věnovány například slovní zásobě románu Karla Čapka *Život a dílo skladatele Foltýna* (1939). Jedná se o dvě studie publikované v časopise *Naše řeč* pod názvy: *Frekvence slov a tvarů ve spise Život a dílo skladatele Foltýna od Karla Čapka* (1948) a *Poznámky k slovní*

124 Svatava Machová ve své stati *Autorský slovník básnického díla Jaroslava Seiferta* (1998) zmiňuje některé dřívější koncepce autorských slovníků: „V České republice dosud projekt autorského slovníku realizován nebyl, nebereme-li v úvahu počáteční etapy počítačového popisu jazyka básnického díla Vítězslava Nezvala, který započal koncem 60. let na filozofické fakultě brněnské univerzity anglista a teoretik překladu Jiří Levý a který byl předčasně ukončen jeho smrtí. Realizovaly se pouze frekvenční slovníky jednotlivých děl některých autorů [autorka zde odkazuje k výzkumu M. Těšitelové a K. Kučery – pozn. R. Z.] nebo frekvenční slovníky vytvořené z náhodně vybraných vzorků textů daného autora.“ (Machová 1997, s. 31); kromě toho Machová ve svém textu uvádí, že na PedF UK je realizován (tedy ve druhé polovině 90. let 20. století) rovněž projekt nazvaný *Index slov a tvarů v rukopisech staročeské kroniky tak řečeného Dalimila* (Ibidem, s. 32); srov. též Janovec 1998, s. 64–65.

zásobě v románě Karla Čapka Život a dílo skladatele Foltýna (1955). Obě práce jsou historicky cenným dokladem zájmu české lingvistiky nejen o vlastní lexikostatistickou analýzu, ale též o interpretaci jazykové materie vycházející z kvantitativních metod spolu s úsilím přiblížit tyto exaktní postupy literární vědě.

Aktuálním projektem, který je věnován problematice korpusového zpracování autorského jazyka, je *Olomoucký korpus české esejistiky přelomu 19. a 20. století*,¹²⁵ ze kterého je zatím přístupný korpus esejů Otokara Březiny *Hudba pramenů* (1903); v přípravě je filozofická esej Ladislava Klímy *Svět jako vědomí a nic* (1904) a korpus českých biblických překladů.¹²⁶

~

Uvedené příklady autorských slovníků nejsou pochopitelně ani zdaleka vyčerpávajícím přehledem všech prací. Mnohé z nich, které náleží k důležitým a reprezentativním, jsme zde samostatně nepojednali; to se týká některých velkých thesaurů, mezi které náleží například Puškinův slovník,¹²⁷ Petöfiho či Ševčenkův slovník. Určitý rámcový přehled podávají i zde přetištěná schémata J. Mattausche, která tak v této publikaci plní dvojí úlohu; jsou nejen demonstrací Mattauschovy taxonomie, ale odkazují na řadu dalších důležitých autorských slovníků. O něco jiná je situace v českém prostředí, kde na rozdíl od zahraničí je možné tuto problematiku z hlediska existujících autorských slovníků *de facto* vyčerpat.¹²⁸ To je samozřejmě dáno malým množstvím takových prací, nikoli jejich náročností a způsobem zpracování.

Klasifikace autorských slovníků zde vychází z úzu, který se opírá o základní distinktivní rysy uvnitř této skupiny slovníků, jež tvoří spíše periferii lexikografie (srov. Hausmann 1989, s. 969). Jedná se o specifické členění, které slouží k popisu

¹²⁵ Podrobněji viz Pořízka – Schäfer 2010, s. 60–66.

¹²⁶ Dostupné z WWW: <<http://corpus.upol.cz/korpusy>>

¹²⁷ O přípravě a prvních výsledcích Puškinova slovníku ve své době informoval například Z. Tyl; na tomto místě vzhledem k dosavadnímu kontextu a způsobu výkladu ocitujeme pouze jeho věcnou část: „Slovník má obsáhnout celou lexikální zásobu Puškinova díla podle šestnáctisvazkového akademického vydání. [...] Jednotlivá hesla slovníku zpravidla mají tuto osnovu: záhlaví, úhrnný počet všech dokladů, význam slova (podle potřeby se stručným výkladem), počet dokladů s tímto významem a vybrané příklady (citáty). Následuje po př. další význam, počet jeho dokladů atd. Závěr hesla tvoří výčet všech míst, kde se dané slovo vyskytuje, seřazený podle osob, pádů apod. Zvláštní pozornost je věnována rčením a vazbám, charakteristickým pro Puškinovu dobu, vytčeny jsou i případy zvláštního užití slova (např. žertovný nebo ironický smysl), upozorňuje na metafory, personifikace, perifráze, slovní hříčky a j.“ (Tyl 1953, s. 80–80)

¹²⁸ Volně by do problematiky autorských slovníků mohla náležet i *Biblická konkordance* (1933). Jedná se o abecední seznam slov v bibli s odkazem na místo v textu a krátkým kontextem. Tato publikace má především praktický význam; slouží pro rychlé vyhledávání konkrétních slov v bibli (viz Daňhelka 1956). V současnosti je dostupná také elektronická biblická konkordance, která uvádí základní frekvenční údaje hledaného výrazu spolu s lexikálním okolím podobných slov, tj. slov se stejným lexikálním kořenem (Dostupné z WWW: <<http://www.fit.vutbr.cz/~michal/kr/>>>).

základních rozdílů mezi jednotlivými autorskými slovníky. Hausmann, jenž se rovněž vyslovil k typologii a charakteristice slovníků, konstatuje, že v základu každé lexikografické typologie se objevují dvě hlavní otázky: „Die Frage: Wozu dient das Wörterbuch? führt zu einer Funktionstypologie. Die Frage: Wie sieht das Wörterbuch aus? führt zu einer phänomenologischen Typologie des Wörterbuchs.“¹²⁹ (Ibidem, s. 970) První otázka směřuje do oblasti užité hodnoty slovníků, zatímco druhá se dotýká konkrétního způsobu jejich realizace. Jak jsme viděli, některé ze slovníků byly sestavovány za účelem reprezentovat jazykový stav příslušné doby, zvláštnosti lexikální zásoby autora, jiné zase usilovaly ukázat na kvantitativní situaci autorského lexika apod. Odtud se následně odvíjí i druhé kritérium, o kterém Hausmann hovoří, totiž způsob sestavení autorského slovníku, a v pozadí epistemologická hodnota lexikografie jako vědecké disciplíny. V průběhu výkladu jsme upozornili na takové slovníky, které podávají maximum informací o lexikální zásobě, a takové, které naopak inklinují ke strojově zpracovaným datům s důrazem na jejich úplnost a exaktnost lexikostatistického zpracování (kvantitativní interpretace). V současnosti je v tomto ohledu progresivní korpusová metoda, jež klade maximální důraz právě na exaktní postupy a materiálovou empirii, což má bezpochyby mnohé nesporné přednosti. Autorský slovník (a ve výsledku nejen tento lexikografický typ), který je založen na postupech a metodách uplatňovaných v korpusové lingvistice, vyžaduje uživatele obeznámeného nejen s korpusovými nástroji a vyhledáváním, ale rovněž s metodologií a možnostmi interpretace výstupních korpusových informací. V opačném případě tato data zůstávají nepřehlednou změť čísel, kategorií a grafů.

Přehled, který jsme podali v této části monografie, měl posloužit především k základní orientaci v problematice autorských slovníků. Jeho cílem nebyla detailní analýza konkrétních příkladů ani vyčerpávající seznam všech dostupných autorských slovníků, ale snaha vytvořit relevantní kontext, do kterého budeme moci s ohledem na jeho nároky umístit specifickou koncepci autorského korpusu a slovníku Jana Čepa.

129 [„Otázka, k čemu je slovník určen, vede k jeho funkční typologii. Otázka, jak má autorský slovník vypadat, směřuje k fenomenologické typologii slovníku.“]

2. Autorský korpus a slovník Jana Čepa: cíle, koncepce, způsob realizace a předpoklady využití v literárněvědném kontextu

2.1 Úvodem

Smyslem autorského slovníku je zmapovat jazykovou (obvykle lexikální) situaci v konkrétní a jasně vymezené oblasti, která je definována slovesnou tvorbou jednoho autora. Uzavřenost díla představuje významný aspekt, neboť autorský korpus tak může být excerptován v úplnosti, což má svůj vliv zejména na jeho reprezentativnost. V minulosti to byl především význam autorova díla, který byl jednou z důležitých motivací, jež vedla k sestavení takové práce. Většina autorských slovníků je v první řadě lingvistickými pracemi, které se soustředí na popis autorova jazyka, jenž je v některých případech chápán jako reprezentant dobové jazykové situace. To však ještě neznamená, že uplatnění autorských slovníků musí být výlučně jazykovědné. Zejména v široké oblasti kontextových vztahů a struktur mezi jazykovými a textovými prvky, které mají významný potenciál pro studium sémantiky, stejně jako obecně pro strukturaci lexikálních souvěsky, se nabízí jejich širší využití, které významně směřuje i do oblasti literárněvědného bádání.¹³⁰ Ačkoli lze se stávajícími autorskými slovníky omezeně pracovat i v literárněvědné oblasti, je takové jejich využití až sekundární a značně omezené. Jak vyplynulo z rámcového přehledu autorských slovníků, způsob, jakým jsou tyto práce koncipovány, výrazně nereflexuje možnosti adekvátního poznání uměleckého diskurzu z jiného než lingvistického stanoviska.¹³¹ Chybí zde totiž koncepce, která by odrážela nároky literární vědy, např. v povaze kategorizace či taxonomizace jazykového materiálu. Na druhé straně i korpus a slovník, který zde navrhujeme a realizujeme, se primárně zakládá na analýze jazykové báze (vrstvy) literárních textů, ovšem s tím rozdílem, že se ji snažíme uchopit a zpracovat tak, aby ve výsledku nejen poskytovala zdroj pro lingvistickou analýzu, ale byla relevantně využitelná také v kontextu literárněvědné práce. Jinými slovy, jazyková rovina je materiálovým fundamentem, který je v rámci návrhu Čepova slovníku

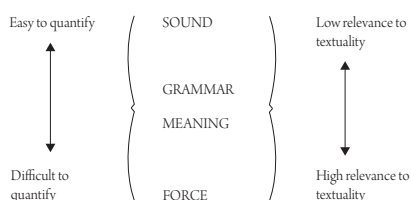
¹³⁰ V předchozí části jsme se seznámili s různými typy autorských slovníků, které odlišným způsobem přistupují k otázce kontextu. Stručně řečeno, kontext je možné reflektovat buď komentářem, nebo jeho formální, popř. statistickou analýzou.

¹³¹ To platí i pro oba nejnovější české autorské slovníky (*Slovník Karla Čapka* a *Slovník Bohumila Hrabala*), u kterých je sice zvolena žánrově-druhovú taxonomie, ovšem ta systematicky neakceptuje literárněvědné kategorie. Proto je i jejich užití v literární vědě omezené.

koncipován jako strukturně-funkční součást literárního textu (zde narativu) zapojující se do jeho vyšších strukturních rovin (vrstev). Naše pojetí vychází ze strukturálního přístupu k uměleckému dílu, který je chápáno jako konglomerát diferencovaných a funkčně provázaných strukturních úrovní, jež ovšem vytvářejí homogenní celek, strukturu uměleckého díla, pro kterou platí to, co formuloval již Jan Mukařovský: „Jsouce složkami jediné vyšší struktury, působí jednotlivé vývojové řady vždy na sebe navzájem; přitom ovšem každý zásah jedné řady do jiné uplatňuje se teprve prostřednictvím imanentního vývoje zasažené řady samé.“ (Mukařovský 1982, s. 504–505)¹³² Přirozeně ovšem nemůžeme počítat s tím, že analýza jazykové vrstvy nám poskytne komplexní odpověď na otázku strukturního uspořádání díla jako celku či jeho sémantiky; nicméně může – jak se domníváme – poskytnout jednu z nezanedbatelných informací o důležité strukturní vrstvě díla, kterou lze zkoumat a analyzovat relativně objektivními metodami, což má své důsledky pro analýzu vyšších vrstev a kategorií, které jsou mnohem obtížnější či zdali vůbec kvantitativně měřitelné.¹³³ Ve výsledku nám tedy půjde o podobu autorského korpusu a slovníku, který bude schopen na tyto problémy reagovat a který současně ponechá prostor pro samostatnou literárněvědnou interpretaci, jež ovšem bude schopna se opřít o poznatky objektivní kvantitativně-korpusové analýzy uměleckého díla. Bylo by nicméně velkým nedorozuměním, pokud by naše koncepce byla chápána ve smyslu substituce stávajících literárněvědných metodologií nebo samostatného interpretačního úsilí. V prvé řadě se jedná o lexikografickou práci, o autorský slovník, který je vždy důležitým nástrojem a nemůže v žádném případě nahradit vlastní analýzu a výklad. Na druhé straně jde však o více než slovník; smyslem naší koncepce není pouhé sestavení autorského lexikonu, ale především snaha funkčně aplikovat kvantitativní a korpusové metody uplatňované v lingvistice na oblast literárního díla s cílem podpořit (ve smyslu

132 Mukařovský tento svůj názor formuloval v souvislosti se strukturní povahou širšího socio-kulturního prostoru. V našem případě se sice jedná o uzavřenou oblast literárního díla, ovšem principy tohoto uvažování jsou platné na všech strukturních úrovních, tedy i v rámci organizace uměleckého díla a jeho složek, jež tvoří hierarchicky a funkčně uspořádaný celek.

133 Willie van Peer využívá ve svém článku argument o obtížně kvantifikovatelných vyšších složkách uměleckého textu ke kritice kvantitativní metody v literární vědě. Zdá se, že Peer plně nedoceňuje funkční a systémovou provázanost strukturních vrstev uměleckého textu, když kvantifikaci přisuzuje *de facto* schopnost referovat výhradně k jazykové oblasti. Na druhé straně má pravdu v tom, že ne všechny vrstvy uměleckého textu jsou adekvátně kvantifikovatelné:



(Peer 1989, s. 305)

exaktního přístupu) literárněvědnou argumentaci, která však sama o sobě zůstane autonomním metodologickým a analytickým výkonem. Konkrétně náš projekt směřuje do vybraných literárněvědných oblastí týkajících se vývojové perspektivy tvůrčího procesu, proměny tematiky, do oblasti textologie nebo (prozatím jen teoreticky) k některým z naratologických kategorií. Na základě systematického uspořádání jazykového materiálu, které by nepodléhalo pouze lingvistickému hledisku, ale rovněž (primárně) požadavkům literárněvědným, se lze dokonce pokusit o určitá generalizovaná tvrzení, která by se podobně, jako je tomu dnes v lingvistice, opírala nikoli o introspekci, ale o speciálně sestavený reprezentativní autorský korpus, jenž je schopen poskytnout objektivní informace o systémových vztazích, jejich modifikacích a proměnách.

Vlastní autorský slovník je *de facto* jedním z produktů kvantitativně a korpusově zpracovaných narativních uměleckých textů Jana Čepa. Základní hodnotou kvantitativně-korpusové analýzy realizované v této práci je důraz kladený na její výsledky, a zejména otázka směřující k její funkční kompatibilitě s oblastí literárněvědných problémů. Naším primárním cílem tedy není výhradně statický produkt lexikografie ve formě autorského slovníku, ale především zkoumání a testování možností funkčního propojení kvantitativně-korpusových postupů s nároky literárněvědné analýzy a interpretace.¹³⁴ Jsme přesvědčeni, že úkolem autorského korpusu a jeho formální analýzy není normativně vykládat umělecké dílo, ale poskytnout teoretickou i praktickou oporu pro jeho hlubší, systematictější a především exaktnější zkoumání. Tím nehodláme nikterak snižovat význam stávajících autorských slovníků. Je třeba si ovšem uvědomit, že současné postupy zakládající se na kvantitativních a statistických metodách, jimiž dnes disponuje moderní lingvistická analýza textu, umožňují nejen práci se sofistikovanými nástroji jazykové analýzy, ale otevírají cestu k novým konceptům a způsobům využití. To je ostatně příklad korpusové lingvistiky, která se v kontextu světové i české jazykovědy významně rozvíjí od 90. let minulého století a jejíž potenciál se dnes ukazuje být mnohem širší.

2.2 Základní metodologická východiska

Při realizaci autorského korpusu a slovníku je možné postupovat různými způsoby. Jak jsme ukázali v první části monografie, jako důležitý aspekt se v tomto ohledu jeví především cíl slovníku, který podmiňuje způsob jeho vypracování. Neméně rozhodující je povaha excerpovaného materiálu. Klademe-li si při zpracování a analýze autorského

¹³⁴ Proto i konkrétní výsledky, které jsou uveřejněny ve třetím oddíle této publikace, nejsou jediným možným způsobem, jak lze s Čepovým autorským korpusem pracovat.

jazyka jako hlavní podmínku maximálně možnou objektivitu, předpokládá to dnes především využití moderních kvantitativních metod, zejména pak práci s jazykovým korpusem. Co se týče materiálu, pracujeme v našem případě s narativními uměleckými texty, tedy s tzv. fikčními narativy. Kvantitativní výzkum jazyka automaticky neznamena, že je předem diskvalifikována otázka významu. Při postupu, kdy se vychází z kvantitativní analýzy, je ovšem nezbytné chápat problematiku významu jako determinovanou kvantitativními aspekty. V případě literárněvědné analýzy propojené s kvantitativně-korpusovým výzkumem to nikterak nezamezuje jejím vlastním analytickým postupům směřujícím k precizaci otázky významu. Důležitým momentem se kromě praktické využitelnosti formálních postupů v rámci literárněvědného bádání stává rovněž teoretická reflexe stávajících literárněvědných metod a výše jmenovaných postupů. Další okolností, která směřovala k volbě moderního autorského korpusu a jeho kvantitativní analýze, byla možnost jeho konfrontace s jinými (externími) korpusy.¹³⁵ V neposlední řadě jsme se pro moderní autorský korpus rozhodli proto, že dovoluje další rozšíření za hranice tvorby jednoho autora směrem k literárním skupinám, směrům apod.¹³⁶

2.3 Návrh a organizace autorského korpusu

Na úrovni celého projektu je třeba rozlišit dvě základní oblasti realizace. První spočívá ve vypracování komplexního korpusu a slovníku Jana Čepa tak, jako je tomu v případě *Slovníku Karla Čapka* a *Slovníku Bohumila Hrabala*. Obecně lze vycházet ze základního principu, který je uplatněn u obou těchto prací. Konkrétně to znamená excerpovat veškeré Čepovy texty, tedy kromě beletrie i esejistiku, korespondenci, rozhlasová pásma, publicistiku apod., a uspořádat tento materiál do žánrových kategorií, které se zakládají na literárněvědných taxonomiích. Druhou a aktuální částí projektu je užší korpus Jana Čepa – tzv. korpus fikčních narativů, který vznikl na základě naší snahy propojit literárněvědné požadavky s formální kvantitativně-korpusovou analýzou. Kompletní projekt tedy budou tvořit dva základní (sub)korpusy definované příslušnými soubory textů, jež lze vymezit na základě distinktivního příznaku dle jejich příslušnosti do množiny fikčních narativů, nebo non-fikčních narativů (+FIK/–FIK).¹³⁷ S ohledem na primárně

135 Jedná se především o srovnání s korpusem Karla Čapka nebo vybranými částmi SYN2010.

136 Viz Závěr.

137 Je třeba uvést, že toto dělení je motivováno metodologickým zaměřením slovníku. Ve skutečnosti je (nejen) u Čepa situace složitější; řada Čepových textů má přechodový charakter, tzn. z hlediska této distinkce (+FIK/–FIK) je nelze zcela jednoznačně definovat tak, aby bylo možné hovořit o „čistém“ žánru beletrie (fikce), nebo eseje (non-fikce). Z důvodů, které vycházejí ze zvoleného přístupu a metodologie, je ovšem nezbytné každý z textů určitým způsobem právě takto formálně taxonomizovat. Proto na nejvyšší úrovni volíme zmíněnou binární klasifikaci, která navíc umožní srovnávat

stanovené cíle, tj. možnost využití kvantitativně-korpusové analýzy v literární vědě, se budeme zabývat koncepcí korpusu fikčních narativů, který tvoří množina deseti povídkových knih Jana Čepa, jež vycházely od roku 1926 do roku 1953. Každý povídkový soubor bude současně dílčím subkorpusem. Ve výsledku tak bude korpus fikčních textů rozdělen na jednotlivé subkorporusy a seřazen chronologicky dle data prvního vydání každého povídkového souboru. Smysl tohoto řazení je dán dvěma základními kritérii: vydání povídkových souborů za Čepova života je zjevným aktem zveřejnění, na kterém se aktivně podílel sám autor; lze tedy na tento akt pohlížet jako na událost, ve které se projevuje vědomá autorova vůle (intence). Kvantitativně-korpusová analýza proto může reflektovat vývojovou křivku – tvůrčí genezi, kterou bude možné statisticky analyzovat a vyhodnocovat. Na straně druhé je třeba se vyrovnat se skutečností, že jmenovaná autorská intence se v daném případě projevuje i opačně. Čep totiž některé své práce modifikoval, dodatečně je upravoval a přepisoval. Jeho zásahy se netýkaly pouze izolovaných povídkových souborů, ale odrážely se také ve způsobu jejich propojení, což je příklad *Zeměžluče* (1931), jak o tom pojednáme dále. Druhý aspekt je čistě formální a vyplývá z uplatnění korpusové metodologie; jednotlivé povídkové soubory totiž představují uzavřené (úplné) subkorporusy, v rámci kterých je možné provádět statistická měření, což samozřejmě platí i z hlediska srovnávání mezi těmito subkorporusy nejen v rámci množiny fikčních narativů, ale v důsledku i v rámci celého autorského slovníku. (Pro konkrétní představu je možné uvažovat komparaci tematických oblastí mezi jednotlivými povídkovými knihami¹³⁸ nebo mezi žánrovými skupinami napříč celým slovníkem apod.)

Teoretický model se ovšem musí vyrovnat s reálnou situací, kterou představuje celé dílo Jana Čepa, jež bylo v minulosti nejednou různým způsobem uspořádáno a publikováno, a to dokonce i samotným autorem. Tvůrčí akt a autorská intence ve skutečnosti nepředstavují prostou kauzální vývojovou linii, ale většinou se realizují ve složitějších strukturních vazbách, které tvoří síť různým způsobem provázaných textů, jež vůči sobě mohou být v rozdílném kvalitativním poměru. Tyto skutečnosti nelze jednoduše eliminovat a ignorovat. Naopak je to jeden z důležitých požadavků, který od autorského korpusu očekáváme. Pro Čepa platí (podobně jako pro jiné autory, i když pro každého vlastním způsobem), že se ke svým literárním pracím opakovaně navracel a revidoval je, což se netýká jen konkrétních povídek, ale i celých souborů. V Čepově literární pozůstalosti najdeme také množství korektur a dalších textových úprav, které jsou nejen z textologického hlediska směrodatné pro interpretaci. Ačkoliv si lze představit realizaci

situaci mezi množinami textů s rozdílnou žánrovou příslušností. Tato komparace může naopak prokázat to, co jsme zde konstatovali, totiž míru provázanosti určitých žánrových skupin; současně může detekovat symptomatické oblasti autorského korpusu, kde se dané analogie vyskytují.

138 Viz třetí oddíl.

takového korpusu, který by obsahoval veškeré verze a různočtení, v první řadě je nezbytné vypracovat základní osnovu, jež bude schopna určitým způsobem reflektovat výše zmíněný tvůrčí proces a na niž bude dodatečně možné napojit jednotlivé textové varianty a další (sub)korpusy.

Situaci, o které hovoříme, si lze názorně demonstrovat na příkladu prvních tří Čepových knih, jejichž analýze je věnován třetí oddíl. Jedná se o *Dvojí domov* (1926), *Vigilie* (1928) a *Zeměžluč* (1931). Již třetí soubor obsahuje kromě nového oddílu pojmenovaného stejně jako celá kniha i první dva, ovšem ve změněné podobě. Druhá vydání, resp. verze *Dvojího domova* (1931) a *Vigilií* (1931) v rámci souboru *Zeměžluč* (1931) jsou oproti prvním značně redukována. Verze *Dvojího domova* z roku 1931 již neobsahuje povídky *Bouře*, *Smrt ševce Nerušila*, *Vzpoura*, *Dobyvatel*, *Justýnka*, *Mámení*, *Křepelka* a podobně je tomu i v případě *Vigilií*, kde byly eliminovány povídky *Na rozchodnou*, *Zbloudilý*, *Zelené jiskry* a *Epilog*. Z hlediska našeho záměru postihnout co nejširší oblast autorova jazyka bylo nezbytné zachovat původní znění těchto povídkových souborů dle prvních vydání. V případě jmenovaných prvních tří Čepových knih jsme ovšem do korpusu začlenili i druhé verze prvních dvou povídkových souborů, neboť jsou součástí autorem koncipovaného celku *Zeměžluč*. Tato strategie dovoluje postihnout nejen vlastní kauzální vývoj, ale rovněž komparovat různé verze *téhož* povídkového souboru a sledovat specifické zákonitosti konkrétní fáze tvůrčího procesu. Kromě toho budeme mít na základě prvních vydání k dispozici poměrně reprezentativní autorský korpus fikčních narativů, který bude založen na textové úplnosti jednotlivých povídkových knih. Tabulka 2 podává přehled o skladbě povídek v prvních třech jmenovaných knihách s ohledem na jejich první a druhé vydání. Z pohledu kvantitativně-korpusové analýzy nás v další části budou konkrétně zajímat vztahy mezi těmito třemi knihami, jež v rámci celé Čepovy tvorby představují jedinečnou situaci.

Dvojí domov (1926)	Vigilie (1928)	Zeměžluč (1931)		
		Dvojí domov	Vigilie	Zeměžluč
Domek	Na rozchodnou	Domek		Starcův smích
Bouře	Zbloudilý**			Hoře z lásky
Smrt ševce Nerušila	Rozárka Lukášová		Rozárka Lukášová	Veselá pohřební
Vzpouza	Husopas		Husopas	Můra
Dvojí domov	Čí bude vítězství***	Dvojí domov	Příčinnivá rodina	Starosvětská
Kozlovice	Zelené jiskry	Kozlovice		Lucie Laurová
Dobývateľ	Vigilie****		Vigilie	Albína Drůzová
Justýnka*	Epilog			Cesta na jitřní
Peněženka		Peněženka		Přízraky
Do města		Do města		
Mámení				
Křepelka				
Elegie		Elegie		

Tabulka 2: Tabulka zobrazuje rozdíly ve skladbě jednotlivých povídkových knih a jejich verzí. Pro účely korpusového a kvantitativního zpracování množiny fikčních narativů zakládáme jednotlivé (sub)korpusy (povídkové soubory) na maximálně relevantním rozsahu, tj. na prvních vydáních.

* Povídka *Justýnka* byla poprvé otištěna v roce 1926 v moravskoslezském deníku *Besedy*. Jak uvádí Mojmir Trávníček, jednalo se o jediný juvenilní text, který Čep zařadil do oficiálního prvního vydání *Dvojího domova* (Trávníček 1996, s. 24–25). Jak je ovšem z tabulky patrné, ve druhém vydání ji Čep vypustil.

** Původně pod názvem *V předvečer neznámé katastrofy*; povídka byla otištěna v časopise *Černá země* (Ibidem, s. 41).

*** Povídka *Čí bude vítězství* je od druhého vydání přetištěna pod názvem *Příčinnivá rodina*.

**** Původně pod názvem *Kříže u cest* (Ibidem, s. 41).

Základ autorského korpusu tedy tvoří původní povídkové soubory, jež obsahují i ty texty, které byly v pozdějších vydáních eliminovány. Spolu s tím bylo třeba vyřešit otázku, jež vyplývá z nezařazených textů, které Čep publikoval v průběhu 20. a 30. let minulého století časopisecky.¹³⁹ Řada těchto textů nebyla nikdy včleněna do jednotného souboru, jiné se naopak staly součástí pozdějších knižních vydání. Tato skupina textů je navíc žánrově diferencovaná. Najdeme zde jak povídky, tzn. fikční narativy, tak eseje,

¹³⁹ Srov.: „V průběhu pěti let – od října 1921 (kdy nastupoval do oktávy) do února 1926 (kdy začínal osmý a poslední semestr vysokoškolského studia) publikoval asi dvacet povídek a črt v *Moravskoslezských besedách* (později *Besedách*), beletristické příloze *Moravskoslezského deníku*, vycházejícího v Ostravě.“ (Trávníček 1996, s. 24)

recenze, zamyšlení, referáty, črty, impresy, překlady apod. Tuto množinu bylo nezbytné podrobněji přehlédnout a rozdělit příslušné žánrové oblasti. Mohli jsme se opírat o dva základní zdroje, které podávají přehled o jmenované Čepově tvorbě. Řadu důležitých informací bylo možné dohledat v Trávníčkově čepovské monografii (viz Trávníček 1996), kompletní seznamy Čepovy tvorby jsou naopak k dispozici v knize Tomáše Kubíčka (viz Kubíček 2014). Při rekonstrukci seznamu výhradně časopisecky publikovaných próz, který se stane nedílnou součástí autorského korpusu i frekvenčního slovníku, jsme vycházeli především z faktografických poznatků v těchto dvou pracích. Kromě toho bylo pochopitelně nezbytné některé problematické texty projít a rozhodnout o jejich vřazení do příslušného (sub)korpusu. Například povídka *Příchod*, kterou Čep publikuje v roce 1933 v *Listech pro umění a kritiku*, se později stala součástí románu *Hranice stínu* (1935). Podobně je tomu i s povídkou *Fůra v dešti* otištěnou v roce 1926 v *Cestě*. Ještě téhož roku je vydána pod názvem *Bouře* v prvním Čepově povídkovém souboru *Dvojí domov*. Analogická situace nastala i v případě textu *Proměny*, který vyšel opět v *Listech pro umění a kritiku* v roce 1935 a později byl autorem včleněn do knihy *Tvář pod pavučinou* (1941). Poněkud specifická je například situace u povídky *Justýnka*, která byla nejprve otištěna časopisecky, posléze se stala součástí první Čepovy knihy a v dalších vydáních byla autorem vyřazena.¹⁴⁰ Naopak texty jako *Žeh světla* (*Našinec*, 1940), *Poznamenaný* (*Akord*, 1941) nebo *Bohuslavice a Vřesina* (*Vyšehrad*, 1946), které vycházejí převážně v první polovině 40. let 20. století, jsme do subkorpusu časopiseckých próz nezařadili, neboť i přes užití ozvlášťujících literárních prostředků se celkově jedná o non-fikční texty či narativy.¹⁴¹ Ve výsledku však budou součástí jiných (sub)korpusů autorského slovníku, které obecně náleží do námi vymezené skupiny non-fikce (–FIK).

V tuto chvíli se budeme soustředit výhradně na texty, které lze zařadit do skupiny fikčních narativů (+FIK). Mojmír Trávníček řadu z nich, jež vyšly ve 20. a 30. letech, dodatečně publikoval v roce 2011 v souboru nazvaném podle jedné z Čepových povídek *Červený muškát*. Bohužel ani tento výbor nekoncentruje veškeré časopisecké povídky. Za přispění výše jmenovaných zdrojů a na základě další dodatečné (archivní) segmentace Čepových časopisecky publikovaných textů jsme vytvořili následující seznam Čepových próz (viz Tab. 3), které vyšly pouze časopisecky, tzn. nebyly nikdy publikovány v ucelených souborech za Čepova života. Jedná se o prózy, u kterých platí kritérium fikčnosti spolu s tím, že se nejedná o původní verze povídek později vydaných knižně.

140 Podobná situace nastala také u dalších textů původně zařazených do prvního vydání *Dvojího domova* a z pozdějších vydání vypuštěných (viz Trávníček 2011, s. 300).

141 Zejména v pozdější Čepově tvorbě se zřetelně projevuje prostupování mezi fikcí a esejistikou, což je i příklad některých časopiseckých textů, např. *Bohuslavice a Vřesina* (1946).

Z uvedených důvodů se tak vedle knižních souborů stanou separátní součástí korpusu fikčních narativů.

20. léta		30. léta
<i>Z vesnice</i> Moravskoslezské besedy III, č. 41–42, 1921	<i>Jindřich, Slávek a ona</i> Moravskoslezské besedy V, č. 4, 1923	<i>Cesta ze školy</i> Kolo, č. 1, 1933
<i>Kde Haná v hory přechází</i> Moravskoslezské besedy IV, č. 1, 1922	<i>Kde jsou lidé?</i> Moravskoslezské besedy V, č. 13, 1923	<i>Jakub Neoral</i> Kolo, č. 1, 1934
<i>Touha po středu, po srdci</i> Moravskoslezské besedy IV, č. 3, 1922	<i>Návrat</i> Moravskoslezské besedy VI, č. 11, 1924	<i>Páter Damián</i> Kolo, č. 3–4, 1934
<i>O pošetilé lásce</i> Moravskoslezské besedy IV, č. 10, 1922	<i>Táta Čtyrák</i> Moravskoslezský deník 27, příl. Besedy VI, 1924	<i>Zastavení</i> Almanach Kmene, 1934/35
<i>Anče</i> Moravskoslezské besedy IV, č. 20, 1922	<i>Dědina</i> Cesta VI, č. 50, 1924	<i>Sedláci v hospodě</i> Kolo, č. 1, 1935
<i>Dětská idyla</i> Moravskoslezské besedy IV, č. 22, 1922	<i>Hrob Janka Dospivy</i> Moravskoslezský deník 28, příl. Besedy VII, 1925	<i>Detektivní příběh</i> Hovory o knihách II, č. 1, 1938
<i>Hřeb v hrušce</i> Moravskoslezské besedy IV, č. 28, 1922	<i>Zmrzlá srdce</i> Cesta VII, č. 14, 1925	
<i>Kovář</i> Moravskoslezský deník, č. 22, 1922	<i>Opilec</i> Černá země I, č. 8, 1924/25	
<i>Oheň</i> Moravskoslezské besedy IV, č. 34, 1922	<i>De profundis</i> Cesta IX, č. 20, 1927	
<i>Píseň zahrádky</i> Moravskoslezské besedy IV, č. 34, 1922	<i>Oči</i> Cesta IX, č. 28, 1927	
<i>Karel Benda</i> Moravskoslezský deník, č. 22, 1922	<i>Putování úvozem</i> Tvar I, č. 4, 1927	
<i>Červený muškát</i> Moravskoslezské besedy IV, č. 38, 1922	<i>Noc filipojakubská</i> Tvar I, č. 7, 1927	
<i>Úzkost</i> Moravskoslezský deník, č. 25, 1922	<i>Holubí princezna</i> rukopis z r. 1927	
<i>Úpění v samotě</i> Moravskoslezské besedy IV, č. 47, 1922		
<i>Oddělené světy</i> Moravskoslezské besedy V, č. 2, 1923		

Tabulka 3: Čepovy časopisecké prózy, které nebyly vřazeny do knižního vydání. Texty jsou uspořádány do dvou oddílů, které se stanou základem pro oddělené (sub)korpusy, jež budou samostatnou součástí autorského korpusu fikčních narativů.

Z toho, co bylo dosud řečeno, je zřejmé, že Čepova tvorba se neodvíjela přímočaře kauzálně, tj. od vydání jedné knihy k vydání knihy další, ale že se jednalo o složitější kontinuum, několik rovin (prostoupených rovněž esejistikou a publicistikou), ze kterých se zformovaly ucelené povídkové knihy. Jestliže hodláme autorský slovník koncipovat co nejúplněji a současně se pokusit využít jeho způsobu nastavení co nejefektivněji pro literárněvědnou práci, musíme akceptovat i soubory „nezařazených“ textů a včlenit je do systému autorského korpusu. Z přehledu zde vyplývá, že časopisecké texty z 20. a 30. let, které Čep s výjimkou povídky *Justýnka* nikdy nezakomponoval do uceleného knižního souboru, pokrývají období let 1921–1938, tedy tu část Čepovy tvorby, ve které vycházejí knihy *Dvojí domov* (1926), *Vigilie* (1928), *Zeměžluč* (1931) a dále povídkové soubory *Letnice* (1932), *Děravý plášť* (1934), jediný román *Hranice stínu* (1935) a soubor *Modrá a zlatá* (1938). V tomto rozmezí, tj. mezi roky 1929–1931 vznikají také eseje, které později vyšly souhrnně pod názvem *Etudy pro paní J.* (1986). Po roce 1938 vycházejí Čepovy povídkové soubory *Tvář pod pavučinou* (1941) a *Polní tráva* (1946), jejíž součástí je oddíl *Lístky z alba*, který obsahuje povídky *Stínohra*, *Stará zahrada*, *Komůrka* a *Rodokmen*. Ovšem již v roce 1944 vyšla Čepovi jiná samostatná knížka nazvaná *Lístky z alba*, která na rozdíl od výše jmenovaného oddílu v *Polní trávě* zahrnuje navíc text *Tajemství našich setkání*. Jak je patrné, Čep do prvního vydání *Polní trávy* včlenil celý soubor *Lístky z alba* s výjimkou posledního textu. Zřejmě tak učinil proto, že tento text je žánrově odlišný od předchozích; nejedná se totiž o fikční narativ, ale o esej, zamyšlení či kontemplaci. Velmi pravděpodobně tak Čep uvedený text z důvodů žánrové koherence při výsledném koncipování *Polní trávy* vypustil. Jak je opět patrné, situace v Čepově tvůrčím procesu je složitější. Po autorově odchodu do emigrace v roce 1948 následuje poslední povídkový soubor *Cikání* (1953), který byl v pozdějším Trávníčkově vydání z roku 1999 doplněn o další dva původně samostatné texty (viz dále).

2.3.1 Relativní chronologie vzniku jednotlivých povídkových souborů

Uvedená kritéria a charakteristika Čepovy tvorby a jejích dílčích fází jsou důležité zejména proto, že výsledná podoba autorského korpusu a slovníku bude představovat lexikálněstatistické hodnoty, jež budou schopny tyto skutečnosti reflektovat formalizovaným způsobem. Abychom mohli jednotlivé hodnoty správně interpretovat, je vedle obeznámenosti s metodologií a postupy korpusové a kvantitativní lingvistiky rovněž důležité adekvátně srovnávat jednotlivé oblasti autorského korpusu. V rámci množiny fikčních narativů ve výsledku nepůjde výhradně o vztahy mezi jednotlivými povídkovými soubory, ale také o vztahy mezi knihami a časopiseckými texty a v neposlední řadě

o poměry mezi vývojovými fázemi vzniku jednotlivých povídkových knih.¹⁴² Proto je nutné stanovit relativní časový rozsah¹⁴³ zahrnující tvůrčí práci na jednotlivých povídkových souborech, jež samy o sobě budou v koncepci slovníku sloužit jako určité kotevní (faktické) body. Jinými slovy, je potřeba chápat tyto soubory povídek nikoli jako statické body, kterými se ve chvíli, kdy jsou transformovány do slovníku, svým způsobem nutně stanou, ale jako procesuální události v čase a prostoru. To je další z požadavků, který je kladen na koncepci a strukturu Čepova korpusu a slovníku, tedy pokusit se reflektovat otázku chronologie a relevantní intertextuální komparace. Problém musí být současně řešen tak, aby byl koherentní s metodologickými postupy formální lexikostatistické analýzy. S ohledem na literárněvědnou oblast, která disponuje vlastními metodologickými technikami, z uvedeného přirozeně vyplývá, že literárněvědná analýza jednotlivých kategorií a jejich vývojových tendencí v rámci Čepova díla bude prováděna sice samostatně a na základě jí vlastních metodologických kritérií, avšak na bázi materiálu, který bude zpracován a *de facto* již analyzován odlišnou technikou a postupy. Uvedené propojení umožní dosáhnout především určité míry objektivizace a zpřesnění poznatků, což se může pozitivně promítnout do výsledné literárněvědné interpretace. Sloučením těchto dvou dimenzí – kvantitativně-korpusové a literárněvědné analýzy – usilujeme také o obecnější teoretickou konfrontaci mezi doposud relativně samostatně se vyvíjejícími disciplínami – korpusovou lingvistikou a literární vědou. V této práci testujeme možnosti funkčního propojení obou sfér, nenabízíme nové teorie, pouze se pokoušíme o poněkud netypický přístup k literárním textům, jehož smyslem je na pozadí konkrétního projektu zkoumat potencialitu funkčního propojení výše jmenovaných disciplín, resp. jejich metodologických a analytických postupů a kritérií. Jestliže jsme výše hovořili o nezbytnosti reflektovat dynamickou povahu vývoje literární tvorby, potom platí, že na první pohled strohá fakta číselných lexikostatistických údajů nedokážou sama o sobě tuto skutečnost v úplnosti postihnout. Na straně druhé jednotlivé číselné hodnoty jako výstupní formát kvantitativně-korpusové analýzy musejí zůstat v maximální možné míře izolovány od externích zásahů interpretační povahy.¹⁴⁴ Primární výpovědní hodnota těchto informací musí vycházet z formálních matematicko-statistických základů specifikovaných pro danou oblast (lexikostatistiku). Teprve posléze v součinnosti s navrženou strategií slovníku a na základě propojení s adekvátními literárněvědnými postupy, které

142 Zcela analogická bude situace i tam, kde se budou srovnávat množiny fikčních narativů a žánrově odlišných textů, tj. esejů, korespondence, deníků apod.

143 Atributem *relativní* dáváme najevo, že naše rekonstrukce z pochopitelných a nezbytných důvodů představuje pouze časovou (formální) chronologii bez nároku na řadu aspektů, které se pojí s tvůrčí genezí, kam náleží například intertextuální vlivy, kulturně-dobový kontext apod.

144 Tato problematika vyplyne zejména v souvislosti s vypracováním jemnější taxonomie autorského korpusu, která bude reflektovat narativní promluvová pásma (viz kap. 2.3.3).

s kvantitativně-korpusovou analýzou vytvářejí koherentní metodologickou bázi, je bude možné smysluplně interpretovat.

Dříve nežli navrhneme a představíme samotnou makrostrukturu autorského korpusu a slovníku, bude nezbytné se podrobněji seznámit s chronologií Čepovy umělecké tvorby. Pro tyto účely se jako vhodná jeví zjištění, která předložil Mojmír Trávníček v publikaci *Pouť a vyhnanství: život a dílo Jana Čepa* (1996), kde se podrobně věnuje historii jednotlivých Čepových knih i textů. Z těchto poznatků budeme účelově vycházet, abychom získali faktický přehled o vzniku Čepových děl a mohli na tomto základě pro potřeby korpusu a slovníku stanovit vývojové etapy. Kromě Trávníčkovy monografie se pro nás stala neméně užitečnou i monografie Tomáše Kubíčka *Dvojí domov Jana Čepa* (2014).

~

V souvislosti s první Čepovou vydanou knihou povídek *Dvojí domov* (1926) Trávníček uvádí, že povídku *Justýnka*, která původně vyšla v *Besedách* v únoru roku 1926, Čep v přepracované podobě zahrnul do prvního vydání jmenovaného povídkového souboru (Trávníček 1996, s. 24). O ostatních prózách, jež Čep napsal před tímto datem a jejichž přehled podává Tabulka 3, Trávníček konstatuje: „Na dvacítce Čepových juvenilií lze zřetelně sledovat jak rychlý vývoj, tak i konstanty jeho tvorby. Zpočátku převládá postpubertální pesimismus, chmurná a dekadentní nálada, realistické pozorování detailů a expresivní výraz. [...] Nyní je Čepovo slovo ještě ztěžklé a kormutlivé, hrnouce před sebou a bezúspěšně absorbující nadbytek neurčitého zoufalství a beznaděje. A již zde zazní, zpravidla v prvních větách nebo v závěru, intonace známá z Čepova zralého díla. [...] Poetika titulů teprve tápavě hledá. Střídmost, melodie a významová síla, integrující text, charakteristické pro názvy Čepových povídek i knih, je v juveniliích výjimkou.“ (Ibidem, s. 25, 26) Mezi ranými Čepovými prózami a jeho první povídkovou knihou *Dvojí domov* sice Trávníček shledává řadu vazeb a anticipací, přesto chápe tuto část tvorby jako svým způsobem samostatnou vývojovou fázi předcházející Čepovy vyzrálé texty: „Čepovy povídky v Moravskoslezském deníku byly psány se syrovější bezprostředností, daleko méně stylizovány bez nadlehčení, které vytváří duchovní atmosféru *Dvojího domova*.“ (Ibidem, s. 40) Některé povídky ze souboru *Dvojí domov* Čep publikoval zvlášť v roce 1926. Konkrétně se to týká povídky *Dvojí domov*, jež vyšla ve druhém ročníku časopisu *Sever a východ*; ve 3. čísle osmého ročníku potom vyšla povídka *Fůra v dešti* (později ve *Dvojím domově* pod názvem *Bouře*). V časopise *Cesta: týdeník pro literaturu, život a umění* ve stejném roce Čep publikoval povídky *Kozlovice*, *Do města* a *Peněženka* (roč. 8). Ve druhém ročníku ostravského časopisu *Černá země* vyšla například úvodní povídka

celého prvního souboru nazvaná *Domek*. Z hlediska časové geneze *Dvojího domova* se tedy stále pohybujeme kolem roku 1926, kdy je obtížné jednoznačně určit přesnou dobu vzniku prvního z textů. Na druhé straně, jak konstatuje Trávníček, je *Dvojí domov* vyústěním předchozích Čepových beletristických snah.

To, co jsme právě konstatovali pro část rané Čepovy tvorby, resp. pro (sub)korpus textů časopisecky publikovaných ve 20. letech a povídkový soubor *Dvojí domov*, bude analogicky platit pro celou Čepovu tvorbu, samozřejmě v závislosti na reálných vztazích mezi jednotlivými tvůrčími fázemi, které reprezentují publikované knihy.

Co se týče druhé sbírky *Vigilie* (1928), Trávníček uvádí, že povídka *Na rozchodnou* byla napsána již v roce 1927 pod názvem *Kdo ví, zdali se sejde* (Ibidem, s. 47) a uveřejněna v devátém ročníku týdeníku *Cesty*. Ve stejném roce „dochází k prvnímu velkému, až horečnému tvůrčímu vzplanutí. Ve *Tvaru* publikuje *Noc filipojakubskou* (v září 1927), *Husopasa* (v prosinci), *Či bude vítězství* (v lednu 1928; do druhého vydání *Vigilií* zařazeno jako *Příčinnivá rodina*), *Starcův smích* (v březnu) a posléze *Hoře z lásky* (v září 1928). Jsou to plody zralého slovesného umění prózy, které někteří (např. Václav Renč) považují za nejčistší a nejlepší v celém Čepově díle“ (Ibidem, s. 48).

Posledními dvěma povídkami se vlastně již posunujeme k následující knize *Zeměžluč* (1931), která je sbírkou navazující na předchozí dvě knihy. Tuto skutečnost dokládá i to, že první dvě knihy jako samostatné dílčí celky autor v upravené podobě zahrnul do prvního vydání *Zeměžluče*, což se projevilo především v redukci původního rozsahu nebo v revizi, jako je tomu v případě povídky *Či bude vítězství*, která byla pro druhé (a další) vydání *Vigilií* přejmenována na *Příčinnivou rodinu*. Tyto skutečnosti, které odrážejí kontinuitu vývoje první Čepovy tvůrčí fáze vymezené třemi knihami próz, musejí být z kvantitativně-korpusového hlediska striktně formalizovány. Naším cílem ovšem bude uvedené relace formálně analyzovat a pokusit se prokázat, na základě jakých strukturních vztahů k sobě náleží a jaký princip zakládá jejich jednotu. Jestliže tedy formalizujeme Čepovy tvůrčí fáze do příslušných kategorií jednotlivých (sub)korpusů, činíme tak z důvodu nezbytné epistemologické koherence, kterou vyžadují postupy kvantitativní a korpusové jazykové analýzy. Jinými slovy, komparativním účelům založeným na konfrontacích statistických čísel musejí odpovídat i veškeré předpoklady spojené s budováním autorského korpusu a navrhováním modelu slovníku, jenž se musí zakládat na přísných kategoriálních stanoviscích; v opačném případě by finální výstupy neměly oporu v epistemologické jistotě, která je dána dodržáním metodologické koherence jednotlivých postupů při budování Čepova autorského korpusu. Případné interpretační konsekvence jsou v této fázi výzkumu nežádoucí, což ovšem přestane platit ve chvíli, kdy budou informace získané ze statistické analýzy dále zpracovávány v rámci literárněvědných analýz. Pokud se zde odvoláváme na Trávníčkovu hledisko, využíváme jej především pro potřebu faktografické

rekonstrukce Čepova díla. Jeho interpretační a evaluativní potenciál, objevuje-li se v námi citovaných odkazech, nemá v danou chvíli pro náš aktuální cíl relevantní hodnotu. Na straně druhé však bude možné jej zpětně konfrontovat s výsledky kvantitativní a korpusové lexikoanalýzy.

Vrátíme-li se zpět k souboru *Zeměžluč*, lze jeho genezi formálně vymezit rokem 1928 (viz první časopisecké publikování povídek *Starcův smích* a *Hoře z lásky*). V následujícím roce podle Trávníčka vznikají povídky *Veselá pohřební* a *Lucie Laurová* (Ibidem, s. 52) a o rok později povídky *Cesta na jitřní* a *Albína Drůzová* (Ibidem, s. 58).

Po vydání *Zeměžluči* Čep připravuje další povídkový soubor *Letnice* (1932). Rok před jeho vydáním Čep koncipoval text *Jakub Kratochvíl* (Ibidem, s. 62), povídku *Letnice* dokončuje v roce vydání knihy (Ibidem, s. 64).

Chronologii *Děravého pláště* (1934), v pořadí pátého Čepova povídkového souboru, Trávníček vysledoval do roku 1933, kdy Čep v časopise *Řád* publikuje prózu *Samomluva*, která se později stane součástí jmenované knihy (Ibidem, s. 65). Ve stejném roce, ve čtvrtém ročníku *Tvaru* vychází rovněž povídka *Noční návštěva*.

O jediném Čepově románu *Hranice stínu* (1935) v souvislosti s jeho chronologií Trávníček napsal: „Před Vánocemi [jedná se o rok 1934 – pozn. R. Z.] odevzdal Fučíkovi větší část rukopisu *Hranice stínu*, na němž intenzivně pracoval dlouhou dobu. Jeho koncepce se vyvíjela a měnila, jak naznačují ukázky definitivně nezařazených zlomků v *Kole*. Základní obrys, jak ho známe z podání Jana Franze a jak ho naznačoval v hovorách s Jitkou Fučíkovou, však nepochybně zůstal.“ (Ibidem, s. 79) Relativní chronologii tohoto románu tedy rekonstruuujeme do roku 1933, kdy Čepovi vycházejí některé prózy v *Kole* (viz Tab. 3).

Do vydání další povídkové knihy uplynuly tři roky, avšak již roku 1935 Čep podle Trávníčka napsal povídku *Ponocný* (Ibidem, s. 85), která se ve výsledku stala součástí knihy *Modrá a zlatá*, jež vyšla v roce 1938. Další texty, které se staly součástí jmenovaného souboru, Čep psal v roce 1937; jedná se o *Zatopenou ves* a *Zápisky Jiljího Klena* (Ibidem, s. 99–100).

Historii prvního z textů následující knihy *Tvář pod pavučinou* (1941) lze datovat již do roku 1935, kdy dle Trávníčkova zjištění Čep napsal drobnou prózu *Proměny* (Ibidem, s. 85). V roce 1937 otiskuje Čep v pátém ročníku časopisu *Listy pro umění a kritiku* povídku *Variant* a o rok později text *Příbuzenstvo* (Ibidem, s. 103). Následující rok potom vzniká povídka *Tvář pod pavučinou*, podle které byl pojmenován celý výbor: „V rozporu s dosud zachovávaným zvykem píše tentokrát na pokračování, což je v jeho praxi ojedinelé – dosud nikdy nedal z ruky nehotový rukopis. Tolik mu tentokrát záleží na tom, aby jeho hlas zazněl bez prodlení. *Tvář pod pavučinou* je jednou z odpovědí české literatury na 15. březen 1939 (první část vyšla 27. března) [...] Poslední prózu pro *Tvář* psal

v lednu a v únoru 1941: *Oldřicha Babora* (dvě ukázky vyšly v *Řádu* a v *Akordu*).“ (Ibidem, s. 107, 116)

Poslední Čepův povídkový soubor, který vyšel před autorovou emigrací do Francie v roce 1948, je *Polní tráva* (1946). Kniha obsahuje šest povídek (respektive devět, počítáme-li zvlášť prózy z oddílu *Lístky z alba*¹⁴⁵), z nichž k nejstarším (dobou svého vzniku) dle Trávníčka náleží *Stínohra*: „Píše postupně jednotlivé obrázky z *Lístků z alba*, v nichž kondenzoval všechnu svou zkušenost; nakonec přistupuje k rozlehlé próze *Polní tráva*, poslední ze série pěti velkých novel, v nichž rozvíjí stejný příběh, stejné podobenství. Jako první napsal *Stínohru*, která svou novou melodií překvapila v lednovém *Akordu* 1943.“ (Ibidem, s. 131) Čep ovšem již na jaře v roce 1942 publikoval v devátém ročníku *Akordu* fragment *Svatojánské pouti*. Počáteční chronologii *Polní trávy* tedy stanovujeme k tomuto roku.

Od roku 1948 pobýval Čep natrvalo ve francouzské emigraci a do Československa se již nikdy nevrátil. V Mnichově mu v roce 1953 vyšla knížka *Cikáni* a v roce 1956 Čep publikoval poslední beletristický text, povídku *Tři pocesťní*, otištěnou v témže roce v antologii Antonína Kratochvíla *Peníz exulantův* (Ibidem, s. 153–155), která vyšla nákladem Kulturního referátu Sdružení československých politických uprchlíků v Německu jako třetí svazek literární edice Kamenný erb. Mojmír Trávníček zařadil poslední Čepův text do souboru *Cikáni a jiné prózy*, který vyšel v nakladatelství Vyšehrad v roce 1999 jako poslední část výboru *Polní tráva* (třetího svazku *Spisů Jana Čepa*). I když původní vydání z roku 1956 tuto prózu neobsahuje, zahrnujeme ji do námi vytvořeného (sub)korpusu *Cikáni*. Důvody pro tento postup jsou především pragmatické: v opačném případě bychom museli akceptovat jeden izolovaný text, což by vzhledem ke způsobu, jakým jsou koncipovány jednotlivé subkorpora, (tj. na základě povídkových souborů), působilo nesystematicky a současně by to komplikovalo adekvátní kvantitativně-korpusovou analýzu, jejíž výsledky jsou relevantnější při aplikaci na rozsáhlejší textové (datové) množiny. Vřazení jmenované povídky do (sub)korpusu *Cikáni* je vhodné i z důvodů komparačních. V lexikostatistice bude vždy adekvátnější, budou-li se srovnávat větší celky než ojedinělé texty, i když ani takový postup není samozřejmě vyloučený. A konečně uvedené stanovisko podporuje skutečnost, že kniha *Cikáni*, jak uvádí Trávníček (Ibidem, s. 376), nevznikla na rozdíl od předchozích knih jako výsledek autorova záměru souborně uspořádat a vydat konkrétní texty. Ty sice vyšly souborně v roce 1956, ovšem k vydání je připravili pracovníci Sdružení politických uprchlíků v Mnichově. Rovněž Trávníček

145 *Lístky z alba* vznikly jako samostatný soubor próz, který byl vydán v roce 1944. První vydání obsahuje texty *Stínohra* (časopisecky v *Akordu*, roč. X., 1942–1943), *Stará zahrada* (časopisecky v *Akordu*, roč. X., 1942–1943), *Komůrka* (časopisecky v *Akordu*, roč. XI., 1943–1944), *Rodokmen* (časopisecky v *Akordu*, roč. XI., 1943–1944) a text *Tajemství našich setkání* (časopisecky v *Akordu*, roč. X., 1942–1943), který byl z *Polní trávy* vypuštěn.

svůj krok zařadit povídku *Květnové dni* (1954) a zmíněnou prózu *Tři pocestní* do vydání z roku 1999 zdůvodňuje ohledem na blízké umělecké postupy, které pojí výše zmíněné prózy s texty původního souboru *Cikáni*: „Konfrontujeme-li například skicu *Květnové dni* (podobně také *Cikáni*) s deníkovými záznamy, korespondencí a vzpomínkami Jana Čepa, objevíme minimální distanci mezi hrdiny příběhu a vypravěčem. Postavy příběhu *Tři pocestní* neváhal identifikovat sám, když v korespondenci s přáteli v šedesátých letech s uspokojením napsal, že „ti živí se v tom hned poznali“. Zůstává tuším nesporné, že vyjadřování prostřednictvím epických dějů, smyšlených příběhů a jejich figur se autorovi stávalo čím dál obtížnější a zřejmě i zbytečnější. Přirozenější, autentičtější mu bylo bezprostřední, osobní oslovení posluchače a čtenáře.“ (Ibidem, s. 376–377) Pro zařazení próz *Květnové dni* a *Tři pocestní* do jmenovaného (sub)korpusu by tedy kromě časové blízkosti mohly hovořit i Trávníčkovy argumenty týkající se analogie v tvůrčím postupu. Ve výsledku tedy zakládáme (sub)korpus *Cikáni* na souboru próz, jež tvoří první vydání, a rozšiřujeme je o dva zmíněné texty. Takový postup podporuje „kompatibilní“ časový horizont. Na druhé straně kvantitativní metody umožňují detailní prověření těch částí (sub)korpusu, které náleží k původnímu souboru *Cikáni*, a těch, jež sem byly dodatečně přiřazeny Trávníčkem.

Nejen v tomto případě ovšem nadále platí distinkce mezi tím, co označujeme termínem (sub)korpus, a reálným prozaickým souborem. Jinými slovy, existuje historické vydání souboru *Cikáni* z roku 1953 a (sub)korpus *Cikáni*, který navíc obsahuje výše uvedené prózy, podobně jako existuje soubor *Zeměžluč* a (sub)korpus *Zeměžluč – oddíl*, který obsahuje pouze třetí oddíl původního povídkového souboru,¹⁴⁶ neboť – jak jsme již konstatovali – druhá vydání *Dvojího domova* a *Vigilií* jako součásti *Zeměžluče* jsou z hlediska původních verzí obou knih značně redukována. (Sub)korpus *Polní tráva* zakládáme na prvním vydání této povídkové sbírky. Samostatnou pozornost zasluhuje oddíl *Lístky z alba*, který oproti svému prvnímu vydání (1944) neobsahuje text *Tajemství našich setkání*. Důvody plynou z toho, co jsme v souvislosti s tímto souborem uvedli výše (viz s. 84). Pojem (sub)korpus neobsahuje mnohé z atributů, které lze v rámci literárněvědného přístupu relevantně akceptovat. Především zde nelze operovat s vnějšími vlivy, které se spolupodílejí na recepčních strategiích, tzn. typ nakladatelství, formát a provedení knihy, její propagace, recenze apod.; jednoduše zde neexistuje celý soubor skutečností, které absorbuje fenomén knihy jako kulturního faktu v reálném historickém kontextu. Naopak (sub)korpus dovoluje takový pohled, který by byl u klasické knihy problematický. Ačkoliv je svým způsobem reprezentantem knihy, zastává ji velice specifickým způsobem, který účelově zdůrazňuje jazykovou složku jako materiálovou oblast

146 První dva tvoří samostatné (sub)korpusy *Dvojí domov* (1931) a *Vigilie* (1931).

1	2	3	4	5	6	7	8	9	10	11	12
Dvojí domov (1926)	Vigilie (1928)	Zeměžluč (1931)	Letnice (1932)	Děravý plášť (1934)	Hranice stínu (1935)	Modrá a zlatá (1938)	Tvář pod pavučinou (1941)	Polní tráva (1946)	Cikáni (1953)	časopisecké prózy z 20. let	časopisecké prózy z 30. let
1 Doměk	Na rozchodnou	Starcův smích	Letnice	Děravý plášť	Hranice stínu	Ponocný	Příbuzenstvo	Polní tráva	Cikáni	Z vesnice	Cesta ze školy
2 Bouře	Zbloudilý	Hoře z lásky	Jakub Kratochvíl	Člověk na silnici		Zápisky Jiřího Klana	Proměny	Staník	Květnové dni	Kde Haná v hory přechází	Jakub Neoral
3 Smrt ševce Nerušila	Rozárka Lukášová	Veselá pohřební	Nedělní odpoledne	Samomluva		Zatopená ves	Tvář pod pavučinou	Host	Tajemství kláry Berdové	Touha po středu, po sráči	Páter Damián
4 Vzpoua	Husopas	Můra		Pouť			Variant	Listky z alba*	Tři pocesní	O pošetilá lásce	Zastavení
5 Dvojí domov	Či bude vítězství	Starosvětská		Noční nověťova			Oldřich Babor	Setkání	Rozptýlení**	Aniče	Sedláci v hospodě
6 Kozlovice	Zelené jiskry	Lucie Laurová		Poslední Velkonoce				Svatolájská pouť		Dětská idyla	Detektivní příběh
7 Dobývateľ	Vigilie	Albina Drúzová								Hřeb v hrůšce	
8 Justýnka	Epilog	Cesta na jitřní								Kovář	
9 Peněženka		Přízraky								Oheň	
10 Do města										Píseň zahrádky	
11 Mámení										Karel Benda	
12 Křepelka										Červený muškát	
13 Elegie										Úzkost	
14										Upění v samotě	
15										Oddělené světy	
16										Jindřich, Slávek a ona	
17										Kde jsou lidé?	
18										Návrat	
19										Táta Čtyrak	
20										Dědina	
21										Hrob Janka Dospívá	
22										Zmizlá srdce	
23										Opilec	
24										De profundis	
25										Oči	
26										Putování úvozem	
27										Noc filipjakubská	
28										Holubi princezna	

Tabulka 4: Chronologický přehled prvních vydání Čepových povídkových knih a jejich základní makrostruktura.

* Oddíl obsahuje prózy *Stínohra*, *Stará zahrada*, *Komůrka*, *Rodokmen*.

** Prózu tvoří dvě části: *Před zavřenými dveřmi* a *Ostrov Ré*.

a strukturní vztahy mezi jejími (korpusově vydělenými) komponenty. Je důležité mít tento rozdíl na paměti, aby nedocházelo k nedorozumění či k očekáváním, jež jsou zcela za hranicemi a možnostmi konkrétních kvantitativně-korpusových analýz.

Relativní chronologie představuje námi koncipovaný model (Schéma 4), který slouží nejen k sestavení autorského korpusu a slovníku, ale který je současně základním orientačním rámcem sloužícím pro primární přehled a výklad statistických hodnot. Dříve nežli tento model představíme, prezentujeme přehledovou tabulku, jež obsahuje veškeré povídkové soubory i texty, které se staly základem pro vybudování jednotlivých (sub)korpusů. V této fázi se tedy jedná o základní množiny textů, které po náležitém mechanickém zpracování (elektronizace, lemmatizace, disambiguace a inkorporace do funkčního pole korpusového systému) budou napříště fungovat jako (sub)korpusy. Spolu se samostatnými, časopisecky publikovanými prózami z 20. a 30. let dostáváme relativně úplný a uspořádaný soupis beletristických textů Jana Čepa, který lze použít jako základ pro autorský korpus a slovník (viz Tab. 4).

Vedle beletrie je Čep autorem řady žánrově odlišných textů. Kromě esejů, korespondence nebo deníků se jedná o publicistiku, v jejímž rámci je třeba rozlišit mezi různými novinovými články (recenze, imprese, úvahy apod.) a souborem textů pro rozhlas Svobodná Evropa, ve kterém Čep po své emigraci každý týden referoval o nových knižních publikacích či uveřejňoval vlastní úvahy. Celek této Čepovy tvorby tvoří (z důvodů, o kterých jsme již hovořili) druhou část autorského korpusu, tj. množinu non-fikčních textů (–FIK). Do uvedené množiny ovšem nepatří veškeré Čepovy esejistické texty, neboť některé z nich jsou psány francouzsky a přeloženy byly až později po Čepově smrti. Korpus veškerých Čepových textů, které je nutné oddělit od beletrie, musí být rovněž strukturován a segmentován, a to ze zcela pochopitelných důvodů dle žánrových kritérií. Výsledkem jsou jednotlivé (sub)korpusy obsahující oblast Čepovy esejistiky, korespondence, deníkových záznamů či publicistiky. Některé z těchto textů byly zpracovány a publikovány, což platí například pro korespondenci, ze které vyšlo několik výborů, jiné (např. deníky) naopak stále zůstávají v Čepově literární pozůstalosti uložené v pražském Památníku národního písemnictví. Specifická je situace i v případě Čepových rozhlasových pořadů v rádiu Svobodná Evropa. Každý týden zde Čep komentoval konkrétní dílo nebo autora. K těmto pořadům se dochovaly Čepovy poznámky a rukopisy, které ovšem nelze automaticky zahrnout do námi budovaného autorského korpusu, neboť se nejedná o homogenní autorovy verbální projevy, ale o rozhlasovou koláž, jež obsahuje i cizorodé texty. Čepovy poznámky byly původně určeny jako podklad pro mluvené slovo v rámci pravidelných rozhlasových vystoupení. Tyto náčrty tedy představují jakýsi scénář, kde jsou vyznačeny jednotlivé hlasové party, včetně dlouhých ukázek z určitého díla, o kterém Čep hovořil. Před samotným včleněním tohoto materiálu do Čepova korpusu

je proto nezbytné jej detailně prohlédnout a z důvodu nežádoucí kontaminace odstranit veškerý text, jehož autorem není Čep. Týká se to například vstupních hlasů, resp. textu, který je citací z aktuálně pojednávaného díla, nebo rámcujících ohlášení komentátora, jež upozorňují na začátek nebo konec Čepova rozhlasového programu.

Následující tabulka (viz Tab. 5) představuje úplné rozvržení Čepova korpusu a slovníku. Princip je analogický tomu, který prezentují *Slovník Karla Čapka* nebo *Slovník Bohumila Hrabala*; organizační strategie makrostruktury slovníku je založena na segmentaci textového materiálu dle žánrových tříd s tím rozdílem, že v případě Čepova slovníku dominují literárněvědné aspekty, což se projevuje již při nejobecnější segmentaci textů do dvou základních skupin. Na této úrovni bude možné Čepův korpus srovnávat nejen s jinými autorskými korpusy, ale *de facto* s jakýmkoli korpusovými hodnotami, které poskytuje ČNK. Z dané základní situace bude vycházet i autorský frekvenční slovník, který je samostatnou přílohou monografie.¹⁴⁷

+FIK	-FIK			
<i>Dvojí domov</i>	eseje a úvahy	korespon- dence	deníky	ostatní publicistika: nekrology, přednášky, poznámky, komentáře, recenze, imprese atd.
<i>Vigilie</i>				
<i>Zeměžluč</i>				
<i>Letnice</i>				
<i>Děravý plášť</i>				
<i>Hranice stínu</i>				
<i>Modrá a zlatá</i>				
<i>Tvář pod pavučinou</i>				
<i>Polní tráva</i>				
<i>Cikáni</i>	texty pro rádio Svobodná Evropa			
časopisecké texty z 20. let				
časopisecké texty z 30. let				

Tabulka 5: Makrostruktura Čepova korpusu a slovníku dle žánrové příslušnosti textů do dvou základních (sub)korpusů definovaných příznakem fikčnosti (+FIK/–FIK). Oblast tzv. fikčních narativů (+FIK) je dále chronologicky segmentována (dle prvního vydání) na jednotlivé povídkové soubory – (sub)korpusy.

2.3.1.1 Strukturace korpusu fikčních narativů (model relativní chronologie)

Stávající přehled založený na textech, které řadíme do tzv. korpusu beletrie (fikční narativy), lze dále funkčně strukturovat. Nadále bude naše pozornost soustředěna již výhradně na tuto oblast. Jedna z možností, jak smysluplně systematizovat a zpřehlednit

¹⁴⁷ Viz příložené CD.

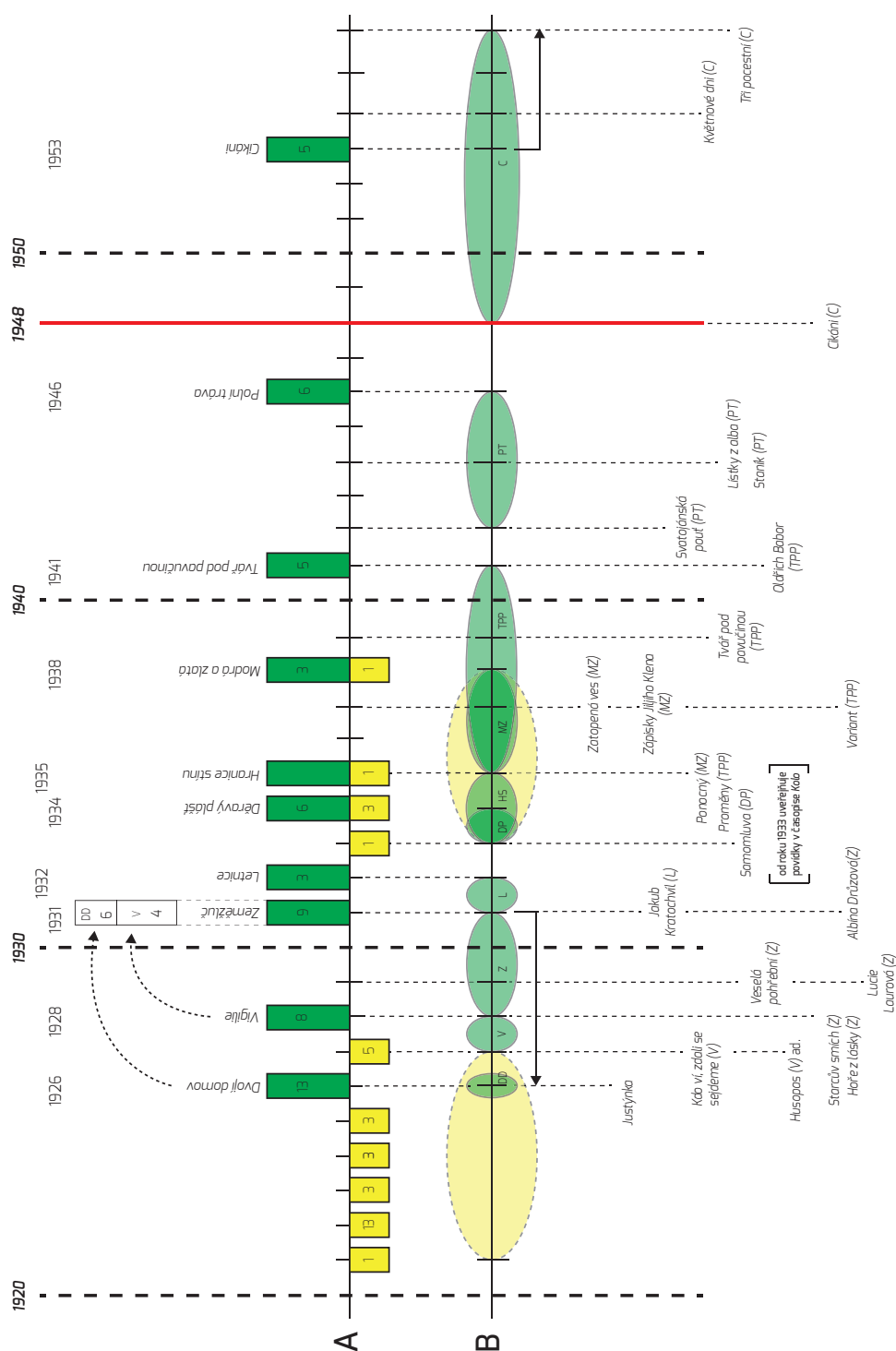


Schéma 4: Model relativní chronologie Čepovy beletristické tvorby se dvěma osami pro interpretaci vztahů mezi jednotlivými subkorpusy. Zeleně jsou vyznačeny povídkové knihy, žlutě časopisecké prózy, čísla v rámečcích představují počet textů v (sub)korpusu.

Čepovu beletristickou tvorbu, zejména pak její vývojovou situaci, je využít schematického znázornění pomocí množin, které budou reprezentovat vždy určitý soubor či skupinu textů. Každá množina bude představovat (sub)korpus k příslušné povídkové knize nebo pouze časopisecky publikovaným povídkám. Pomocí schematického zobrazení (Schéma 4) lze nejen přehledně uspořádat relativní chronologii vzniku jednotlivých textů a knih, ale současně získat rámcový přehled o základní strukturaci a organizaci korpusu fikčních narativů.

Model představuje relativní chronologii Čepových beletristických textů, které jsou distribuovány do dvou základních množin: nejpočetnější množinu tvoří povídkové soubory (celkem 58 textů), vedle kterých jsou situovány časopisecké povídky (celkem 34 textů), jež nebyly zařazeny do žádné z povídkových knih. Základními orientačními prvky v modelu jsou osy A a B. Vrchní horizontální osa A určuje punktální časovou rovinu v rozpětí mezi roky 1920–1956. Na ní jsou umístěna jednotlivá první vydání Čepových povídkových souborů spolu s nejstaršími povídkami, které byly otištěny zvlášť před vznikem příslušného finálního souboru. Osa B je pomocná a slouží k zobrazení relativních časových rozsahů vzniku jednotlivých knih nebo tvůrčích etap (v případě časopiseckých povídek). Její charakter je procesuální, což dovoluje sledovat vztahy mezi jednotlivými tvůrčími fázemi. Z formálních důvodů, o kterých již byla řeč, definujeme časové určení každé množiny vznikem prvního textu příslušného povídkového souboru a rokem jeho prvního vydání. V případě nezařazených povídek tyto limity navrhuje na základě časové kumulace rozptylu textů. Takto tedy vzniká dvanáct množin. Vyše jsme totiž konstatovali, že z hlediska kvantitativní a korpusové analýzy je vhodné, aby (sub)korpusy nebyly příliš malé, jelikož může docházet ke zkreslení výsledků, jež nemusí vykazovat systémové, ale náhodné jevy platné pouze pro jediný text. Proto například navrhuje vytvořit jeden (sub)korpus z Čepových časopiseckých próz ze 30. let, ačkoliv mezi poslední povídkou z roku 1938 (*Detektivní příběh*) a povídkou předposlední (*Sedláci v hospodě*) je rozdíl tří let, což v případě knižních souborů představuje hranice mezi jednotlivými (sub)korpusy. Model nám ukazuje, jakým způsobem lze postupovat při sestavení základní koncepce slovníku, aby jeho výpovědní hodnota byla v souladu s jednou z jeho hlavních funkcí, reflektovat vybrané literárněvědné požadavky. Cílem totiž není, jak bylo řečeno, vytvořit výhradně frekvenční slovník, který by lexikostatisticky zpracoval autorský korpus a podal pouze číselný přehled o počtu lexikálních jednotek – jakkoli je takový úkol důležitý pro navazující lingvistické bádání –, ale pokusit se využít prostředků moderní lingvistické kvantitativní a korpusové metodologie v kontextu literární vědy. Vrátime-li se k našemu modelu, lze v něm vysledovat přibližně dvě oblasti, ve kterých dochází k významnější kumulaci textů. První bychom mohli vydělit lety 1921–1932, druhou lety 1933–1946. Tato je centrem Čepovy literární tvorby,

kteřá vznikla před rokem 1948, tedy před autorovým definitivním odchodem do emigrace. K ní připojujeme texty, jež spadají zhruba do první poloviny 50. let a které zakládají (sub)korpus *Cikáni*.

Ze zobrazení vyplývá, že Čepova tvorba se nevyvíjela zcela jednoduše lineárně, ale spíše v jakýchsi pulzačních etapách,¹⁴⁸ jež jsme vymezili uvedenými množinami. Pokusme se nyní toto schéma interpretovat s ohledem na čistě formální aspekty.¹⁴⁹ Mohli bychom využít matematického zápisu množin, do kterého bychom převedli vztahy mezi jednotlivými (sub)korpusy. Je však třeba upozornit, že uvedené vztahy jsou relacemi časových horizontů, nikoli průniků mezi prvky (texty). Současně se nejedná o vztahy intertextuální, ale čistě formální. V našem případě totiž pro veškeré texty, které tvoří korpus fikčních narativů, platí, že se vždy vyskytují pouze v jedné z množin. Konfrontace formálních a kvalitativních (sémantických) kritérií vytváří prostor pro analýzu různých vztahových aspektů. Vzájemné poměry mezi množinami vyjádříme následovně (viz Tab. 6):

ZÁPIS	INTERPRETACE
$DD \subseteq \text{časopisecké texty 20. let}$	<i>Dvojí domov</i> náleží do časového období vzniku Čepových juvenilií.
$DP \subseteq \text{časopisecké texty 30. let} \wedge DP \subseteq HS$	<i>Děravý plášť</i> náleží do časového období vzniku časopiseckých próz z 30. let a současně do období vzniku románu <i>Hranice stínu</i> .
$HS \subseteq \text{časopisecké texty 30. let}$	<i>Hranice stínu</i> náleží do časového období vzniku časopiseckých próz z 30. let.
$MZ \subseteq \text{časopisecké texty 30. let} \wedge MZ \subseteq TPP$	<i>Modrá a zlatá</i> náleží do časového období vzniku časopiseckých próz z 30. let a současně do období vzniku <i>Tváře pod pavučinou</i> .
$TPP \cap \text{časopisecké texty 30. let}$	<i>Tvář pod pavučinou</i> se časově prostupuje s tvorbou časopiseckých povídek z 30. let.

Tabulka 6: Zápis vybraných vztahů mezi množinami osy B (viz Schéma 4) a jejich formální interpretace.

Uvedené matematické výrazy musíme chápat jako pomocný nástroj interpretace formálních vztahů v rámci autorského korpusu fikčních narativů a jeho jednotlivých segmentů – (sub)korpusů, nikoli jako absolutní predikaci reálné a komplexní situace

¹⁴⁸ Inspirací nám je publikace Dalibora Turečka *České literární romantično* (2012), v níž autor rozvinul tzv. synopticko-pulzační model literárního vývoje, který aplikuje na vývojové etapy české literatury 19. století (viz Tureček 2012). Teoretický model představil Peter Zajac (viz Zajac 2006, s. 13–23).

¹⁴⁹ Pouze podotýkáme, že kvalitativní hodnocení je v tuto chvíli zcela vyloučeno, neboť usilujeme o vytvoření formálního modelu slovníku.

Čepova tvůrčího vývoje. Ve výsledku, resp. v rámci literárněvědné analýzy a interpretace, ke které směřujeme, nelze pochopitelně zacházet s takto striktními výrazy; zde neoperujeme s čísly jako diskretními jednotkami, ale s uměleckými texty, jejichž hierarchie a vzájemná provázanost je řízena dle pravidel uměleckého strukturování, jež v případě beletrie zahrnuje komplikovanou problematiku intertextuálních vztahů, které se promítají (nestejným způsobem) do různých vrstev narativu, např. do jednotlivých oblastí fikčních světů stejně jako do výrazových prostředků narace. Jinými slovy, v umění neplatí přísné matematické vztahy, přesto – jak se domníváme – je inspirující a podnětné využít původně exaktních matematicko-statistických kritérií k pokusu uchopit některé z jevů, jež jsou nedílnou a vlastní součástí umělecké oblasti, tj. především jazykové materie, která je jako jeden ze strukturních elementů výsledného textu či díla schopna odrážet tvůrčí procesy a změny, stejně jako intertextuální vazby. Při výsledné literárněvědné interpretaci bude pochopitelně nutné překvalifikovat chápání vztahů mezi vyznačenými množinami z matematických relací na procesuální pojetí tvůrčího vývoje, které by měla reprezentovat epistemologicky koherentní literárněvědná metodologie vycházející ze strukturně-sémiotické tradice. Z hlediska vývojových tendencí v literatuře operuje základní strukturní metoda rámcově s myšlenkou vývojově starších a mladších fází, stejně jako s přechodovými oblastmi, přičemž dynamiku vývoje spatřuje ve funkčně-významovém působení strukturních prvků jednotlivých řad a v jejich potencialitě vytvářet strukturní uskupení nových (progresivních) kvalit. Dané tvrzení, se kterým se setkáváme například v raném literárněvědném strukturalismu, nemusí platit striktně, jak dokládá i jmenovaná Turečkova publikace (viz pozn. 148) ukazující spíše na pulzující než na přísně kauzálně-lineární průběh literárního vývoje. Proto i model, který zde prezentujeme, je třeba chápat jednak jako předpoklad nejen pro smysluplné rozvržení vlastního slovníku, ale především jako určitý „klíč“ k interpretaci statistických hodnot.¹⁵⁰ Na základě uvedeného modelu bude například možné sledovat frekvenční a korpusové hodnoty lexikálních jednotek nejen v závislosti na chronologii (osa A), ale také s ohledem na kumulaci textů (osa B), což se reálně promítne do otázky, nakolik se ze statistického hlediska od sebe jednotlivé soubory textů liší, popřípadě jak příznakové je určité rozložení. Hlavním smyslem je tedy statisticky postihnout a zaznamenat vývojové fáze Čepovy tvorby, které se mohou vůči sobě ocítat v různých modálních vztazích (kauzálních, variačních, regresivních, pulzačních apod.). Je proto nezbytné reflektovat také aspekty kvalitativní, které představují distribuce lexémů, slovních tvarů, kontextů či lexikálních syntagmat. Neplatí tedy, že by kvantitativní hledisko zcela vyloučilo a nahradilo rozměr kvalitativní. Oba

150 Konkrétní hodnoty představují číselné údaje seřazené v tabulce; samy o sobě nemají přílišnou výpovědní hodnotu.

póly musejí být funkčně propojeny a vzájemně se podmínovat.¹⁵¹ Další otázky vyplývající z modelu mohou směřovat k porovnání konkrétních vývojových celků, například se lze ptát, jak spolu souvisí tvorba 20., 30. a 40. let nebo jaké vztahy panují mezi skupinami, které jsme definovali pomocí matematického zápisu. Současně je třeba si klást otázky po vztazích napříč celým dílem (korpusem), kdy se naopak na základě kvantitativně-korpusových analýz mohou prokázat nové relace a vazby, jež jednoznačně nevyplývají z časové příbuznosti a *vice versa*. S uvedeným konstatováním souvisí i obecnější úvahy směřující k systémovým otázkám lexikální situace v Čepově beletrii a k možnostem její kvalitativní interpretace na základě kvantitativní analýzy. Samozřejmě se také nabízí prostor pro širší komparace, ať již v rámci celého autorského korpusu mezi jeho jednotlivými žánrovými částmi,¹⁵² nebo ve vztahu k externím hodnotám, např. ČNK nebo jiných autorských korpusů.

Nyní je třeba konstatovat, že model množin se pokouší především systematicky uspořádat relativní chronologii a ukázat, jak se k sobě časově vztahují jednotlivé části Čepovy tvorby. Na jedné straně vytváří určité chronologické kontinuum, které realizují jednotlivé povídkové soubory, na druhé straně dokládá, že toto kontinuum není ideální ve smyslu prosté linearit; Čepova beletristická tvorba (a nejen ona) je v mnoha ohledech paralelní. Model tak ukazuje kromě jiného i oblasti shluků a kumulací. Vyjdeme-li při koncipování základních parametrů frekvenčního slovníku z uvedených poznatků, lze řazení jednotlivých kategorií ve slovníku uspořádat tak, aby ve výsledku na sebe navazovaly a aby frekvenční hodnoty mohly být srozumitelně srovnávány.

2.3.2 Základní osnova frekvenčního slovníku

Protože výslednou podobu frekvenčního slovníku představuje tabulka číselných hodnot, bylo i z tohoto důvodu nezbytné, abychom dříve, nežli přistoupíme k vlastnímu návrhu slovníku a jeho parametrů, koncipovali model relativní chronologie Čepových beletristických textů. Číselnou matici základních frekvencí bylo potřeba strukturovat do přehledné podoby. Jelikož základní slovník vychází z excerpce prozaických textů, hlavní pozornost je věnována lexikálním jednotkám, a to zejména obecné (systémové) úrovni,

151 Srov.: „Kvantifikace předloží problém a morfologie najde řešení. Rád bych však dodal, že jen tehdy, pokud se vše odehrává, jak má. Asymetrie mezi kvantitativním *explanandem* a kvalitativním *explanans* je tak markantní, že zde máme často skvěle definovaný problém, avšak s nulovou možností řešení.“ (Moretti 2014, s. 33) Moretti ovšem hovoří o velkých datech, která reprezentují epochy literárních žánrů. V takovém případě je možné, že pro formálně detekovaný jev chybí náležitě vysvětlení. Náš případ je mnohem skromnější; omezuje se na konkrétní povídkové soubory a tvorbu jednoho autora.

152 V konečné fázi realizace celého Čepova korpusu půjde o vypracování obdobného modelu, který bude napojen na stávající model a který bude analogicky řešit situaci v oblasti žánrových skupin.

kteřou reprezentují lemmata. Náš zájem se bude soustředit na řadu aspektů, jež se s touto oblastí pojí: kromě lexémů budou analyzována především lexikální syntagmata (kolokace) či lexikální paradigmata (sémantické třídy). Na základě výše uvedeného modelu (viz Schéma 4) lze sestavit základní (rámcovou) podobu Čepova frekvenčního slovníku fikčních narativů (viz Tab. 7).

pořadí	lemma	prózy z 20. let		Dvojí domov		Vigilie		Zeměžluč		Letnice		prózy z 30. let		Děravý plášť		Hranice stínu		Modrá a zlatá		Tvář pod pavučinou		Polní tráva		Čikání	
		Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf	Af	Rf
#	-	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#

Tabulka 7: Rozvržení frekvenčního slovníku fikčních narativů Jana Čepa. Af = absolutní frekvence, Rf = relativní frekvence. Hodnoty Af a Rf budou sledovány i na úrovni celého korpusu fikčních narativů.

Tabulka, která představuje způsob řazení jednotlivých (sub)korpusů, vychází z modelu relativní chronologie a snaží se uspořádat Čepovu tvorbu dle formálních vztahů časové souslednosti a následnosti. Vedle sebe se tak ocitají (sub)korpusy *časopisecké prózy 20. let* a *Dvojí domov*, stejně jako (sub)korpusy *časopisecké prózy 30. let* a *Děravý plášť*, *Hranice stínu*, *Modrá a zlatá*. Tento způsob organizace umožňuje dvojí základní komparaci: chronologickou (diachronní) a synchronní, jež vyplývá z kumulace textů v určitém časovém období. Jelikož se v této části projektu bude jednat o frekvenční slovník, je třeba každý (sub)korpus chápat jako realizaci frekvenčního seznamu lexikálních jednotek. Jako základní (a samozřejmě nikoliv jediný možný) interpretační klíč k jejich čtení slouží model relativní chronologie.

Hlavní myšlenka Čepova korpusu a slovníku se odvíjí od dvou základních aspektů. Zprvče usilujeme postihnout chronologii autorova díla, která je určena přirozeným časovým průběhem v podobě prostého formálního následování jednotlivých textů a knih. Již dříve jsme však uvedli, že vývojové aspekty literární věda nereflektuje jako prostou formální kauzalitu diskrétních jednotek v čase, ale soustředí se na problematiku dynamického utváření významu a smyslu díla, které odpovídají kvalitativně různorodé strukturní vztahy jak uvnitř literárního systému (imanentní vztahy), tak na úrovni vnějších společensko-kulturních kontextů. Zadrugé: Ačkoliv kvantitativní metody samy o sobě takový komplexní analytický přístup neumožňují, přesto – jak se domníváme – se nám v tuto chvíli alespoň částečně podařilo propojit oba hlavní požadavky zastoupené jednak nárokem na formální exaktnost, kterou zpřístupní kvantitativní analýza autorského korpusu, jednak literárněvědnou pozicí, reprezentovanou např. textologickými či intertextuálními otázkami.

Jiným důležitým požadavkem je především striktní dodržení formálních kritérií z důvodů relevance výstupních informací kvantitativní analýzy autorského korpusu.

Hodláme-li současně testovat možnosti funkční kooperace mezi kvantitativně-korpusovou metodologií a literární vědou, musíme na druhou stranu důsledně zachovávat jejich výchozí epistemologické parametry, což v případě formální analýzy autorského korpusu znamená eliminaci subjektivní interpretace a její nahrazení formálními statistickými algoritmy.

V této fázi budování autorského korpusu se již otevírá možnost pro řadu konkrétních srovnávání, která mohou nejen přinést relevantní zjištění o lexikální situaci v Čepově beletrii, ale na jejím pozadí se vyslovit k rámcové strukturní povaze textových souborů. Z tohoto hlediska se jako velice zajímavá jeví situace prvních tří Čepových povídkových souborů, které se detailně budeme věnovat ve třetím oddíle. Obecně však je možné srovnávat nejen jednotlivé (sub)korpusy, např. z hlediska různě definovaných frekvenčních pásem lexémů a jejich lexikálních syntagmat (kontextů), ale také celý autorský korpus nebo jeho jednotlivé části s externími korpusy.¹⁵³ V neposlední řadě je v této fázi na úrovni korpusu možné sledovat některé specifické oblasti související například s tematickým zaměřením¹⁵⁴ a na jejich základě analyzovat proměnu tematického paradigmatu a také vztahů mezi povídkovými soubory. Představit si lze i jiný přístup, kdy na bázi korpusových hodnot budou modelovány strukturní poměry uvnitř Čepovy tvorby.

2.3.3 Segmentace na úrovni jednotlivých (sub)korpusů

Jednou z otázek, které si (nejen) v souvislosti s Čepovým slovníkem lze klást, je možnost jemnější taxonomie korpusu, která by byla schopna reflektovat zvolená literárněvědná kritéria. To však představuje nemalý problém, kterým je snaha sloučit dva hlavní a důležité aspekty: ryze formální přístup při textové segmentaci, který vyžaduje objektivní kvantitativně-korpusová analýza, a postup, který reflektuje celkovou funkční i sémantickou složitost promluv v narativu, zejména pak jejich modálně-sémantické dispozice a strategie. Naratologie poskytuje řadu modelů, které analyzují různé promluvové segmenty v uměleckých textech, avšak žádný z nich není v úplnosti striktně převoditelný na čistě formální princip, a navíc takovým způsobem, aby byl vyjádřený konkrétním algoritmem. Nejbliže má k této úvaze Doleželův systém, neboť systematicky kombinuje jazyková kritéria v podobě základních distinktivních rysů promluvoových úseků spolu s kritérii funkčně-sémantickými. Ačkoli se zdá, že Doleželův model má ze všech naratologických modelů skutečně nejbliže k možné formalizaci, přes veškeré úsilí, které bylo spojeno s jeho transformací za účelem dosažení maximálně možné míry formalizace

¹⁵³ Viz třetí oddíl.

¹⁵⁴ Tzv. tematické koncentraci textu se budeme věnovat ve třetí části monografie.

a objektivizace daného systému v rámci korpusové segmentace, se ve výsledku ukázalo, že je obtížné některé promluvvé segmenty jednoznačně klasifikovat.¹⁵⁵

Proces modifikace Doleželovy taxonomie spočíval v redukci modálních markerů u jednotlivých typů promluvy. Výsledkem byly základní narativní typy určené gramatickou osobou a přechodová oblast, kterou tvoří neznačená přímá řeč a polopřímá řeč (Schéma 5). Vycházeli jsme z původního Doleželova členění promluvy ve fikčním narativu (viz Tab. 8), které je založeno na distribuci omezené skupiny markerů, jejichž rozložení signalizuje konkrétní promluvvý typ.

	znaky P (V)	znaky P (P)
1. přímá řeč		grafické gramatické výpovědní sémantické slohové
2. neznačená přímá řeč	grafické	gramatické výpovědní sémantické slohové
3. polopřímá řeč	grafické gramatické	výpovědní sémantické slohové
4. smíšená řeč	grafické gramatické výpovědní sémantické slohové	výpovědní sémantické slohové (jako signály)
5. objektivní vyprávění	grafické gramatické výpovědní sémantické slohové	

Tabulka 8: Doleželova tabulka promluvvých typů ve fikčních narativech (Doležel 1993, s. 39).

Vztahy mezi jednotlivými oblastmi Doležel systematizoval do kruhového schématu (viz Schéma 5), u kterého jsme z požadovaných formálních důvodů provedli redukci, která sloučila subkategorie a pokusila se zachovat promluvvá pásma na základě nepřekrývající se distribuce jejich distinktivních rysů (viz Tab. 9 a Schéma 6).

typ promluvy	znaky P (V) A	znaky P (P) B	neutralizace	matematický zápis	vztah mezi množinami A a B	distinkce
T1	0/5	5/5	0	$T1 = \{A = \emptyset, B = \{1, 2, 3, 4\}\}$	$A \neq B$	pozitivní
T2	1/5	4/5	1	$T2 = \{A = \{1\}, B = \{2, 3, 4, 5\}\}$	$A \neq B$	pozitivní
T3	2/5	3/5	2	$T3 = \{A = \{1, 2\}, B = \{3, 4, 5\}\}$	$A \neq B$	pozitivní
T4	5/5	3/5	3	$T4 = \{A = \{1, 2, 3, 4, 5\}, B = \{3, 4, 5\}\}$	$A = B$	negativní
T5	5/0	0/5	0	$T5 = \{A = \{1, 2, 3, 4, 5\}, B = \emptyset\}$	$A \neq B$	pozitivní

Tabulka 9: Transformace Doleželovy tabulky do jednoduché numerické matrice s vyznačením míry neutralizace a matematickým zápisem vztahů mezi množinami A a B. Čísla ve složených závorkách představují jednotlivé distinktivní rysy tak, jak jsou řazeny v Doleželově tabulce, tzn. 1 = znak grafický, 2 = znak gramatický atd. Čtvrtý promluvový typ T4 je založen na rovnosti obou množin (v matematickém smyslu), neboť |A| a |B| obsahují identické prvky.

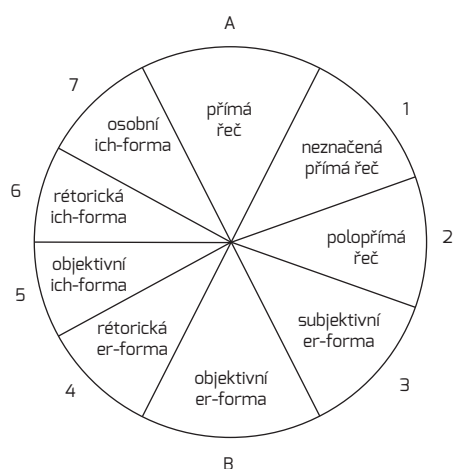


Schéma 5: Doleželova soustava promluv ve fikčním narativu.

Vztahy, které jsme vyznačili v Tabulce 9, ukazují, že pouze promluvové typy T1, T2, T3 a T5 disponují takovou distribucí rozlišovacích markerů, které se nepřekrývají. Doležel navíc u každé základní vypravěčské promluvy rozeznává její jednotlivé modifikace, jež jsou výsledkem míry vlivu subjektivních nebo objektivních rysů (subjektivní er-forma, objektivní ich-forma atd.). Jestliže uvedená klasifikace funkčně obohacuje rejstřík dvou základních narativních typů (ich-formy a er-formy), z hlediska čistě formální taxonomie se potom jedná o nejednoznačně klasifikovatelné segmenty, jejichž určení je závislé na určité míře interpretace. Pro potřeby ryze formální segmentace jsme proto Doleželův systém transformovali tak, aby byla zachována jeho podstata, ale abychom současně získali takové promluvové úseky, jejichž vyznačení v korpusu by spočívalo na jednoznačných formálních ukazatelích (viz Tab. 9, Tab. 10 a Schéma 6). Je totiž nutné zachovat kritérium formalizace autorského korpusu na všech jeho úrovních, což

je základní předpoklad kvantitativně-korpusové metody. Předem interpretovaná data by *de facto* zkreslovala objektivní parametry analýzy.

vztah	typ
$A \neq B$	přímá řeč
	objektivní Er-forma
$A \cap B$	neznačená přímá řeč
	polopřímá řeč
$3 \sim B \sim 4$	objektivní Er-forma
	subjektivní Er-forma
	rétorická Er-forma
$5 \sim 6 \sim 7$	objektivní Ich-forma
	osobní Ich-forma
	rétorická Ich-forma

Tabulka 10: Vztahy mezi promluvovými typy.

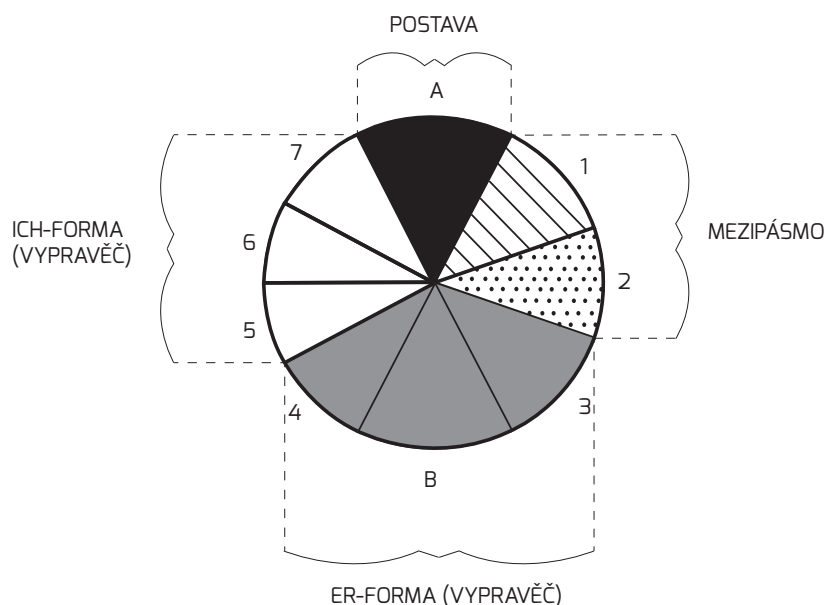


Schéma 6: Redukce Doleželovy soustavy promluv ve fikčním narativu.

I přes pokročilý proces formalizace, který je dosažen úpravou Doleželova systému, zůstává základní problém s oblastí tzv. mezipásma,¹⁵⁶ kterou jsme vymezili neznačenou přímou řeč a polopřímou řeč. Ačkoli Doležel stanovil konkrétní kritéria při klasifikaci těchto promluv, pro potřeby maximálně formalizované analýzy autorského korpusu se

¹⁵⁶ Jedná se o námi zavedený pojem.

v praxi ukázaly být tyto parametry přece jen poněkud komplikované. Při jejich vymezení totiž není možná relativně spolehlivá automatická formalizace tak, jako je to možné při lematizaci. Tím by na úrovni formální analýzy textu docházelo k nežádoucí metodologické konfuzi, neboť by při rozlišování konkrétního promluvového typu mezipásma přece jen docházelo k určité míře subjektivní interpretace těchto úseků, jež se zakládají na ne vždy zcela jasných hranicích.¹⁵⁷

Z důvodu zachování jednoznačné segmentace, která je při korpusové a kvantitativní analýze nezbytná, zbývá *de facto* možnost, kterou představuje klasifikace narativních typů dle gramatické osoby, tzn. ich-forma, er-forma a v případě Čepových textů i du-forma. Vedle toho lze jednoznačně vydělit přímou řeč. Korpusová a kvantitativní analýza uvedených segmentů by pak mohla umožnit sledovat míru subjektivizace či objektivizace základních narativních typů stejně jako analyzovat povahu řečové promluvy postavy. V žádném případě ovšem nelze tuto potenciální segmentaci konstruktivně chápat mimo tento rámec, tj. v kontextu vlastního naratologického diskurzu, kde je naopak takové zjednodušení nežádoucí.¹⁵⁸ Distribuce, která respektuje pouze elementární formální dichotomii mezi vypravěčskými způsoby, by mohla v rámci autorského korpusu detekovat jevy, jež se týkají jednak typologického výskytu jazykových jednotek (syntaktika, autosemantika) v jednotlivých narativních pásmech v závislosti na frekvenčním zatížení a jeho průběhu, jednak stanovit specifická lexikální paradigmaty, která by bylo možné interpretovat jako paradigma pro danou narativní kategorii. Jednotlivé vztahy mezi zmíněnými kategoriemi by následně podléhaly statistickému testování, které by specifikovalo významnost změn, ke kterým by v jednotlivých korpusech docházelo. Jak jsme uvedli na příslušném místě, korpus ani frekvenční slovník neodráží výhradně jen kvantitativní vlastnosti, ale spolu s nimi (v přímé návaznosti) také aspekty kvalitativní (lexikální). Obě hodnoty, tj. sémantiku a četnost, bude možné funkčně využít nejen při interpretaci příslušné kategorie, ale rovněž se zde v návaznosti na celkovou strukturu slovníku otevírá možnost komparace těchto kategorií z hlediska jejich vývoje uvnitř Čepovy prozaické tvorby. Vývojové hledisko našeho modelu (Schéma 4) – jak bylo konstatováno – dovoluje realizaci jak s důrazem na diachronii (relativní chronologie), tak s ohledem

157 Na tuto skutečnost v případě Čepova *Dvojího domova* (1926) upozornil zejména Kubíček (2014, s. 53).

158 Viz např. studii Wayne C. Bootha *Typy vyprávění*, kde autor konstatuje: „Když začneme probírat četné narativní prostředky, které se nám v próze nabízejí, brzy pocítíme, jak zoufale nedostačující je naše tradiční klasifikace ‚vypravěčských hledisek‘ na tři nebo čtyři druhy, které se vzájemně liší pouze ‚osobou‘ a mírou vševědouceho. Když si jako příklad vezmeme některé ze slavných vypravěčů – řekněme třeba Cervantesova Cida Hameta Benengeliho, Tristrama Shandyho, já‘ románu Middlemarch a Strethera, jehož očima vnímáme většinu děje románu *The Ambassadors* Henryho Jamese, uvědomíme si, že popíšeme-li kteréhokoli z nich termíny jako ‚vypravěč v první osobě‘ a ‚vševědoucí vypravěč‘, nedozvíme se nic o způsobu, jakým se tito vypravěči od sebe navzájem liší nebo proč jsou zrovna tito vypravěči tak úspěšní, zatímco jiným, které bychom mohli popsat stejnými termíny, se tolik nedaří.“ (Booth 2007, s. 42)

na kumulativní (pulzující, synchronní) momenty v Čepově tvorbě, což by platilo i pro potenciálně nejnižší korpusové segmenty.

Než dojde k reálnému zapojení jmenovaných aspektů (promluvových pásem v narativu ve formě korpusových úseků) do konkrétního (sub)korpusu, musí být nejprve tato problematika detailněji zkoumána a testována s ohledem na její relevanci vůči alespoň elementárním nárokům literárněvědné naratologie v přímé souvislosti s požadavky formální korpusové taxonomizace. Důležité je zejména rozlišení sémantických modalit narativních promluv, které by umožnilo sledovat míru subjektivizace či objektivizace dané promluvy.¹⁵⁹

Jestliže jsme až doposud hovořili o korpusu fikčních narativů Jana Čepa, zbývá uvést, že výsledná podoba Čepova korpusu a frekvenčního slovníku bude zahrnovat veškeré autorovy texty, tzn. kromě beletrie i publicistiku, esejistiku, korespondenci, deníkové záznaky apod. (viz Schéma 7).

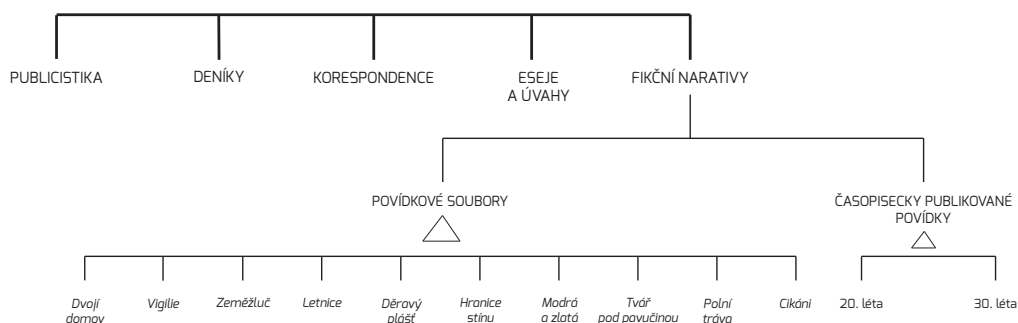


Schéma 7: Celkový model Čepova slovníku.

2.3.4 Kontext a sémantika

Doposud jsme se zabývali formálními kritérii při budování autorského korpusu. Hlavní jazykovou jednotkou, kterou hodláme sledovat, je slovo, resp. lexém. Automatický (strojový) výpis lexikální databáze příslušného (sub)korpusu spolu s frekvenčními hodnotami podává důležité informace o zatížení daného lexému nebo slovního tvaru. Kromě jiného lze z takového seznamu generovat jednotlivá frekvenční pásma, jež informují například o centru či periférii příslušného lexikonu a zejména o proměnách lexikální skladby určitého frekvenčního pásma v závislosti na změně (sub)korpusů (viz relativní chronologii). Taková zkoumání na materiálu české beletrie v minulosti učinila Marie Těšitelová, která výsledky svého výzkumu publikovala mimo jiné v monografii *Otázky*

159 Měření těchto segmentů by vykazovalo určité výraznější rozdíly zejména na velkých korpusech, kde by bylo možné sledovat distribuci základních narativních forem a jejich výrazových prostředků a kontextů, stejně jako oblasti přechodů dílčích paradigmat těchto kategorií v závislosti na epoše, autorovi apod.

lexikální statistiky (1974), ve níž definovala tři základní frekvenční pásma lexikonu – oblast s nejvyšší a vyšší frekvencí, oblast se střední frekvencí, oblast s nižší a nejnižší frekvencí, jež dále podrobovala zkoumání z hlediska jejich struktury a distribuce. Na základě frekvenčních hodnot lexika Těšitelová přistupuje k dalším statistickým mírám, pomocí kterých lze lexikon analyzovat; k takovým náleží například bohatství, jeho rozptyl či koncentrace ad. Kromě jiného se Těšitelová věnuje proporčním charakteristikám nominálních a verbálních skupin u jednotlivých výběrových autorských lexikonů.¹⁶⁰ Zkoumání, kterým se Těšitelová věnuje, jsou přece jen limitována izolovanými jazykovými jednotkami – jednotlivými slovy a jejich frekvenční hodnotou. Nás ovšem musí zajímat i jejich kontext, a to především z toho důvodu, abychom se mohli blíže věnovat vybraným otázkám sémantiky. Pro moderní jazykové korpusy je kontextová analýza již naprostou samozřejmostí a nezbytností. Naopak lze říci, že teprve elektronický korpus umožnil systematictější zkoumání této problematiky na bázi strojového algoritmu.

Kontext je možné v korpusu sledovat na úrovni konkordance, kterou František Čermák definuje jako „různě velké dílčí seznamy slovních forem v jejich přirozeném (co do rozsahu volitelném) kontextu, často s dodatečnou informací o místě výskytu v původním textu, frekvenci výskytu apod. Hlavní výhodou konkordance, užívané obvykle v běžném standardu KWIC (Key Word in Context), je možnost studia slova (popř. jen jeho části), resp. všech jeho forem vedle sebe (obvykle se řadí abecedně), jejich kontextů, a tím také různých pevných a habituálních kolokací, obecněji pak **kolokability** (spojitelnosti) slova, resp. jeho formy a valence“ (Čermák 1995, s. 126). Automatická konkordance umožňuje kompletní či filtrovaný výpis kontextu hledané slovní jednotky (lemma, token).¹⁶¹ Analýzu konkordancí lze provádět samostatným manuálním prohledáváním jednotlivých konkordančních řádků, nebo je pro výpis kontextu možné využít některou z funkcí pro vyhledávání příznakových lexikálních souvýskytů. Mezi takové funkce, jimiž disponuje i Český národní korpus, náleží především t-score, MI-score,¹⁶² log-likelihood, Dice a další.¹⁶³ Příznakovými souvýskyty jsou potom takové struktury, které tvoří ustálenější lexikální (a samozřejmě smysluplná) syntagmata. Za kolokaci lze

160 Těšitelová nepracuje s úplným autorským lexikonem, ale pouze s jeho výběrem, který je pro ni dostatečně reprezentativní, aby reflektoval obecné jevy celého lexikonu.

161 Srov. Baker – Hardie – McEnery 2006, s. 42–44.

162 „[Č]íselná hodnota MI-score vyjadřuje asociativní sílu kolokace, jde o poměr skutečné frekvence kolokace a frekvence očekávané v případě, že by byla slova v korpusu rozložena náhodně. Hodnota t-score naopak vyjadřuje míru jistoty, se kterou není nalezená kolokace dílem náhody, ale naopak pravidelným jevem; neříká však téměř nic o její síle.“ (Křen 2006, s. 224) Viz dále např. Čermák 2006a, 2006b, s. 13, 95.

163 K jejich srovnání viz Křen 2006, s. 223–247.

považovat spojení lexikálních jednotek,¹⁶⁴ které vykazuje těsnější vazbu mezi dvěma lexémy či slovními tvary nacházejícími se většinou v nejbližším okolí: „[C]ollocation is the phenomenon surrounding the fact that certain words are more likely to occur in combination with other words in certain contexts. A collocate is therefore a word which occurs within the neighbourhood of another word.“¹⁶⁵ (Baker– Hardie – McEnery 2006, s. 36–37) Rozdíl mezi konkordancí a kolokací spočívá v tom, že pro konkordanční korpusový výpis je relevantní jakýkoli kontext, resp. veškerý kontext hledaného výrazu, tzn. i kontext na mezivětných předělech apod., zatímco pro specifikaci vlastní kolokace je naopak tento jev nežádoucí (viz Cvrček 2006), neboť smysl kolokace spočívá v zachycení příznakového a relevantního lexikálního souvýskytu, který je tvořen těsnějším lexikálně-sémantickým syntagmatem.¹⁶⁶ Důležitým ukazatelem kolokace je kolokabilita lexémů (viz Čermák 2006a, s. 11), tedy schopnost systémově vstupovat do vztahu s jinými slovními jednotkami a spolu s nimi vytvářet kompatibilní a sémanticky koherentní celky (jednotky). Toto spojení může být systémové (pevné), nebo nesystémové (volné), náhodné.

Při analýze kontextu v autorském korpusu je nezbytné zaznamenávat veškeré souvýskyty, tj. nejen výhradně ty, které lze z hlediska národního jazyka jako globálního celku považovat za systémové. Primárním úkolem je sledovat lexikálně-sémantické jevy uvnitř specifického systému, který je utvářen literárními texty. Ten nad obecnou jazykovou rovinou staví vlastní strukturu a síť vztahů mezi různými komponenty (jednotkami) celku. V rámci literárního systému platí jedinečné zákonitosti, které ovlivňuje například tvůrčí záměr ve spojení s historickou potencialitou literárního systému, narativní strategie, estetika apod. Tyto aspekty je třeba mít na paměti ve chvíli, kdy provádíme korpusovou analýzu autorského lexikonu s cílem využít jejích výsledků nejen pro lingvistické účely, ale pokusit se je zapojit také do kontextu literárněvědného bádání.

Studium kontextu v korpusové lingvistice náleží k hlavním oblastem jejího zájmu,¹⁶⁷ neboť pomocí takového výzkumu je schopna na ose systémovost/nesystémovost

164 F. Čermák konstatuje, že při vyhledávání slov je vhodnější se soustředit na konkrétní tvary: „Kolokace jsou vytvářeny primárně slovními tvary (tedy lexý), které často na celý lexém nelze zobecnit [...] Proto je ve většině případů při hledání v korpusu mnohem přínosnější se soustředit na souvýskyty tvarů v kolokacích místo na méně říkající souvýskyty lemmat...“ (Čermák 2006a, s. 12)

165 „[K]olokace se týká té okolnosti, že jistá slova se v určitém kontextu vyskytují častěji v kombinacích s jinými slovy. Kolokát je proto slovo, jež se vyskytuje v blízkosti jiného slova.“]

166 Záměrně nepoužíváme pojem *promluva*, neboť je širší než pojem *věta* a zahrnuje pod sebe skupinu vět. Na druhé straně každá věta v našem označeném korpusu náleží do určité promluvy, a typologickým zobecněním souvýskytů lze korpusově charakterizovat dané promluvě pásmo.

167 Aktuální lingvistickou publikací systematicky se zabývajícím kvantitativně-korpusovou analýzou kontextu je monografie Václava Cvrčka *Kvantitativní analýza kontextu* (2013), ve které autor zavádí tři hlavní míry pro jemnější kontextovou analýzu: proximitu, variabilitu a obligatornost kontextu. Jmenované jednotky se ukazují jako velice vhodné doplnění

analyzovat posuny a vývoje v kombinaci lexikálních jednotek, což má rovněž své funkční přesahy do oblasti gramatiky.¹⁶⁸ Postupy kontextové analýzy, které nabízí korpusová lingvistika, lze funkčně využít také s ohledem na literárněvědné účely. Jak již bylo řečeno, kromě zjišťování kolokací, pevných vazeb lexikálních souvýskytů, zde půjde o zjišťování všech příznakových souvýskytů, které nemusejí být nutně ustálenými kolokacemi. Tento požadavek je dán jednak specifícností materiálu, kterým je literární jazyk, jednak cíli, jež zde sledujeme a jsou primárně literárněvědné. Lexikální kombinace (obecně *n-gramy*¹⁶⁹) vytvářejí v systému konkrétního literárního jazyka zvláštní *autorské kolokace*, jež disponují symptomatickými rysy v rámci literárního systému. Ty však nemusejí být na úrovni národního jazyka jako celku stejně příznakové. Uvedená skutečnost z našeho pohledu představuje jev nadmíru pozoruhodný, neboť se v ní odráží schopnost bezprostřední reference k umělecké strukturní intenci textu. Četnost i charakter kontextových souvýskytů jsou pro literárněvědně orientovanou korpusovou analýzu důležité z hlediska primárního popisu symptomatických rysů diskurzivního narativního pole. Jakýkoli souvýskyt je z literárněvědného stanoviska relevantní, neboť uvažování o těchto souvýskytech vychází z pravidel umělecko-narativní syntagmatiky a je výrazem například narativní intence vypravěče nebo promluvového pásma postavy. Souvýskyty ve fikčních narativech mají z pohledu kvantitativně-korpusové analýzy dvojdimenzionální povahu: realizují se na bázi obecného jazykového systému a současně uvnitř literárního systému (viz Schéma 8). Jako takové pak odkazují ke strukturní povaze obou systémů s tím, že je

asociačních měr především proto, že poskytují detailnější pohled na problematiku kontextu a sémantiky. Cvrček rozlišuje rovněž různé stupně kontextu, které specifikují výběr příslušného lexika, jež determinuje sémantiku sledovaného výrazu. Jedná se o tzv. bezprostřední, blízký a vzdálený kontext: „Čím je kontextová pozice vzdálenější od klíčového slova (KWIC), tím jsou omezení ovlivňující výběr jejího obsazení, která plynou z použití daného klíčového slova, menší.“ (Cvrček 2013, s. 28) A dále: „[Z]atímco blízký a zejména bezprostřední kontext reflektuje systémové (gramatické a lexikální) vlastnosti zkoumané jednotky, kontext vzdálený odráží vlastnosti vyplývající z jejích charakteristik odvozených z použití v textu. Můžeme tedy dodat, že blízký kontext ukazuje na vlastnosti, které se realizují v celních svojí velikosti nepřesahujících úroveň souvětí, vzdálené pozice poukazují na stylové a textotvorné charakteristiky, jež jsou určující pro povahu hierarchicky větších celků.“ (Cvrček 2013, s. 159)

168 „Především díky studiu korpusu se znovu naléhavě ukazuje v jazyce prioritou **pravidel sémantických**, která řídí kombinace slov a kolokace a která ukazují postupně, co je tu (proto)typické a nejčastější, průměrně běžné a všední a co už marginální, netypické, popřípadě možné, až co už je nemožné. Až na tomto základě fungují pak pravidla gramatická, formální, která jsou takto vždy druhotná, což nebývá dosud obecně reflektováno; bez splněných pravidel sémantických formálně (gramaticky) smysluplné kombinace slov ani věty vzniknout nemůžou.“ (Čermák 2006a, s. 10)

169 „Od kolokace je třeba odlišit bigram, trigram, tetragram atd., tj. dvou-, troj-, čtyř- či více-člennou kombinaci, resp. postupně a po sobě jdoucí kombinace dvou, tří, čtyř atd. všech sousedních slov, kterou z celého textu, resp. korpusu mechanicky, postupně a poměrně jednoduše vyčleňuje některý z mnoha už dostupných programů (srov. např. komerční *Collocate* nebo na webu dostupný *kfNgram*). Výsledkem jsou tedy např. všechny dvojice (trojice atd.) sousedních slov. To znamená, že v takto mechanickém a obvykle velkém úhrnu kombinací je i mnoho kombinací nesmyslných, náhodných apod. Teprve z úhrnu bigramů lze pak získávat, obvykle poměrně pracně, smysluplné kolokace.“ (Čermák 2006a, s. 11)

otázkou konkrétního výzkumu a způsobu vytěžování korpusu, ke kterému z nich mají získané informace primárně referovat.¹⁷⁰

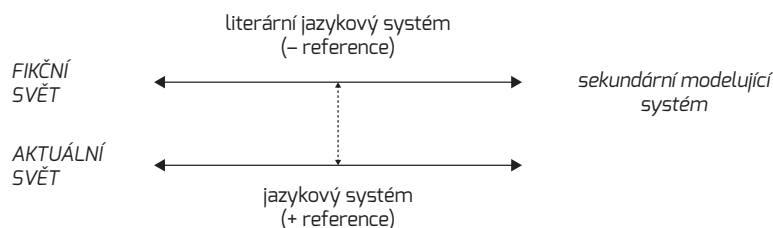


Schéma 8: Dorrit Cohnová definuje literární fikci rovněž na základě charakteristiky jejího jazyka, který je podle ní zásadně nerefereční, tzn. neodkazuje k předmětným či jevovým entitám reálného světa, ale do oblasti světa fikčního. Konkrétní denotáty jsou substituovány tzv. fikčními entitami (Cohnová 2009).

Lexikální souvýskyty a kontexty budou poměřovány mezi sebou navzájem (jako v případě frekvencí lexikálních jednotek), což umožní jejich komparaci mezi (sub)korpusy situovanými na časové ose. Tento postup dovolí sledovat nejen kvalitativní proměnu nejfrekventovanějších struktur souvýskytů pro danou oblast a předpokládat určité literárněvědné konsekvence, ale například na základě tohoto srovnání bude možné sestavovat významová paradigmatata pro danou korpusovou oblast, jež budou specifikovat příslušnou sémantickou třídu. Typologické komparace kontextových struktur budou prováděny v součinnosti s frekvenčními kritérii. Zajímavým aspektem pak v této souvislosti bude otázka paradigmatické jednotlivých kontextových struktur a jejich příslušnosti do sémantické třídy. Autorský korpus bude vedle zjišťování formální paradigmatické konkrétního lexika disponovat i paradigmatickou kontextů.

Prakticky budou kontexty a kolokace vyhledávány standardní korpusovou metodou KWIC a pomocí základních asociačních měr. Takové vyhledávání má tu přednost, že je systematictější a přesnější než běžná introspektivní zjištění prováděná v průběhu recepce textu (díla) literárním badatelem, a zejména umožní exaktní statistické vyhodnocování jevů. Sémantika lexikální jednotky z pohledu korpusové analýzy se vztahuje k poměrně jasně a přesně definovanému okolí, avšak tematika literárního díla jako významový komplex strukturně-kompoziční výstavby narativu, který zasahuje jak celkovou oblast diskurzu narativu, tak rovinu obsahu, příběhu (v Chatmanově smyslu¹⁷¹), je konstruována celým souborem vztahů uvnitř strukturovaného fikčního narativu. Zde dochází k vytváření vlastních paradigmaticko-syntagmatických vztahů mezi různě rozsáhlými oblastmi a rovinami, tedy nejen na rovině lexikální. Je tedy třeba opakovaně upozornit na

¹⁷⁰ S touto problematikou souvisí i způsob, jakým je konkrétní korpus navržen.

¹⁷¹ Viz Chatman 2008.

to, že samotná korpusová analýza je částečným obrazem strukturní povahy textu/textů a v první řadě ji musíme vnímat jako analýzu korpusu, nikoli jako konečný literárněvědný počín. Žádná kvantitativní a korpusová analýza se neobejde bez interpretace, jež vyplývá z konkrétní disciplíny a její metodologie. Analogicky to platí i v lingvistice, která se nemůže spokojit s čistými a izolovanými výsledky korpusové analýzy, ale potřebuje tyto informace v první řadě adekvátně interpretovat.

Budeme-li se v další části výkladu zabývat otázkou tematiky literárního textu, je třeba upozornit na zásadní rozdíl, který představuje korpusový a literárněvědný přístup. Tematiku literárního díla lze vedle řady rozdílných literárněvědných metod nahlížet rovněž jako problém, který přesahuje vnitřní hranice literárního díla, jak prokázaly výzkumy zejména tartusko-moskevské sémiotické školy. Relační vztahy typu literární dílo jako sémiotický objekt a okolí jsou v kontextu tohoto literárněvědného a kulturního uvažování promýšleny jako vztah sémantické jednotky (literárního díla) k sémiosféře, tj. kulturně-historickému prostředí nahlíženému ve smyslu heterogenní sémiotické sítě. Korpusová analýza tematiky textu naopak směřuje ke snaze za pomoci formálních algoritmů postihnout exponovanou oblast textu, ve které se tematika reflektuje na úrovni strukturní organizace lexika a jeho kontextů. Smyslem tedy není primárně se vyjadřovat k literárněvědné kategorii tématu či motivu, ale objektivně demonstrovat ty symptomy textové struktury (slova a kontexty), které na základě frekvenční příznakovosti logicky souvisí s problematikou ústřední tematiky textu. Samotné propojení kvantitativně-korpusové analýzy a literární vědy je *de facto* spojením sofistikovaného analytického nástroje a vlastní autonomní literárněvědné metody. Na druhé straně se nejedná pouze o pasivní využívání nástrojů korpusové lingvistiky. Uvědomíme-li si, že současné postupy a směry korpusové lingvistiky jsou založeny na strukturálních pozicích,¹⁷² nebude zdánlivý rozpor mezi kvantitativně-korpusovou a literárněvědnou analýzou natolik problematický, jak se na první pohled může jevit. Je ovšem nezbytné mít vždy na paměti, že kvantitativní a korpusová analýza je specifickým úkonem, který se bezprostředně týká pouze vlastní jazykové (lexikální) vrstvy uměleckého textu. Analýza jeho vyšších rovin bude záležitostí navazujícího (konkrétního) literárněvědného přístupu, ke kterému korpus může vytvářet vhodné funkční doplnění a zajištění.¹⁷³ Strukturálně-sémiotická

172 Srov.: „Její pojetí [korpusové lingvistiky – pozn. R. Z.], které se tak vyvíjí, v zásadě navazuje na základní strukturalistické názory na jazyk (vnímá se takto jako jisté pokračování či modifikace osvědčeného strukturalismu) a dnes se rysuje především i jako pojetí komplexní *metodologie* získávání a interpretace informací získaných z jazyka, protože staré cesty a způsoby už nestačí a selhávají.“ (Čermák 2011, s. 11–12)

173 „A přece je zřejmé, že samotný diskurs (jako soubor vět) je uspořádán a že právě svou uspořádaností se jeví jako sdělení jiného jazyka, vyššího, než je jazyk lingvistů: diskurs má své jednotky, svá pravidla, svou „gramatiku“: jako projev nadvětný, a třebaže složen výhradně z vět, diskurs se musí přirozeně stát předmětem nějaké jiné lingvistiky. Tato lingvistika diskursu nesla dlouho slavný název: rétorika; avšak v důsledku jakési historické souhry okolností rétorika se minula

analýza fikčního narativu se pochopitelně bude opírat o kvantitativní analýzu jazykové složky, která bude fungovat jako jeden z pilířů, jehož prioritní hodnotou je exaktnost. Tuto spolupráci si nelze představovat pasivně ve smyslu nekritického přejímání výsledků z jedné metodologické oblasti do druhé. Smyslem jejich propojení je nejen empiricky systematizovat a zpřesnit popis a analýzu jazykové (lexikální) vrstvy fikčního narativu, ale především extrapolovat tyto výsledky do vyšších strukturních „pater“ fikčního narativu. Kvantitativně-korpusová analýza uměleckých textů může nalézt uplatnění rovněž v otázkách textologických, v problematice geneze a atribuce textů, intertextuality apod.

Vrátíme-li se ještě jednou k základní otázce významu, lze konstatovat, že konkrétní slovní jednotky detekované v rámci korpusové analýzy musejí být ve výsledku zvažovány vzhledem ke svému komplexnímu narativnímu rámci, který formují kategorie vyššího řádu, např. narativní strategie či kompozice. Z tohoto důvodu, neboť se zde jedná o zvláštní typ „reference“ (viz Cohnová 2009), je vlastní kontextová analýza vyplývající z korpusu obzvláště důležitá. Na druhé straně nesmí být ani přeceněn její význam. Jednotlivé aspekty determinující smysl lexikální jednotky musejí být v umělecké fikci uvažovány v rámci celého narativu. V centru pozornosti tak ve výsledku stojí procesy sémiózy uvnitř fikčního narativu.¹⁷⁴

s krásnou literaturou, a protože ta se oddělila od lingvistiky, bylo třeba vrátit se v nedávné době k tomuto problému znovu: nová lingvistika diskursu se ještě nerozvinula, ale je alespoň požadována, a to i samými lingvisty. Tato skutečnost není bezvýznamná: třebaže tvoří objekt autonomní, diskurs musí být studován právě z hlediska lingvistiky; je-li třeba stanovit pracovní hypotézu pro analýzu, jejíž úkol je nesmírný a materiál nekonečný, bude nejrozumnější postulovat vztah homologie mezi větou a diskursem v té míře, že stejné formální uspořádání pravděpodobně řídí všechny sémiotické systémy, ať už je jejich materiál a rozsah jakýkoli: podle toho by diskurs byl veliká ‚věta‘ (jejímiž jednotkami nemusí nutně být věty), podobně jako věta je jistý specifický malý ‚diskurs‘. [...] Obecný jazyk vyprávění je zřejmě jen jeden z těch, které má lingvistika diskursu k dispozici, a spadá tudíž do hypotézy homologické: z hlediska strukturálního má vyprávění povahu věty. Přitom se nikdy nemůže redukovat na množinu vět: vyprávění je veliká věta, podobně jako každá vypovídací věta je v jistém smyslu náčrtem malého příběhu.“ (Barthes 2002, s. 12, 13)

[Pozn.: V této citaci zachováme původní znění překladu, tj. psaní slova *diskurs*.]

174 Sémantika literárního díla i jeho částí je samozřejmě spojena i s mimoliterárním polem; tato oblast ovšem přesahuje naši problematiku.

2.3.5 Strukturace a hierarchizace autorského korpusu

Konkrétní strategie vyhledávání v autorském korpusu bude řízena následujícími parametry, které vyplývají z dosavadního pojednání Čepova korpusu.

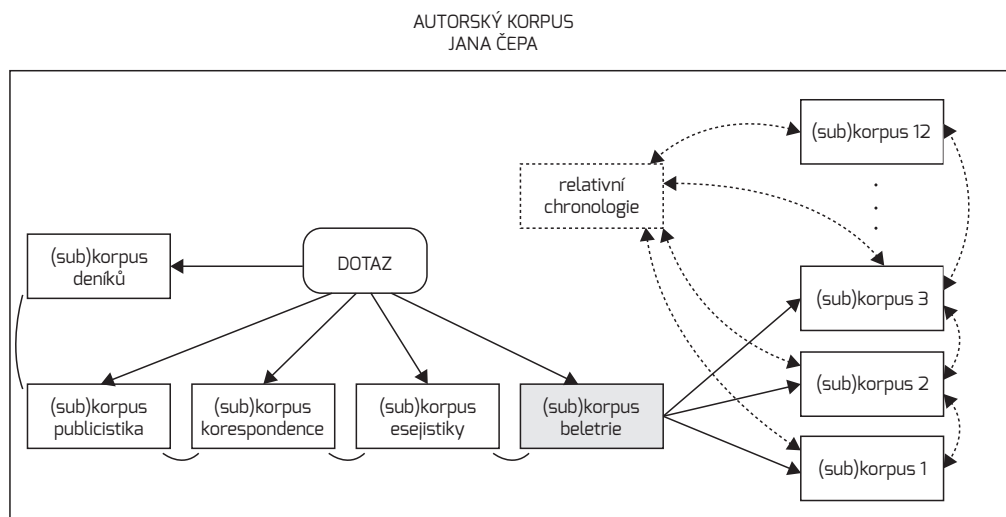


Schéma 9: Funkční strukturace autorského korpusu Jana Čepa.

Vlastní vyhledávání v autorském korpusu bude možné jak na úrovni jedné jednotky – lemmatu, slovního tvaru, morfologické kategorie, metajazykové informace (tzv. CQL dotaz) apod., tak na úrovni lexikálního syntagmatu. Výsledná a komplexní korpusová analýza bude ovšem dána teprve součinností všech aspektů, tzn. konkrétní jednotky, jejího kontextu, frekvence a příslušnosti k danému segmentu – sémantické třídě a (sub)korpusu, stejně jako ohledem na relativní chronologii (Schéma 10). Zcela zásadní úlohu má především srovnání výsledků se situací v jiných (sub)korpusech (Schéma 11). Výstupní korpusové hodnoty, které budou mít povahu hodnot absolutních, budou dále formálně strukturovány, systematizovány a taxonomizovány, aby byly funkčním předpokladem extrapolace do vyšších strukturních rovin uměleckého narativu.

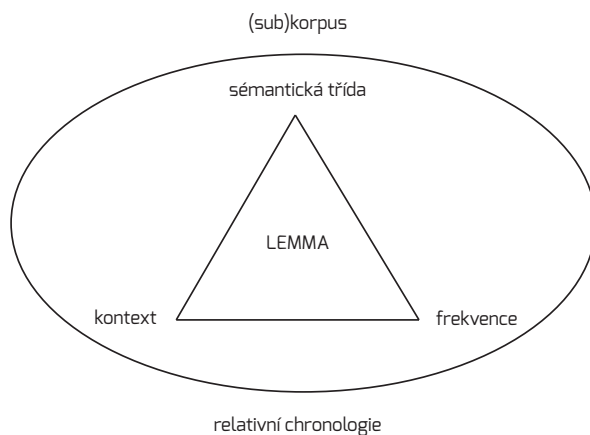


Schéma 10: Funkční propojení základních kritérií korpusové lexikální analýzy.

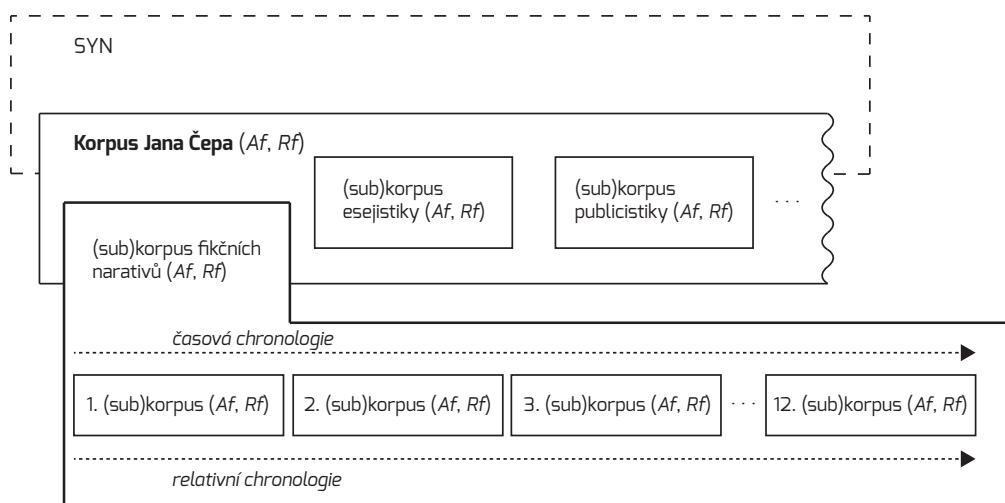


Schéma 11: Hierarchie autorského korpusu Jana Čepa.

2.4 Závěrem

Dosavadní výklad umožňuje učinit reálné předpoklady o způsobu, jakým bude možné s Čepovým korpusem a slovníkem pracovat. Jak jsme již konstatovali, jádrem celého autorského korpusu je tzv. korpus fikčních narativů, jehož koncepcí tvoří hlavní spojnici mezi kvantitativně-korpusovou metodologií a literárněvědnými aspekty. Základní způsob, kterým bude možné slovník využívat, spočívá v možnosti hierarchizovaného a systematického řazení lexikálních jednotek podle jednotlivých korpusových kritérií, která jsme zde prezentovali. Ve výsledku tedy bude možné analyzovat jazykovou situaci nejen uvnitř každého (sub)korpusu, ale především srovnávat dosažené výsledky mezi vybranými skupinami (sub)korpusů a také napříč celým (sub)korpusem beletrie Jana Čepa. Z takto získaných a dále filtrovaných dat bude možné vyvozovat některá obecná zjištění, jež jako signály umožní soustředit pozornost na vybrané literárněvědné problémy (např. tematické aspekty textu), jejichž interpretace nalezne oporu v exaktní kvantitativně-korpusové analýze. Zde ovšem musí platit zásadní stanovisko určující způsob práce s korpusem, na které jsme několikrát upozornili.

Autorský korpus, který zde navrhneme, není pouhou evidencí lexikálních komponentů jazykové vrstvy uměleckého díla, ale především obrazem strukturní povahy uměleckého díla jako znakového celku. Je samozřejmé, že jazyková rovina nemůže suplovat poznání díla v jeho komplexnosti, jak to ostatně jednoznačně vyjádřil také Roman Jakobson ve svém ohlédnutí za historií ruského formalismu,¹⁷⁵ avšak je rovněž jasné, že dílo je celistvostí, kterou určuje její strukturní strategie a způsob uspořádání vztahů mezi jednotlivými úrovněmi.¹⁷⁶ Daná skutečnost, kterou dostatečně explicitně vyjádřili především formalisté a strukturalisté, se dnes znovu významně projevuje v kvantitativně-korpusovém výzkumu orientovaném na problematiku diskurzu, jak například dokládá (mezi jinými) práce Johna F. Burrowse, který ve studii *Computers and the Study of Literature* (1992) konstatuje přínos kvantitativního přístupu v literární vědě, jenž spočívá v tom, že zdůrazňuje význam funkčního propojení mezi jednotlivými formativními rovinami uměleckého díla.¹⁷⁷

175 Viz Jakobson 2005, s. 80–81.

176 Zcela oprávněně lze komplexnost díla vztahovat i k dalším aspektům, mezi které náleží například dobová recepce. Naše stanovisko a hledisko je však z metodologických důvodů striktně zaměřeno na vnitřní strukturní vlastnosti díla.

177 Jmenovaný aspekt je implicitně přítomen také v následujícím Burrowsově konstatování: „In literary studies, as elsewhere, we examine and compare patterns. They may be sets of taxonomic differentiae, uniting certain texts in a genre or sub-genre or as moments in the evolution of a tradition or a school. They may be signs of individual authorship or of the larger cultural forces that inscribe themselves in and through the writings of individual authors. They may be large thematic patterns, established less often on the basis of direct statements in the text than by such inferences as underlie those readings of *King Lear* which emphasize the parallels between the main plot and the Gloucester sub-plot. They may lie in the juxtaposition of characters or the modulations of a narrative. They may, at their most concrete, consist in sequences

Lze tedy konstatovat, že kvantitativní obraz lexikální situace v uměleckém díle je jedním z přímých a explicitních dokladů jeho strukturních vlastností a zákonitostí. V případě analýzy jazykového plánu díla se z hlediska komplexní strukturní analýzy uměleckého textu samozřejmě jedná pouze o dílčí, nikoli komplexní reflexi. To je třeba mít neustále na paměti, aby se z vlastní kvantitativně-korpusové a statistické analýzy nestal samoúčelný nástroj literárněvědné dezinterpretace. Tím by taková analýza ztratila zcela svůj smysl. Lexikální (jazykovou) vrstvu uměleckého díla nelze vzhledem k jeho komplexní struktuře absolutizovat, nicméně podstatné je, že strukturní vlastnosti lexikální (jazykové) úrovně jsou funkčně provázány s vyššími segmenty a stavebními elementy fikčního narativu, které sama determinuje a je jimi determinována. Strukturně podmíněné vztahy mezi jednotlivými úrovněmi (vrstvami) fikčního narativu systematicky a přehledně formuloval Wolf Schmid ve svém generativním modelu fikčního narativu (viz Schéma 12).

of images, sets of tropes, poetic rhythms, shifts of semantic register, or even (as I shall go on to argue) the frequency patterns of very common words. They may also, though it is not fashionable to say so, help clarify our understanding of what we value and so contribute to the assessment of artistic excellence or inadequacy." (Burrows 1992, s. 168–169)

[„V literární vědě, podobně jako jinde, hodnotíme a srovnáváme schémata. Těmi mohou být soubory taxonomických diferencí sdružující některé texty do žánru nebo subžánru nebo do vývojových řad určité tradice nebo školy. Mohou jimi být znaky individuálního autorského stylu nebo širších kulturních sil, které se podepsaly na způsobu psaní jednotlivých autorů. Mohou jimi být velké tematické předlohy, jež se často neutvářejí na základě přímých konstatování v textu, ale spíše na základě důsledků, které posilují takové čtení *Krále Leara*, jež zdůrazňuje paralelitu mezi hlavní osnovou a gloucesterskou subosnovou. Takovými modely mohou být srovnávání postav nebo modulace narativu. Nejkonkrétněji jimi mohou být vztahy v řadě zobrazení, soubory tropů, básnických rytmtů, posunů v sémantickém registru, nebo dokonce (jak se pokusím doložit) frekvenční modely nejčastěji se vyskytujících slov. Mohou tedy, ačkoli není módní to tvrdit, pomoci objasnit způsob našeho chápání toho, čeho si ceníme, a tím přispět k posouzení umělecké dokonalosti nebo nedostatečnosti.“]

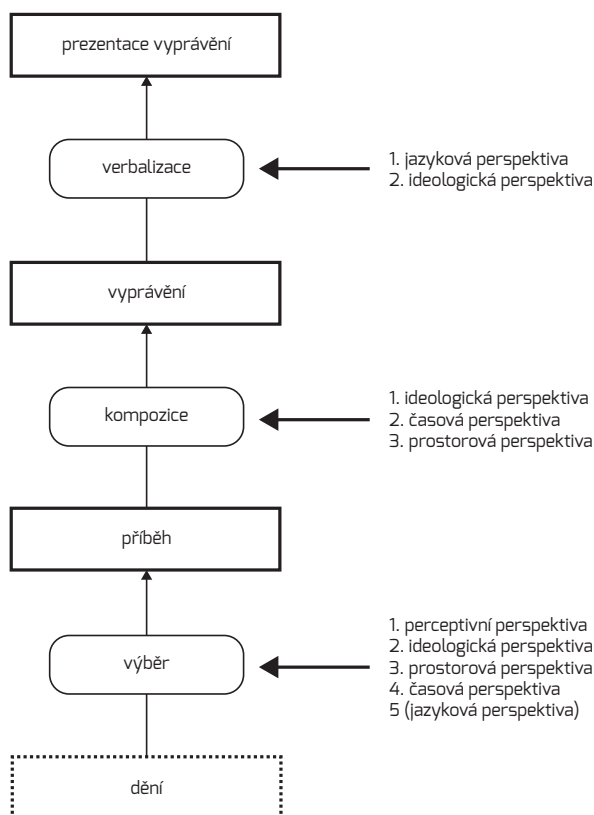


Schéma 12: Schmidův generativní model perspektivy narativu (Schmid 2004, s. 45).

Schmid konstatuje, že ačkoli jsou jednotlivé segmenty v modelu odděleny (fázovány), ve skutečnosti koexistují v homogenní struktuře a realizují se současně.¹⁷⁸ Můžeme tedy uvažovat, že každá z rovin disponuje vlastními a specifickými strukturními vazbami, avšak vzájemně se k sobě funkčně i významově vztahují a současně se determinují.¹⁷⁹ Kvantitativně-korpusová analýza lexikální (jazykové) vrstvy, na rozdíl od problematických možností analýzy jiných vyšších oblastí uměleckého narativu, poskytuje velkou

¹⁷⁸ „Generativní model nepostihuje, jak již bylo řečeno, ani skutečný proces vzniku díla ani proces jeho recepcce, nýbrž modeluje s pomocí časových metafor ideální, atemporální genezi narativního díla s cílem vydělit postupy řídící vyprávění a ukázat je v jejich logických vztazích. [...] Jedinou rovinou přístupnou našemu pozorování je *text narativního díla*. Všechny ostatní roviny jsou abstrakce a konstrukty.“ (Schmid 2004, s. 47)

¹⁷⁹ Tyto vztahy ostatně vyplývají i z uvedeného Schmidova modelu, kde například prostorová perspektiva je dána jak na úrovni výběru, tak kompozice; ideologická perspektiva jako komplexní kategorie zasahuje roviny výběru, kompozice i verbalizace apod. Schmid uvádí, že jeho model je abstrakcí; můžeme jej pochopitelně modifikovat a kupříkladu promýšlet možnosti prostorových nebo časových relací nejen na úrovni výběru a kompozice, ale stejně tak i na úrovni verbalizace reprezentované příslušným lexikem apod.

výhodu spočívající v exaktnosti. Epistemologicky je potom koherentní se strukturálně-sémioticky zaměřenou literárněvědnou metodou.

Využití korpusových, statistických a kvantitativních metod neslouží pouhému záznamu a evidenci lexikálních kvalit uměleckého textu či díla, ale především snaze vybudovat výchozí přesnou a objektivní bázi, jež by umožnila funkční extrapolaci do vyšších struktur textu (viz Schéma 13).

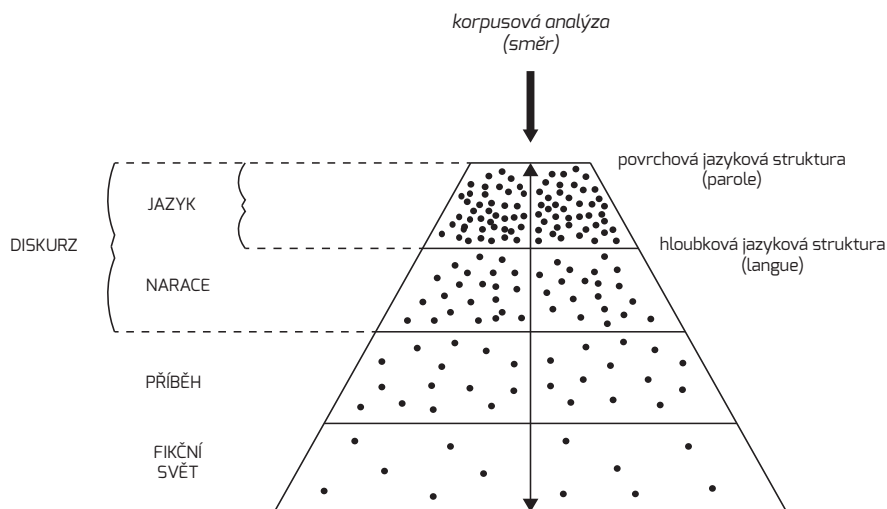


Schéma 13: Základní strukturální segmenty uměleckého narativu. Hustota vyznačuje míru reference, kterou je korpus schopen k určité rovině poskytnout. S narůstající abstrakcí se snižuje referenční spolehlivost korpusu.

Z uvedeného ovšem vyplývá, že výsledky kvantitativní a korpusové analýzy stejně jako statistická měření je nezbytné náležitě interpretovat. Samostatná data v podobě číselných hodnot jsou pouze předpokladem pro interpretaci založenou na exaktním pozorování příslušného jevu, nikoli finální odpovědí na kladené otázky. Ve výsledku je proto důležitá metodologická kooperace exaktních kvantitativně-korpusových technik s příslušnými metodologicky orientovanými přístupy v literární vědě, např. s naratologickými či sémiotickými. Výsledky formální analýzy by měly být interpretovány teprve až v rámci příslušného literárněvědného kontextu; naopak jako zcela nedostačující se jeví pouze jejich prostý výčet či prezentace bez dalšího výkladu propojeného s konkrétními literárněvědnými požadavky.¹⁸⁰

¹⁸⁰ Totéž pochopitelně platí i v lingvistice.

3. První tři Čepovy povídkové knihy z hlediska kvantitativně-korpusové analýzy tematických oblastí

Literární věda má k dispozici řadu vlastních metodologických postupů, jimiž lze problematiku tematiky uměleckého díla zkoumat.¹⁸¹ Následující oddíl se hodlá věnovat otázce, jakým způsobem lze k uvedené problematice přispět z perspektivy kvantitativně-korpusové analýzy. V pravém slova smyslu ovšem nepůjde o přeformulování toho, co literární věda chápe pod pojmem *téma*, ale o snahu za pomoci speciálních formalizovaných postupů analýzy autorského korpusu konceptualizovat možnosti funkčního propojení kvantitativně-korpusového a literárněvědného aspektu. Ve výsledku pak naše úvahy nebudou směřovat výhradně do oblasti tematiky, ale dotknou se i některých dalších literárněvědných kategorií a otázek, jež jsou s existencí literárního uměleckého díla spojeny – zejména kategorií vyprávěče, času a prostoru.

V této kapitole hodláme demonstrovat jeden z možných přístupů kvantitativně-korpusové analýzy uměleckého textu, který vychází z měření tzv. *tematické koncentrace textu*. Postup, který teoreticky formuloval Ioan-Iovitz Popescu, byl testován rovněž na některých českých autorech.¹⁸² V našem případě ovšem nepůjde výhradně o aplikaci této metody, tzn. o demonstrování vlastního analytického postupu, ale budeme se také ptát, jaké možnosti poskytuje metoda literární vědy a jak s ní lze v kontextu této disciplíny pracovat. Význam kvantitativní analýzy opírající se o měření tematické koncentrace textu je širší než pouhé stanovení rozhraní mezi dvěma frekvenčními pásmy, které vymezuje distribuční oblast příznakového lexika. Ve spojení s korpusovou analýzou překračuje hranice dílčího postupu a vybízí k obecnějšímu zamyšlení, jehož součástí je otázka, nakolik je možné také v literární vědě (tedy nejen v lingvistice) využívat exaktní postupy dosud výhradně spojené s moderní kvantitativní a korpusovou lingvistikou a nakolik je takový přístup schopen obohatit či alespoň významně doplnit literárněvědný výzkum.

Materiál, na kterém hodláme uvedený postup předvést, tvoří první tři povídkové soubory Jana Čepa: *Dvojí domov* (1926), *Vigilie* (1928) a *Zeměžluč* (1931). Tyto knihy

181 Vedle strukturální analýzy se jedná například o interpretaci fenomenologickou, sémiotickou, archetypální či psychologickou. Tento výčet lze rozšířit o široký seznam dalších možných přístupů, které vyrůstají z konkrétně orientovaného literárněvědného nebo obecně kulturního bádání, např. gender studies, spatial turn, postkolonialismus apod.

182 Davidová-Glogarová – David – Čech (2013), Davidová-Glogarová – Čech (2013), David – Čech – Davidová-Glogarová – Radková – Šustková (2013).

nerepresentují pouze konkrétní chronologickou linii počáteční fáze Čepovy tvorby, jejich vztah je mnohem komplikovanější. Kniha *Zeměžluč* totiž obsahuje tři oddíly, z nichž první dva tvoří podstatně přepracované verze prvních dvou výše jmenovaných titulů (viz Tab. 2). Jak dokládá situace prvních Čepových povídkových souborů, nelze v jejich případě uvažovat výhradně o časové následnosti, kterou tvoří roky prvního vydání, ale je nezbytné na danou situaci nahlížet rovněž jako na komplexní jev (událost). Autor se ke svým předchozím textům vracel a s odstupem času znovu reflektoval podobu prvních souborů, jež neváhal ve značně upravené verzi včlenit do nově vznikající knihy *Zeměžluč*. To samozřejmě předpokládá určitý autorský záměr. Situace prvních Čepových knižních titulů a jejich podob je tedy projevem nejen časové následnosti, ale i návratnosti. Uvedené knihy tak společně vytvářejí úzce propojený a komponovaný celek, na jehož pozadí se manifestuje autorova tvůrčí intence.

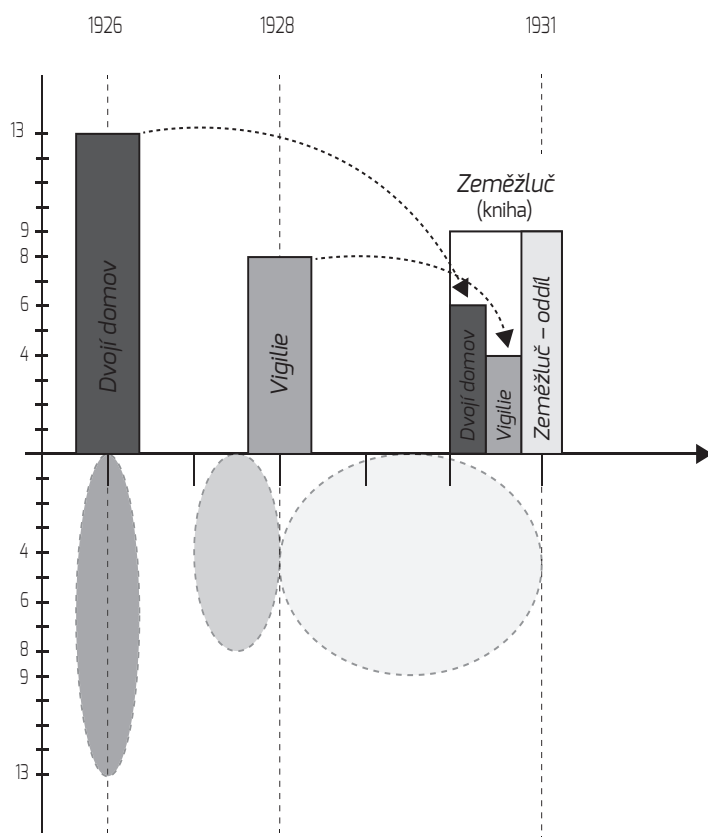


Schéma 14: Chronologická situace Čepových prvních povídkových souborů. Vodorovná osa zobrazuje časové průběhy vzniku jednotlivých povídkových souborů, tzv. relativní chronologii, na svislé ose je vyznačen počet povídek.

Jedna z otázek, které si literární věda může klást, je vzájemný vztah mezi jednotlivými částmi této rané fáze Čepovy tvorby. Co vedlo autora k tomu, že ve druhém vydání *Dvojího domova* (1931) a *Vigilií* (1931) některé povídky vyřadil? Nezbytně se musíme ptát, jaký je vztah mezi jednotlivými povídkovými soubory a jejich variantami. Analýza uvedených vztahů, na jejichž pozadí se pokusíme zhodnotit rámcovou situaci Čepovy rané povídkové tvorby, bude prováděna pomocí dat, která jsme získali vytvořením autorského korpusu Jana Čepa, jenž obsahuje veškerou Čepovou prozaickou tvorbu.¹⁸³ Abychom získané statistické a korpusové hodnoty dokázali relevantně interpretovat, což v našem případě znamená učinit tak primárně s ohledem na literárněvědné aspekty, vycházíme ze strategie tzv. relativní chronologie, která zachycuje relativní průběhy vzniku jednotlivých povídkových souborů. Jak jsme již konstatovali, relativní chronologie poskytuje přehled o vzniku a časové situovanosti jednotlivých textů a povídkových souborů s tím, že umožňuje danou situaci zobrazit nikoli punktálně, ale fázově (viz Schéma 4). Modelace na ose B je schopna v rámci nutné formalizace zobrazit dynamický vztah jednotlivých oblastí Čepovy tvorby. Dovolí mimo jiné sledovat podobnosti a rozdíly souborů, které se překrývají, a ověřovat dané situace na korpusových výsledcích za pomoci kvantitativní analýzy. Ze schématu je zřejmé, že tento model umožňuje klást si celou řadu dílčích otázek, které vyplývají z konkrétní situace určité vývojové fáze, a současně dovoluje sledovat celkovou situaci jako lineárně komplikovaný (nesamozřejmý) tvůrčí proces. Náš přístup spojuje dva základní aspekty náhledu na analyzovaný materiál:

1. externí kritérium vyplývající z časové charakteristiky, zejména relativní chronologie,
2. interní kritérium, které je určeno imanentními strukturními vztahy uvnitř povídkových souborů.

3.1 Metodologie

Z celkového chronologického schématu Čepovy beletrie (viz Schéma 4) je patrné, že analýza může být prováděna nejen s ohledem na celé autorovo prozaické dílo, ale i v kontextu jeho jednotlivých vývojových fází. Pochopitelně teprve až komplexní analýza celého díla je schopna přinést obecnější a ucelenější závěry, což ovšem nikterak nevyklučuje dílčí sondy a analýzy, které jsou nezbytné pro relevantní poznání celkové situace

¹⁸³ Aktuálně Čepův korpus obsahuje pouze beletrii. Do budoucna se předpokládá realizace celého Čepova autorského korpusu, který bude kromě beletrie zahrnovat autorovu esejistiku, korespondenci, deníky a drobnější publicistické texty rozdílné žánrové povahy.

autorovy tvorby. Jak bylo konstatováno, jednou z metod, které jsou vhodné pro analýzu korpusového materiálu s ohledem na jeho potenciál v rámci literárněvědného bádání, je měření tematické koncentrace textu, které vychází z definování tzv. *h-bodu*, jenž tvoří předěl mezi dvěma částmi kvantitativně segmentovaného a uspořádaného lexikonu.

„*h*-bod je místo, které odděluje dvě sousedící oblasti frekvenční distribuce slov: udává vztah mezi pořadím slova dle jeho frekvence r a jeho frekvencí f_r . Řešením dvou souběžných rovnic dostaneme definici *h*-bodu

$$h = \begin{cases} r, & \text{pokud } r = f(r) \\ \frac{f(i)r_i - f(j)r_j}{r_j - r_i + f(i) - f(j)}, & \text{pokud } r \neq f(r). \end{cases}$$

h-bod je tedy místo, pro které platí $r = f(r)$. Pokud takový bod neexistuje, použijeme dvě sousedící frekvence $f(i)$ a $f(j)$ tak, aby platilo $f(i) > r_i$ a $f(j) < r_j$. Většinou platí $r_i + 1 = r_j$.“

(Popescu – Mačutek – Altmann 2014, s. 37)

Pro stanovení uvedené hodnoty je zprvu nutné sestavit frekvenční seznam příslušného lexikonu. Následující tabulka, která je pouze částí úplného frekvenčního seznamu Čepova prvního vydání *Dvojího domova* (1926), je rozdělena do tří sloupců, které představují pořadí (r), lemma a frekvence (Af).¹⁸⁴

r	lemma	Af	r	lemma	Af	r	lemma	Af	r	lemma	Af
1	a	1011	16	jít	109	31	o	57	46	ještě	45
2	se	882	17	jako	105	32	stát	54	47	tak	45
3	být	582	18	svůj	104	33	ty	53	48	pak	45
4	na	361	19	po	98	34	matka	53	49	vidět	44
5	on	264	20	už	97	35	moci	53	50	já	43
6	v	252	21	oko	93	36	co	53	51	muset	43
7	s	206	22	mít	88	37	okno	52	52	pod	43
8	do	193	23	oni	85	38	u	51	53	jeho	43
9	ten	165	24	jenž	80	39	otec	50	54	starý	42
10	ale	137	25	který	76	40	nad	50	55	cesta	42
11	ona	134	26	hlava	75	41	před	50	56	žena	41
12	z	133	27	jak	70	42	den	50	57	i	41
13	k	121	28	když	68	43	Rudolf	47	58	člověk	40
14	že	116	29	však	63	44	od	46	59	Jeník	40
15	za	112	30	ruka	58	45	Ludvík	45	60	její	39

Tabulka 11: Prvních 60 lemmat Čepova *Dvojího domova* (1926) řazených sestupně dle klesající absolutní frekvence (Af).

¹⁸⁴ Absolutní frekvence – celkový počet všech tvarů příslušného slova v daném (sub)korpusu.

Pro stanovení *h-bodu* je možné v tomto případě použít první pravidlo, které platí ve chvíli, kdy existuje identická korelace mezi *r* a *Af*. V našem případě tedy $h = 45$. Jiná je však situace ve druhém (sub)korpusu *Vigilie* (1928).

<i>r</i>	lemma	<i>Af</i>	<i>r</i>	lemma	<i>Af</i>	<i>r</i>	lemma	<i>Af</i>	<i>r</i>	lemma	<i>Af</i>
1	a	1111	16	za	122	31	oko	66	46	země	45
2	se	886	17	ona	110	32	cesta	61	47	cítit	44
3	být	655	18	ale	110	33	nad	58	48	však	44
4	v	381	19	už	106	34	její	54	49	co	44
5	na	338	20	který	101	35	tvář	53	50	moci	43
6	s	223	21	po	90	36	chvíle	49	51	ještě	42
7	on	212	22	tak	88	37	všecek	49	52	bez	41
8	ten	191	23	já	82	38	jeho	49	53	matka	41
9	jenž	183	24	mít	77	39	člověk	48	54	u	41
10	do	164	25	když	77	40	stát	48	55	ruka	40
11	z	144	26	jak	75	41	až	47	56	starý	40
12	že	144	27	jít	74	42	pak	47	57	den	40
13	jako	144	28	oni	72	43	od	47	58	hlava	38
14	svůj	136	29	tento	69	44	pod	46	59	všechn	38
15	k	128	30	Rozárka	68	45	před	46	60	jen	37

Tabulka 12: Prvních 60 lemmat Čepových *Vigilií* (1928) řazených sestupně dle klesající absolutní frekvence (*Af*).

V uvedeném případě neplatí předchozí vztah $r = Af$, proto použijeme alternativní výpočet:

$$r_i = 45 \quad f(i) = 46$$

$$r_j = 46 \quad f(j) = 45$$

$$h = \frac{(46 \times 46) - (45 \times 45)}{(46 - 45) + (46 - 45)} = \frac{91}{2} = 45,5$$

Analogicky postupujeme v případě oddílu *Zeměžluč* (1931)¹⁸⁵ i celého souboru, tj. celé knihy téhož jména tvořené druhými verzemi *Dvojího domova* (1931) a *Vigilií* (1931). Smyslem *h-bodu* je vymezit hranici, která rozděluje lexikum na oblast s převažujícím potenciálem synsémantik (nad *h-bodem*) a autosémantik (pod *h-bodem*). Přesto vidíme, že i nad tímto bodem se objevují některá autosémantika. Výskyt těchto autosémantik v uvedené části frekvenčního seznamu je z hlediska jejich obvyklé distribuce příznakový, a svědčí tak o jejich významném postavení v rámci slovesného strukturního systému textu nebo

¹⁸⁵ Pracovně budeme tento (sub)korpus označovat *Zeměžluč – oddíl* (1931). Pojmenování (sub)korpusu *Zeměžluč – celý soubor* (1931) zahrnuje všechny tři části knihy *Zeměžluč*.

díla. Sestavíme-li samostatný seznam plnovýznamových slov, situace u prvního Čepova (sub)korpusu vypadá následovně (viz Tab. 13).

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	jít	109	4999,31
2	oko	93	4265,47
3	hlava	75	3439,89
4	ruka	58	2660,18
5	matka	53	2430,86
6	okno	52	2384,99
7	otec	50	2293,26
8	den	50	2293,26
9	Rudolf	47	2155,67
10	Ludvík	45	2063,94

Tabulka 13: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Dvojí domov* (1926).

Kromě absolutní frekvence (*Af*) jsme do tabulky připojili hodnoty relativní frekvence (*Rf*), která je nezbytná pro adekvátní srovnání frekvencí lexikálních jednotek z různých a nestejně velkých korpusů. Relativní frekvence je normovanou hodnotou, jejíž výpočet je v tomto případě stanoven s ohledem na parametry standardně používané ČNK, kde je relativní frekvence přepočítána na milion slov (instances per milionem – i. p. m).

$$Rf = Af \times \frac{1000000}{N}$$

Symbol *N* ve vzorci označuje celkový počet všech jednotek v daném korpusu, tzn. v tomto případě velikost Čepova (sub)korpusu *Dvojí domov* (1926).¹⁸⁶ Analogicky budeme postupovat při sestavení dalších seznamů tematických slov, což se týká (sub)korpusů *Vigilie* (1928), *Dvojí domov* (1931), *Vigilie* (1931) a *Zeměžluč – oddíl* (1931) a *Zeměžluč – celý soubor* (1931). Pro přehlednější orientaci následně převedeme číselné hodnoty *Rf* do grafů.

¹⁸⁶ (Sub)korpus *Dvojí domov* (1926) má celkem 21 803 pozic. Celý autorský korpus fikčních narativů má 511 026 pozic.

Vigilie (1928)

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	jít	74	3247,32
2	Rozárka	68	2984,03
3	oko	66	2896,26
4	cesta	61	2676,85
5	tvář	53	2325,79
6	chvíle	49	2150,25
7	člověk	48	2106,37

Tabulka 14a: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Vigilie* (1928).

Zeměžluč – oddíl (1931)

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	oko	103	4464,29
2	tvář	65	2817,27
3	hlava	60	2600,55
4	ruka	57	2470,53
5	srdce	54	2340,50
6	cítit	52	2253,81

Tabulka 14b: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Zeměžluč – oddíl* (1931).

Dvojí domov (1931)

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	jít	38	5174,29
2	oko	28	3812,64

Tabulka 14c: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Dvojí domov* (1931).

Vigilie (1931)

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	Rozárka	68	4259,05
2	oko	46	2881,12
3	tvář	44	2755,86
4	jít	43	2693,22
5	matka	39	2442,69
6	ruka	37	2317,42

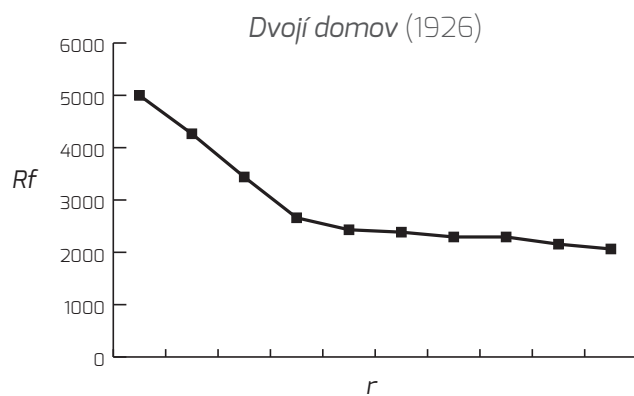
Tabulka 14d: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Vigilie* (1931).

Zeměžluč – celý soubor (1931)

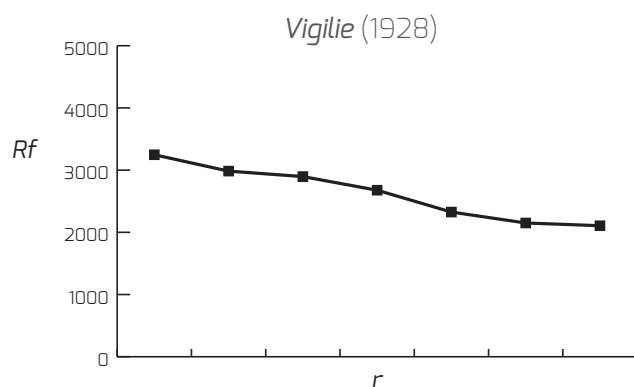
<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	oko	177	3816,14
2	tvář	112	2414,73
3	ruka	107	2306,93
4	hlava	107	2306,93
5	cesta	90	1940,41
6	člověk	86	1854,17
7	vědět	86	1854,17

8	cítit	85	1832,61
9	den	82	1767,93
10	matka	81	1746,37
11	chvíle	79	1703,25
12	srdce	75	1617,01
13	starý	73	1573,89
14	Rozárka	68	1466,09
15	svět	65	1401,41

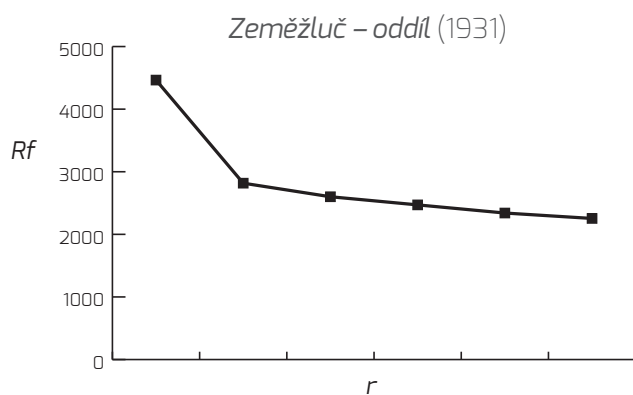
Tabulka 14e: Autosémantika (nad *h-bodem*) v Čepově (sub)korpusu *Zeměžluč – celý soubor* (1931).



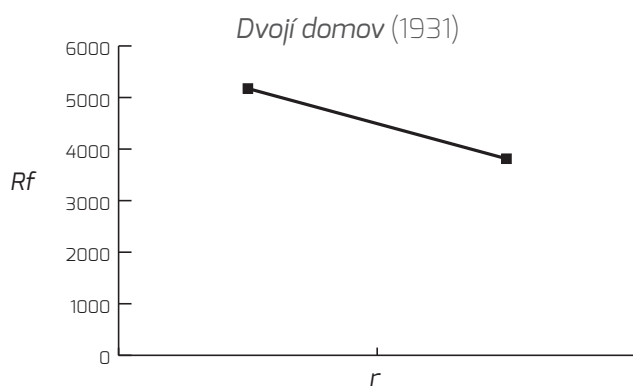
Graf 1a: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Dvojí domov* (1926).



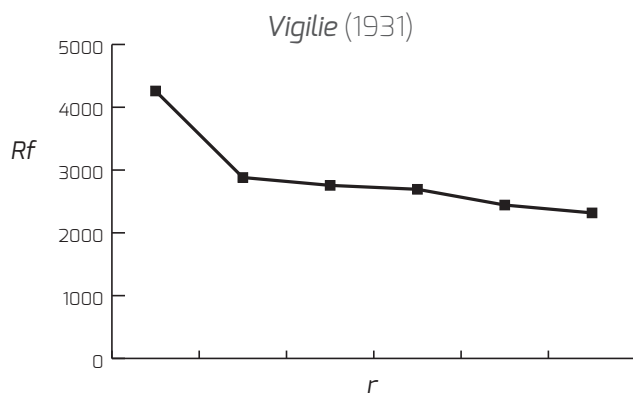
Graf 1b: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Vigilie* (1928).



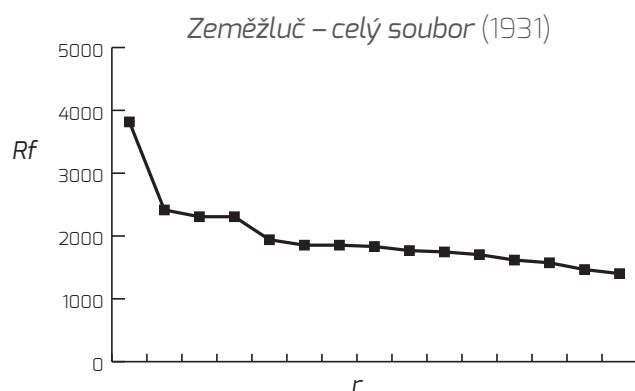
Graf 1c: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Zeměžluč – oddíl (1931)*.



Graf 1d: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Dvojí domov (1931)*.



Graf 1e: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Vigilie (1931)*.



Graf 1f: Grafické znázornění průběhů relativních frekvencí autosémantik nad *h-bodem* u (sub)korpusu *Zeměžluč – celý soubor* (1931).

Tato zobrazení nám poskytují první srovnání, která jsou založena na výběru tematických slov a jejich relativních frekvencí (*Rf*). Interpretaci kvantitativních i kvalitativních aspektů se budeme věnovat později. V danou chvíli je třeba vymezit ostatní hlavní kritéria, která nás budou zajímat a která využijeme v pokročilé analýze tematické oblasti a jejích prvků (lemmat). Jedná se o následující aspekty:

1. počet autosémantik v oblasti nad *h-bodem*;
2. disperze tematických oblastí (*DTO*);
3. typ autosémantik a jejich distribuce do sémantických tříd (paradigmatika);
4. vzájemné sémantické vztahy těchto tříd;
5. kontexty (syntagmatika).

Zvolená kritéria jsou stanovena tak, aby rozšířila referenční platnost vlastní analýzy tematické koncentrace a pokusila se v rámci této analýzy postihnout širší strukturní výšeč příslušného textového útvaru. Pomocí uvedených parametrů a především jejich následnými interpretacemi lze analyzovat a srovnávat distribuci a povahu tematických oblastí Čepových (sub)korpusů.¹⁸⁷ Grafické znázornění je přehlednou a systematickou transformací frekvenčních údajů. Pro jejich přehlednější komparaci zavádíme další hodnotu, tzv. *disperzi tematické oblasti* (*DTO*). Jejím smyslem je numericky vyjádřit vzdálenosti mezi krajními frekvenčními hodnotami vymezené tematické oblasti. Pro výpočet je stanoven jednoduchý vzorec:

¹⁸⁷ Jak bylo uvedeno v předchozím oddíle, v rámci celého autorského korpusu fikčních narativů je možných kombinací celá řada. Lze se například zaměřit na rozdíly mezi tvůrčími fázemi 20. a 30. let nebo na souvislosti mezi oblastmi, které na ose relativní chronologie signalizují významné průniky apod.

$$DTO = \frac{Rf(y)}{Rf(x)} (\times 100 \%)$$

kde $Rf(y)$ je nejnížší hodnota relativní frekvence příslušného autosémantika nad *h-bodem* a $Rf(x)$ naopak hodnota nejvyšší. Čím je hodnota *DTO* větší, resp. čím více se blíží 1 (100%), tím je rozptyl mezi autosémantiky menší. Hodnota tohoto čísla upozorňuje na povahu této vymezené oblasti. Je-li nižší, lze se domnívat, že vztahy mezi tematickými slovy budou nerovnoměrné, což může ve srovnání s jejich grafickým průběhem ukazovat na příznakovost konkrétních lexémů a jejich sémantických tříd a *vice versa*.

(sub)korpus	DTO
<i>Dvojí domov</i> (1926)	0,41
<i>Vigilie</i> (1928)	0,65
<i>Zeměžluč – oddíl</i> (1931)	0,50
<i>Dvojí domov</i> (1931)	0,74
<i>Vigilie</i> (1931)	0,54
<i>Zeměžluč – celý soubor</i> (1931)	0,37

Tabulka 15: Hodnoty *DTO* u sledovaných (sub)korpusů.

3.2 Vyhodnocení

Frekvenční analýza není pouhou formální kvantifikací, ale přirozeně obsahuje i rozměr kvalitativní, který je dán již výběrem konkrétního lexika, jež se v daném případě vyskytuje nad *h-bodem*. Skutečnost, že vzájemný poměr obou aspektů (kvantitativní a kvalitativní) není věcí náhody, ale projekcí imanentních strukturních a sémantických pravidel příslušného textu, dokládá i následující test, ve kterém jsme k frekvenčním hodnotám tematických slov u jmenovaných tří Čepových (sub)korpusů připojili hodnoty relativních frekvencí ze tří zcela odlišných (sub)korpusů. První, který zahrnuje vybrané a korpusově zpracované dílo Karla Čapka, jež vzniklo v průběhu 20. let 20. století, jsme označili jako ČAPEK 1, (sub)korpus ČAPEK 2 zahrnuje vybranou tvorbu téhož autora z 30. let. Následně jsme hodnoty srovnávali s referenčními korpusy SYN2010:BEL¹⁸⁸ a SYN2010.¹⁸⁹

¹⁸⁸ Korpus SYN2010:BEL je korpusem české beletrie, který je součástí SYN2010. Struktura Českého národního korpusu je přístupná zde: <<http://ucnk.ff.cuni.cz/struktura.php>>.

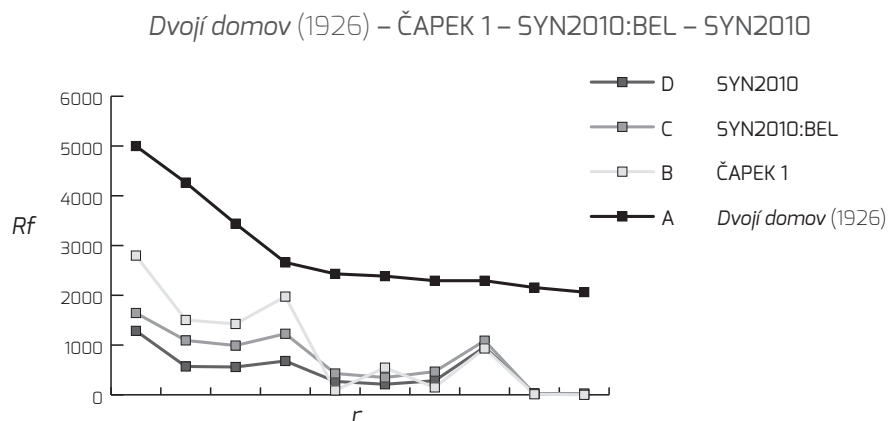
¹⁸⁹ Korpus SYN2010 je českým synchronním referenčním korpusem, který vzniká na pracovišti Českého národního korpusu v Praze. Dostupné z WWW: <<http://ucnk.ff.cuni.cz/struktura.php>>.

Jan Čep	ČAPEK 1	SYN2010:BEL	SYN2010
<i>Dvojí domov</i> (1926)	<i>Továrna na absolutno</i> (1922)		
<i>Vigilie</i> (1928)	<i>Krakatit</i> (1922)		
	<i>Povídky z jedné a druhé kapsy</i> (1929)		
Jan Čep	ČAPEK 2	SYN2010:BEL	SYN2010
	<i>Hordubal</i> (1933)		
	<i>Povětroň</i> (1934)		
<i>Zeměžluč – oddíl</i> (1931)	<i>Obyčejný život</i> (1934)		

Tabulka 16: Srovnávané (sub)korpusey.

	<i>Dvojí domov</i> (1926)			ČAPEK 1	SYN2010:BEL	SYN2010
<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>
1	jít	109	4999,31	2798,36	1644,20	1285,68
2	oko	93	4265,47	1504,17	1093,59	569,57
3	hlava	75	3439,89	1424,00	990,25	559,25
4	ruka	58	2660,18	1973,74	1227,33	679,97
5	matka	53	2430,86	87,81	428,42	269,29
6	okno	52	2384,99	545,93	347,38	212,18
7	otec	50	2293,26	148,89	466,46	283,49
8	den	50	2293,26	931,51	1088,58	960,28
9	Rudolf	47	2155,67	11,45	20,79	26,61
10	Ludvík	45	2063,94	0,00	23,36	20,75

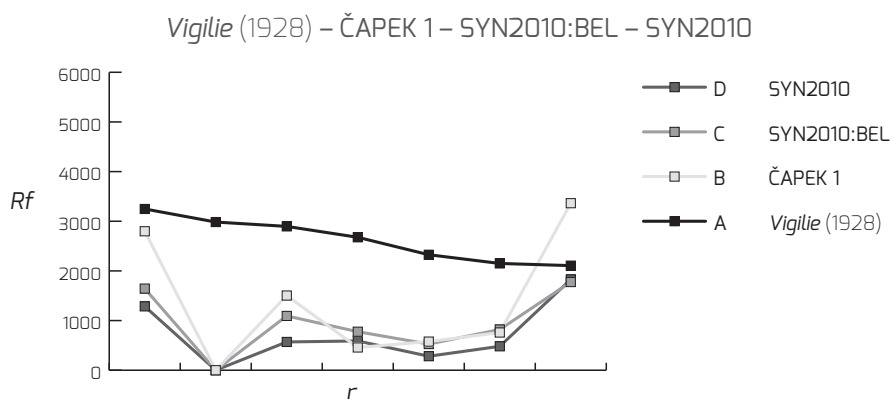
Tabulka 17a: Hodnoty relativních frekvencí ve vybraných (sub)korpusech.



Graf 2a: Grafický průběh relativních frekvencí ve vybraných (sub)korpusech.

	<i>Vigilie</i> (1928)			ČAPEK 1	SYN2010:BEL	SYN2010
<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>
1	jít	74	3247,32	2798,36	1644,20	1285,68
2	Rozárka	68	2984,03	0,00	0,54	0,42
3	oko	66	2896,26	1504,17	1093,59	569,57
4	cesta	61	2676,85	458,12	775,99	589,16
5	tvář	53	2325,79	576,47	526,42	281,85
6	chvíle	49	2150,25	759,72	820,20	481,40
7	člověk	48	2106,37	3363,38	1775,62	1827,81

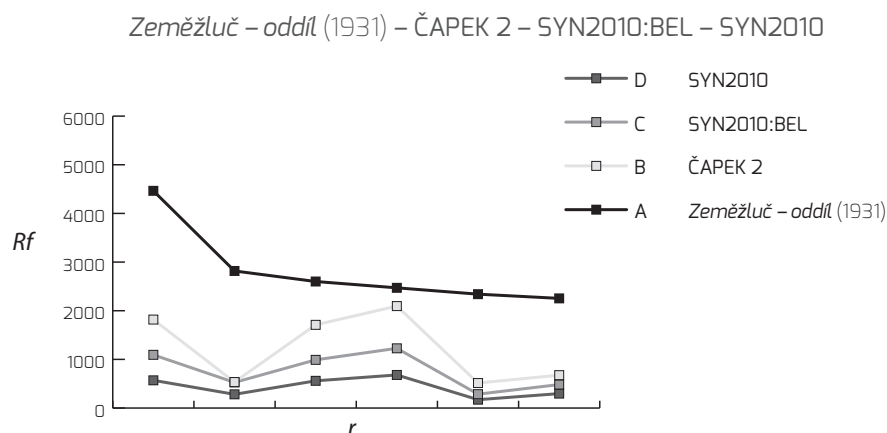
Tabulka 17b: Hodnoty relativních frekvencí ve vybraných (sub)korpusech.



Graf 2b: Grafický průběh relativních frekvencí ve vybraných (sub)korpusech.

	<i>Zeměžluč – oddíl</i> (1931)			ČAPEK 2	SYN2010:BEL	SYN2010
<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>	<i>Rf</i>
1	oko	103	4464,29	1818,12	1093,59	569,57
2	tvář	65	2817,27	533,95	526,42	281,85
3	hlava	60	2600,55	1709,98	990,25	559,25
4	ruka	57	2470,53	2095,23	1227,33	679,97
5	srdce	54	2340,50	513,67	285,05	171,77
6	cítit	52	2253,81	675,88	482,63	296,56

Tabulka 17c: Hodnoty relativních frekvencí ve vybraných (sub)korpusech.



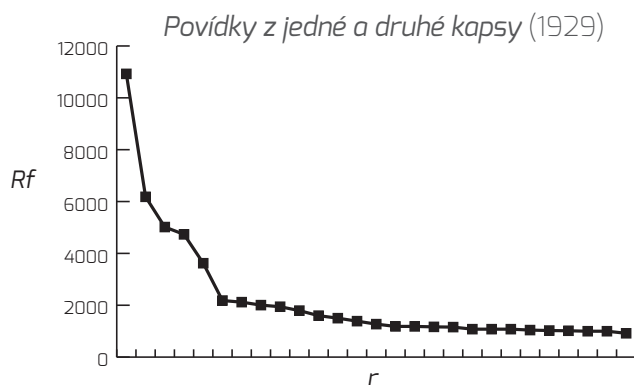
Graf 2c: Grafický průběh relativních frekvencí ve vybraných (sub)korpusech.

Jiným dokladem jedinečné symptomatickosti tematických slov, jejich distribuce a frekvenčních průběhů v Čepových (sub)korpusech je srovnání s tematickými oblastmi v jiném autorském (sub)korpusu. Jako příklad jsme utvořili frekvenční seznam lemmat, která se vyskytují nad *h-bodem* v (sub)korpusu *Povídky z jedné a druhé kapsy* (1929) Karla Čapka. Jen při zběžném porovnání výsledků, tj. volby lexika a frekvenčního průběhu, jsou zřejmé rozdíly mezi oběma autory.¹⁹⁰

<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>	<i>r</i>	lemma	<i>Af</i>	<i>Rf</i>
1	pan	1237	10924,18	15	případ	134	1183,38
2	řící	700	6181,83	16	příjit	134	1183,38
3	vědět	568	5016,12	17	prosit	132	1165,72
4	člověk	536	4733,52	18	starý	131	1156,89
5	jít	410	3620,79	19	rok	122	1077,41
6	vidět	247	2181,30	20	doktor	122	1077,41
7	paní	240	2119,49	21	rozumět	122	1077,41
8	dát	227	2004,68	22	ruka	118	1042,08
9	říkat	220	1942,86	23	den	116	1024,42
10	povídat	203	1792,73	24	dostat	115	1015,59
11	dělat	181	1598,45	25	pravit	113	997,92
12	věc	170	1501,30	26	udělat	113	997,92
13	myslet	157	1386,50	27	podívat	104	918,44
14	celý	144	1271,69				

Tabulka 18: Frekvenční seznam autosémantik (nad *h-bodem*) v (sub)korpusu Karla Čapka *Povídky z jedné a druhé kapsy* (1929).

¹⁹⁰ Na tomto srovnání by bylo zajímavé sledovat například odlišné pojetí kognitivních vlastností a strategií fiktivních postav u obou autorů. Čapek explicitněji tematizuje valenci „subjekt – kognitivní vlastnosti“ (viz prvních šest pozic a jejich frekvence).



Graf 3: Grafický průběh relativních frekvencí autosémantik (nad *h-bodem*) v (sub)korpusu *Povídky z jedné a druhé kapsy* (1929).

Z uvedených hodnot a srovnání je patrné, že tzv. tematická slova jsou jednoznačně příznaková nejen z hlediska své frekvence, která je oproti jejich rozložení v jiných korpusích ztlačeně vyšší,¹⁹¹ ale i z hlediska výběru, resp. typu slova. Z toho lze usoudit, že v dostatečně velkém a reprezentativním korpusu jsou tematická slova, která přísluší ke svému (konkrétnímu) korpusu, distribuována relativně analogicky a mimo korpus, ve kterém dosahují příznakových hodnot, nabývají neutrálního rozložení.¹⁹² V obou případech potom platí, že výskyt těchto slov (v našem případě lemmat¹⁹³) je odrazem jejich strukturní situace v příslušném (sub)korpusu. Jinými slovy, jejich frekvenční distribuce i kvalitativní výběr v námi sledovaných (sub)korpusech Jana Čepa je projevem strukturní intence uměleckého textu, resp. řady strukturních činitelů, které se podílejí na budování a formulování umělecké sémantiky fikčního narativu. Zjištění, která jsme doposud učinili, umožňují nejen smysluplně srovnávat jednotlivé oblasti výskytu tzv. tematických slov, ale rovněž detailněji interpretovat jejich frekvenční (kvantitativní) i kvalitativní aspekty. Frekvence (*Rf*) u sledovaných prvních tří (sub)korpusů Jana Čepa ukazují, že jejich průběh je odlišný, což se dotýká struktury tematických pásem. Rozdíly sledujeme na prvních třech pozicích.

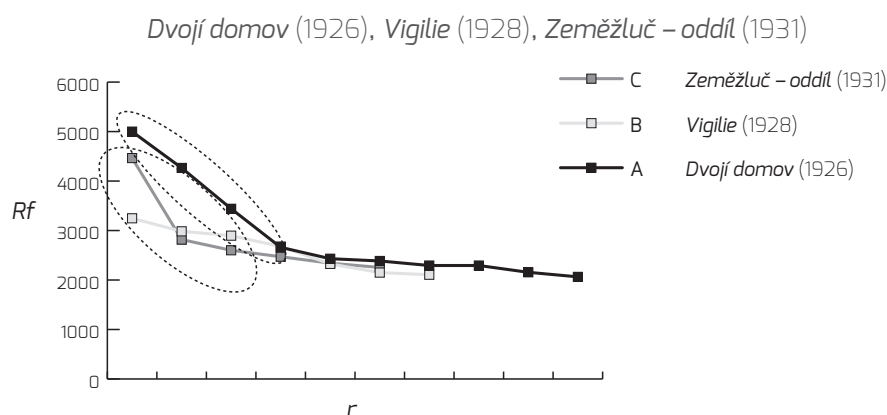
191 Žádná z naměřených hodnot *Rf* z Čepových (sub)korpusů neklesá pod hranici 2000 *Rf*, zatímco u jiných korpusů se pouze minimum lexémů dostává nad tuto hranici. Konkrétně se jedná o lemmata *jít*, *člověk* (v korpusu ČAPEK 1) a lemma *ruka* (v korpusu ČAPEK 2), která mají významnější frekvenční zatížení, což ovšem souvisí s tím, že uvedené pojmy mají samy o sobě v prozaické tvorbě Karla Čapka významnou funkci i sémantiku (viz vyšší pozice lemmat *jít* a *člověk* v Tab. 18).

192 Je pochopitelné, že křivky ČAPEK 1 a ČAPEK 2 vykazují oproti korpusové řadě SYN určité odlišnosti, které jsou způsobené tím, že tyto korpusy nemají obecně referenční povahu na rozdíl od SYN2010:BEL a SYN2010. Přesto i zde je patrná analogická tendence s průběhy křivek SYN2010:BEL a SYN2010.

193 Lze samozřejmě pracovat i se slovními tvary.

	<i>Dvojí domov</i> (1926) [Rf]	<i>Vigilie</i> (1928) [Rf]	<i>Zeměžluč – oddíl</i> (1931) [Rf]
1	4999,31	3247,32	4464,29
2	4265,47	2984,03	2817,27
3	3439,89	2896,26	2600,55
4	2660,18	2676,85	2470,53
5	2430,86	2325,79	2340,50
6	2384,99	2150,25	2253,81
7	2293,26	2106,37	
8	2293,26		
9	2155,67		
10	2063,94		

Tabuka 19: Srovnávací tabulka relativních frekvencí tematických slov prvních tří Čepových (sub)korpusů.



Graf 4: Srovnávací graf průběhu relativních frekvencí tematických slov (autosémantik nad *h-bodem*) v prvních třech Čepových (sub)korpusech.

Grafické znázornění nám umožňuje interpretovat vztahy uvnitř těchto pásem jako poměr kvality (výběr lexému), kvantity (frekvence) a struktury (uspořádání lexémů a frekvenční průběh). V (sub)korpusu *Dvojí domov* (1926) dominují první tři lemmata (*jít, oko, hlava*), zatímco v (sub)korpusu *Vigilie* (1928) je mezi tematickými slovy pozvolný průběh, čemuž odpovídá i vyšší hodnota *DTO*. V (sub)korpusu *Zeměžluč – oddíl* (1931) je v rámci analyzované oblasti frekvenčně příznakové již jen jediné lemma – *oko*. Sloučení prvních tří tabulek relativních frekvencí uvedených (sub)korpusů ukazuje rozdíly v průbězích křivek *Rf*. Výkyvy se týkají oblastí mezi pořadím 1–3, zatímco od pořadí 4 jsou průběhy relativně analogické. Rozdíly jsou zejména mezi průběhy frekvenčních hodnot u (sub)korpusu *Dvojí domov* (1926) na jedné straně a (sub)korpusy *Vigilie* (1928) a *Zeměžluč – oddíl* (1931) na straně druhé. Uvedené rozdíly ukazují tendenci k oslabování původně vysoce příznakových lemmat (*oko, jít, hlava*). Toto oslabování má vliv na strukturní proměnu vztahů mezi tematickými slovy, což se projevuje rovněž v modifikaci

výběru typu slov. Lemma *jít se* v (sub)korpusu *Zeměžluč – oddíl* (1931) odsouvá mimo oblast nad *h-bodem*. V (sub)korpusu *Vigilie* (1928) sice toto lemma je nad *h-bodem* a stejně jako v (sub)korpusu *Dvojí domov* (1926) také na první pozici, avšak zaznamenáváme zde významný frekvenční rozdíl, který mezi hodnotami *Rf* uvedeného lemmatu činí 53,95 %. Této situaci rovněž odpovídá příslušná hodnota *DTO*. Ačkoli se jedná spíše o modifikaci invariantního paradigmatu, uvedené aspekty signalizují jistou proměnu umělecké strategie, která se dotýká vlastní textové intence. Zde se již nabízí potenciál pro samostatnou literárněvědnou analýzu, která bude směřovat k otázkám, nakolik se daná změna projevuje v jednotlivých oblastech strukturní výstavby uměleckého narativu.¹⁹⁴

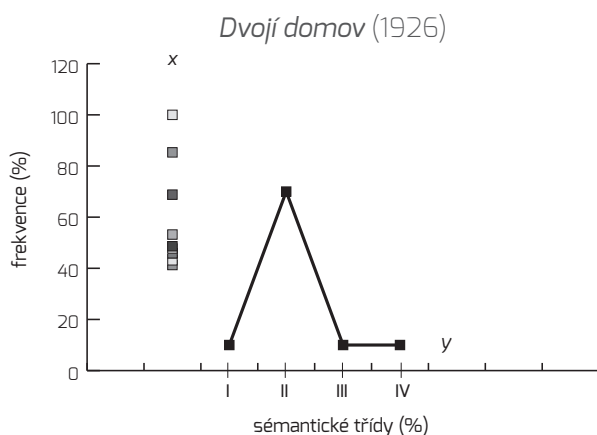
Jak jsme uvedli dříve, vedle frekvence lexika (kvantitativní kritérium) jsou důležité i další postupy, jež nabízí formální analýza. Propojení kvantitativního kritéria s kvalitativním umožňuje taxonomizaci lexika do sémantických tříd na základě sémému¹⁹⁵ jako distinktivního rysu. Ve sledovaných tematických oblastech Čepových (sub)korpusů tak můžeme specifikovat následující sémantické kategorie a jejich struktury.

¹⁹⁴ Konkrétně se na příslušném místě budeme věnovat vztahům kvantitativně-korpusových výsledků ke kategorii vypravěče, času a prostoru.

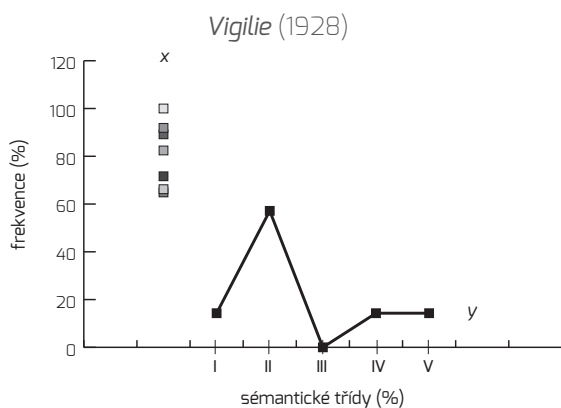
¹⁹⁵ Viz Karlík – Nekula – Pleskalová 2002, s. 386.

		I	II	III	IV	V	VI	VII	% (x)
Dvojí domov (1926)	1	jít							100,00
	2		oko						85,32
	3		hlava						68,81
	4		ruka						53,21
	5		matka						48,62
	6			okno					47,71
	7		otec						45,87
	8				den				45,87
	9		Rudolf						43,12
	10		Ludvík						41,28
% (y)		10,00	70,00	10,00	10,00				
Vigilie (1928)	1	jít							100,00
	2		Rozárka						91,89
	3		oko						89,19
	4					cesta			82,43
	5		tvář						71,62
	6				chvíle				66,22
	7		člověk						64,86
% (y)		14,29	57,14	0,00	14,29	14,29			
Zeměžluč – oddíl (1931)	1		oko						100,00
	2		tvář						63,11
	3		hlava						58,25
	4		ruka						55,34
	5		srdce						52,43
	6						cítit		50,49
% (y)		0,00	83,33	0,00	0,00	0,00	16,67		
Zeměžluč – celý soubor (1931)	1		oko						100,00
	2		tvář						63,28
	3		ruka						60,45
	4		hlava						60,45
	5					cesta			50,85
	6		člověk						48,59
	7						vědět		48,59
	8						cítit		48,02
	9				den				46,33
	10		matka						45,76
	11				chvíle				44,63
	12		srdce						42,37
	13		starý						41,24
	14		Rozárka						38,42
	15							svět	36,72
% (y)		0,00	60,00	0,00	13,33	6,67	13,33	6,67	

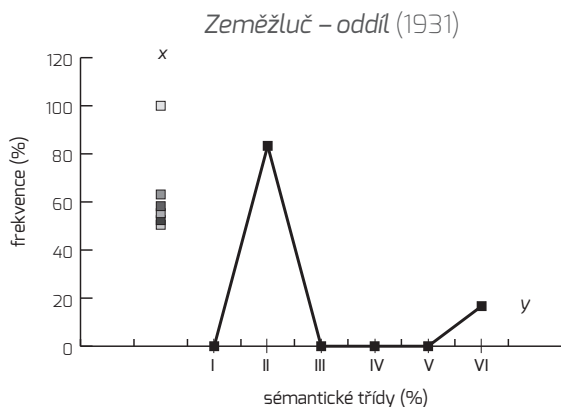
Tabulka 20: Taxonomie sémantických tříd lemmat nad *h-bodem* v jednotlivých (sub)korpusech. Procenta ve sloupci vyznačují proporční rozdíly mezi *Rf* lemmat, procenta v řádku vyznačují zastoupení lemmat v sémantické třídě. Vysvětlivky: **I** – pohyb, **II** – subjekt, **III** – předmět, **IV** – čas, **V** – prostor I (užší rámec prostorového horizontu – prostředí), **VI** – vlastnosti subjektu, **VII** – prostor II (širší rámec prostorového horizontu).



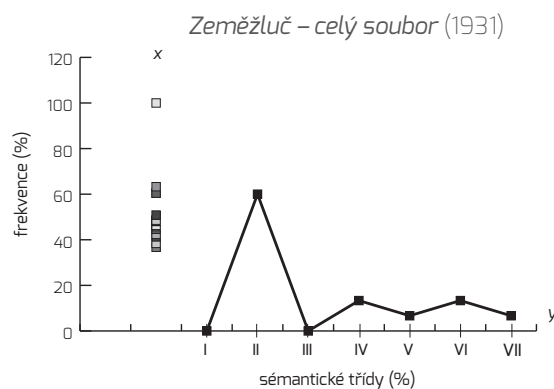
Graf 5a: Grafické znázornění vztahu mezi frekvenčními průběhy (x) a rozložením lemmat (autosémantik nad *h-bodem*) do sémantických tříd (y) v (sub)korpusu *Dvojí domov* (1926).



Graf 5b: Grafické znázornění vztahu mezi frekvenčními průběhy (x) a rozložením lemmat (autosémantik nad *h-bodem*) do sémantických tříd (y) v (sub)korpusu *Vigilie* (1928).



Graf 5c: Grafické znázornění vztahu mezi frekvenčními průběhy (x) a rozložením lemmat (autosémantik nad *h-bodem*) do sémantických tříd (y) v (sub)korpusu *Zeměžluč – oddíl* (1931).



Graf 5d: Grafické znázornění vztahu mezi frekvenčními průběhy (x) a rozložením lemmat (autosémantik nad *h-bodem*) do sémantických tříd (y) v (sub)korpusu *Zeměžluč – celý soubor (1931)*.

V uvedené tabulce (Tab. 20) a v příslušných grafech (Graf 5a–5d) můžeme sledovat distribuci sémantických tříd (v tabulce řádek s procenty – y) a frekvenční intenzitu jednotlivých lemmat vyjádřenou v procentech (v tabulce sloupec s procenty – x). V (sub)korpusu *Dvojí domov (1926)* dominuje sémantická třída +HUM (subjekt), která se úzce vztahuje ke kategorii pohybu.¹⁹⁶ Následuje kategorie předmětu (lemma *okno*) a temporality (lemma *den*). Ačkoli jako sémantické třídy mají tyto kategorie shodné procentuální zastoupení (10 %), dominantní význam kategorie pohybu určuje frekvenční zatížení lemmatu *jít*. V korpusu *Vigilie (1928)*, kde opět převládá třída +HUM (subjekt) a kategorie pohybu, je vedle temporality posílen motivický prvek prostoru (lemma *cesta*). To však automaticky nepředpokládá, že by prostorový aspekt ve *Dvojím domově (1926)* nebyl přítomný; lemma *cesta* se zde pohybuje těsně pod hranicí *h-bodu*, což dokazuje, že tento motiv je i zde významný.¹⁹⁷ Zdá se však, že ve *Vigiliích* došlo k jeho posílení, s čímž souvisí nejen otázka, nakolik se tento motiv proměňuje, ale také jak se vztahuje například ke kategoriím vypravěče, postavy, kompozice a konečně co uvedená situace znamená pro sémantiku prostoru. V oddíle *Zeměžluč (1931)* opět převládá kategorie +HUM (subjekt), ke které se pojí třída vlastností, jimiž je tato kategorie schopná disponovat (lemma *cítit*). V celku souboru *Zeměžluč (1931)* jsou potom kategorie temporality (lemmata *den, chvíle*), prostoru (lemma *cesta*) a vlastností subjektu (*vědět, cítit*) distribuovány relativně rovnoměrně, což odpovídá i frekvenčnímu rozložení těchto lemmat. Na uvedeném srovnání vidíme, jak jsou u Čepa strukturně propojeny frekvence s typem lemmatu a sémantickými

¹⁹⁶ Vypustili jsme vysoce frekventované lemma *stát*, které má v korpusu homonymní významy: vyskytuje se jako sloveso pohybu, spona (*stát se někým/něčím*) a ve významu finanční hodnoty. Ve skutečnosti by tedy mělo být procento zastoupení této sémantické třídy o něco vyšší.

¹⁹⁷ Hranici *h-bodu* je třeba chápat nikoli striktně, ale jako orientační mezník, který pochopitelně předpokládá, že i pod jeho hranicí bude určité významové pásmo lemmat, která se spolupodílejí na tematice textu.

třídami. Dominantní úlohu má zcela pochopitelně vždy sémantická třída +HUM (subjekt), ke které se s různou mírou intenzity a distribuce připojují sémantické kategorie pohybu, času, prostoru a vlastností, jimiž subjekt (postava) disponuje. Ve výsledném souboru *Zeměžluč* (1931) jsou tyto kategorie – s výjimkou stále dominující skupiny +HUM (subjekt) – distribuovány relativně rovnoměrně, což vede k úvaze o proměně strukturní strategie narativu a jeho částí. Rovněž je zřejmé, že jednotlivé kategorie mají schopnost se prostupovat, resp. hierarchizovat. Kategorie času a prostoru ze sémantického hlediska vytváří ve fikčním narativu vyšší strukturní vrstvu, která se přirozeně funkčně váže na kategorii +HUM (subjekt) a s ní spojené atributy (pohyb, vlastnosti, předmět). Navíc je každá z kategorií tvořena hyponymně-hyperonymními paradigmaty s tím, že frekvence ukazuje, že v rámci paradigmatu jsou symptomatická hyponyma. Ve výsledku jsme nyní schopni modelovat vztahové relace založené na sémantických bázích (kategoriích) s ohledem na jejich prototypickou funkci v literární fikci spolu s ústředním strukturním hyponymně-hyperonymním principem. Jednání postav se ve fikčních světech primárně odvíjí v časoprostorových rámcích. Dosavadní kvantitativní analýza Čepových (sub)korpusů signalizuje, že právě tato problematika zde bude ústředním tématem, které charakterizují centrální hyponymně-hyperonymní struktury.

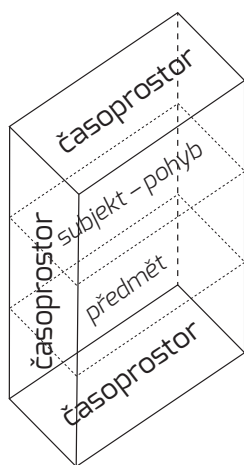


Schéma 15: Strukturace prototypických sémantických kategorií v prvních třech Čepových (sub)korpusech.

Na pozadí provedených analýz lze pozorovat, že jednou ze strukturních zákonitostí v prvních třech Čepových povídkových souborech je synekdochičnost,¹⁹⁸ avšak s proměnlivou mírou umělecké evokace. Synekdochičnost na úrovni kvantitativně-korpusové

¹⁹⁸ Význam principu synekdochy v Čepově tvorbě aktuálně zdůraznil Tomáš Kubíček v monografii *Dvojí domov Jana Čepa* (2014, s. 31).

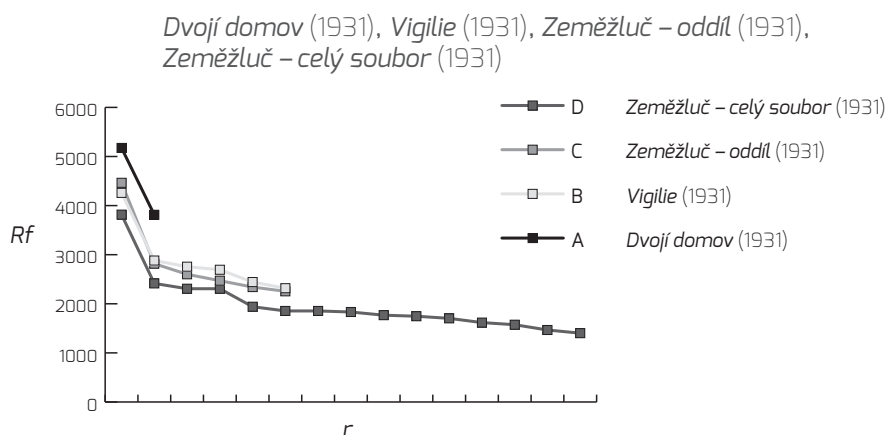
analýzy manifestují hyponymně-hyperonymní vztahy mezi lemmaty. Co to však znamená, uvažujeme-li tyto vazby mezi sémantickými kategoriemi? Především půjde o výslednou (sémantickou) projekci textově-narativní intence, která určuje specifickou podobu fikčního světa, jehož centrálním úběžníkem se princip synekdochy stává.

Synekdochičnost má ovšem u Čepa ještě jeden rozměr. Projevuje se v autorské intenci, jejíž realizací je radikální modifikace prvních dvou povídkových souborů a jejich vřazení do knihy *Zeměžluč* (1931).¹⁹⁹ Pokusme se na závěr tento aspekt ozřejmit rovněž pomocí kvantitativní analýzy tematických slov. Konkrétně srovnáme průběhy tematických slov u druhého vydání *Dvojího domova*, *Vigilií* a oddílu *Zeměžluč* a následně výsledky porovnáme s celým tímto souborem. Srovnání frekvenčních průběhů tematických slov u prvních vydání *Dvojího domova*, *Vigilií* a oddílu *Zeměžluč* ukázala, jak jsou rozkolísány zejména první pozice. Následující tabulka (viz Tab. 21) dokládá, jak se tyto vztahy mezi jednotlivými frekvenčními průběhy proměnily v případě druhých verzí *Dvojího domova* a *Vigilií* (viz Graf 6a, Graf 6b).

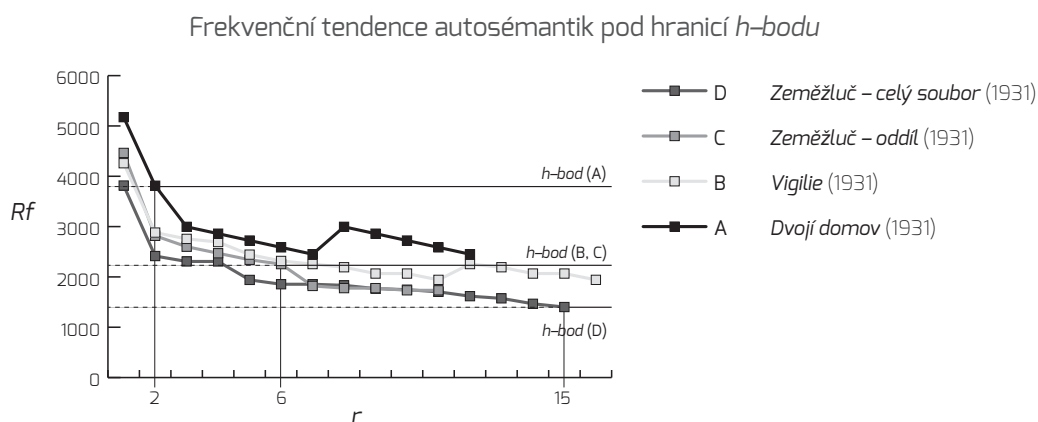
r	<i>Dvojí domov</i> (1931) [Rf]	<i>Vigilie</i> (1931) [Rf]	<i>Zeměžluč – oddíl</i> (1931) [Rf]	<i>Zeměžluč – celý soubor</i> (1931) [Rf]
1	5174,29	4259,05	4464,29	3816,14
2	3812,64	2881,12	2817,27	2414,73
3	2995,64	2755,86	2600,55	2306,93
4	2859,48	2693,22	2470,53	2306,93
5	2723,31	2442,69	2340,50	1940,41
6	2587,15	2317,42	2253,81	1854,17
7	2450,98	2254,79	1820,39	1854,17
8	2995,64	2192,16	1777,05	1832,61
9	2859,48	2066,89	1777,05	1767,93
10	2723,31	2066,89	1733,70	1746,37
11	2587,15	1941,63	1733,70	1703,25
12	2450,98	2254,79		1617,01
13		2192,16		1573,89
14		2066,89		1466,09
15		2066,89		1401,41
16		1941,63		

Tabulka 21: Srovnávací tabulka vybraných Čepových (sub)korpusů. V tabulce jsou uvedeny Rf tematických slov (v mřížce) a vybraných autosémantik, která se nacházejí pod *h-bodem* (mimo mřížku).

199 Viz Tab. 2.



Graf 6a: Grafické průběhy Rf autosémantik nad hranicí h -bodu ve vybraných (sub)korpusech.



Graf 6b: Grafické průběhy Rf autosémantik ve vybraných (sub)korpusech s vyznačením h -bodu.

Průběhy frekvencí v tematických oblastech mají u druhé verze *Dvojího domova* (1931) a *Vigilií* (1931) tendenci k analogičnosti, což se projevuje rovněž v poměru k oddílu *Zeměžluč*. Na druhém grafu (viz Graf 6b) je tato tendence zřetelnější díky zapojení autosémantik pod hranicí h -bodu. Z tohoto hlediska se vzájemné vztahy jednotlivých částí celého souboru *Zeměžluč* jeví mnohem více „harmonizovány“, než je tomu mezi prvním vydáním *Dvojího domova*, *Vigilií* a oddílem *Zeměžluč*.

3.3 Kolokace vybraných lemmat tematických oblastí

Základní frekvenční distribuce lemmat v prvních třech Čepových (sub)korpusech nám ukázala, že se většinou jedná o tematicky variantní řešení s tím, že to, co se proměňuje, není ani tak vlastní téma jako spíše intenzita určitých tematických aspektů. V (sub)korpusu *Dvojí domov* (1926) jednoznačně frekvenčně dominuje lemma *jít*, po kterém následují lemmata *oko* a *hlava*. Všechna (spolu s ostatními) vytvářejí užší sémantické korelační skupiny, zejména pak ale základní osnovu vyznačenou vztahem subjektu, jeho pohybu a perspektivy. Důležité je, jak je kategorie subjektu, kterou na úrovni korpusu reprezentují ostatní lemmata nad *h-bodem* (s výjimkou lemmatu *okno*), koncipována. Subjekt – postava/y příběhu jsou umělecky tematizovány skrze své základní přirozené atributy (*oko*, *hlava*, *ruka*), což platí pro všechny sledované (sub)korpusy. *De facto* se jedná o synekdochický princip. Rozdíly však nacházíme v kontextovém okolí lemmat jednotlivých (sub)korpusů. Jako příklad uvedeme tři shodná lemmata, která se vyskytují u všech zde analyzovaných (sub)korpusů a která náleží do rozdílných sémantických skupin (*oko*, *jít*, *cesta*). Kontextové okolí budeme zjišťovat korpusovou analýzou pomocí tří asociačních hodnot: t-score, MI-score a log-likelihood.²⁰⁰

Lemma *oko*, jež náleží do sémantické třídy subjektu, tvoří v prvních třech (sub)korpusech následující kolokace.

200 Vzájemným srovnáváním těchto a dalších asociačních měr se zabýval Michal Křen, z jehož poznatků v případě řazení hodnot v tabulkách vycházíme: „MI-score vyjadřuje asociativní sílu kolokace, jde o poměr skutečné frekvence kolokace a frekvence očekávané v případě, že by byla slova v korpusu rozložena náhodně. Hodnota t-score naopak vyjadřuje míru jistoty, se kterou není nalezená kolokace dílem náhody, ale naopak pravidelným jevem; neříká však téměř nic o její síle.“ (Křen 2006, s. 224) Parametry funkce log-likelihood Křen charakterizuje následovně: „Dá se tedy říci, že log-likelihood se chová jako konzervativní vylepšení t-score, jehož extrémy má tendenci potlačovat a místo nich dávat prostor relevantnějším bigramům. [...] Celkově však platí, že log-likelihood je v zásadě míra s podobnými vlastnostmi jako t-score [...]. Preference bigramů s vysokou frekvencí však není u log-likelihood tak silná, rozdíly mezi ním a t-score vyznívají většinou ve prospěch log-likelihood.“ (Ibidem, s. 245)

<i>Dvojí domov (1926)</i>					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	vytřeštit	3	1,73	11,01	41,18
2	otrhánek	3	1,73	10,84	40,31
3	Ludvík	5	2,23	9,19	54,55
4	Rudolf	3	1,73	8,42	29,34
5	lůžko	3	1,72	7,47	25,29
6	blízko	3	1,72	7,02	23,41
7	bledý	3	1,72	6,76	22,33
8	slza	4	1,98	6,63	29,09
9	hrůza	3	1,71	6,46	21,09
10	hledět	6	2,42	6,35	41,41
11	obrátit	3	1,71	6,18	19,90
12	veliký	5	2,21	6,16	33,15
13	tma	3	1,70	5,85	18,58
14	bez	4	1,95	5,34	22,03
15	okno	4	1,94	5,12	20,82
16	hlava	5	2,16	4,92	24,70
17	matka	3	1,67	4,91	14,73
18	mít	11	3,19	4,71	52,00
19	do	17	3,96	4,68	80,66
20	vidět	4	1,92	4,58	17,90
21	chtít	3	1,66	4,54	13,24
22	jít	5	2,13	4,43	21,44
23	jeho	6	2,32	4,29	24,65
24	ona	7	2,48	4,04	26,51
25	on	19	4,08	3,96	72,55
26	oko	4	1,86	3,82	13,91
27	před	3	1,61	3,81	10,36
28	a	52	6,70	3,81	210,76
29	její	3	1,61	3,81	10,35
30	po	5	2,07	3,79	17,23
31	oni	4	1,83	3,59	12,72
32	!	4	1,83	3,59	12,72
33	z	7	2,42	3,56	22,19
34	který	8	2,59	3,55	25,36
35	k	6	2,24	3,55	18,86
36	v	16	3,66	3,54	51,84
37	s	11	3,03	3,51	34,55
38	jak	3	1,57	3,45	8,99
39	.	55	6,73	3,43	198,03
40	ještě	3	1,57	3,41	8,85
41	za	4	1,80	3,31	11,30
42	na	14	3,33	3,20	38,92
43	se	32	5,03	3,16	93,74
44	jako	5	1,99	3,16	13,26
45	svůj	3	1,52	3,06	7,52

46	ale	4	1,76	3,03	9,91
47	,	48	5,81	2,63	117,29
48	"	7	2,09	2,24	11,03
49	být	7	1,52	1,23	4,07
50	ten	3	0,94	1,13	1,46

Tabulka 22a: Kolokace lematu *oko* v (sub)korpusu *Dvojí domov* (1926), řazeno podle MI-score.

<i>Dvojí domov</i> (1926)					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	a	52	6,70	3,81	210,76
2	.	55	6,73	3,43	198,03
3	,	48	5,81	2,63	117,29
4	se	32	5,03	3,16	93,74
5	do	17	3,96	4,68	80,66
6	on	19	4,08	3,96	72,55
7	Ludvík	5	2,23	9,19	54,55
8	mít	11	3,19	4,71	52,00
9	v	16	3,66	3,54	51,84
10	hledět	6	2,42	6,35	41,41
11	vytřeštit	3	1,73	11,01	41,18
12	otrhánek	3	1,73	10,84	40,31
13	na	14	3,33	3,20	38,92
14	s	11	3,03	3,51	34,55
15	veliký	5	2,21	6,16	33,15
16	Rudolf	3	1,73	8,42	29,34
17	slza	4	1,98	6,63	29,09
18	ona	7	2,48	4,04	26,51
19	který	8	2,59	3,55	25,36
20	lůžko	3	1,72	7,47	25,29
21	hlava	5	2,16	4,92	24,70
22	jeho	6	2,32	4,29	24,65
23	blízko	3	1,72	7,02	23,41
24	bledý	3	1,72	6,76	22,33
25	z	7	2,42	3,56	22,19
26	bez	4	1,95	5,34	22,03
27	jit	5	2,13	4,43	21,44
28	hrůza	3	1,71	6,46	21,09
29	okno	4	1,94	5,12	20,82
30	obrátit	3	1,71	6,18	19,90
31	k	6	2,24	3,55	18,86
32	tma	3	1,70	5,85	18,58
33	vidět	4	1,92	4,58	17,90
34	po	5	2,07	3,79	17,23
35	matka	3	1,67	4,91	14,73
36	oko	4	1,86	3,82	13,91
37	jako	5	1,99	3,16	13,26

38	chtít	3	1,66	4,54	13,24
39	oni	4	1,83	3,59	12,72
40	!	4	1,83	3,59	12,72
41	za	4	1,80	3,31	11,30
42	"	7	2,09	2,24	11,03
43	před	3	1,61	3,81	10,36
44	její	3	1,61	3,81	10,35
45	ale	4	1,76	3,03	9,91
46	jak	3	1,57	3,45	8,99
47	ještě	3	1,57	3,41	8,85
48	svůj	3	1,52	3,06	7,52
49	být	7	1,52	1,23	4,07
50	ten	3	0,94	1,13	1,46

Tabulka 22b: Kolokace lemmatu *oko* v (sub)korpusu *Dvojí domov* (1926), řazeno podle log-likelihood.

<i>Vigilie</i> (1928)					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	log-likelihood
1	zpod	3	1,73	8,92	31,43
2	zavřít	4	1,99	8,07	37,18
3	slza	6	2,44	7,71	52,90
4	Rozárka	4	1,99	7,36	33,21
5	klobouk	3	1,72	7,27	24,45
6	obrátit	4	1,99	7,09	31,66
7	bílý	4	1,97	6,02	25,77
8	její	6	2,39	5,30	32,97
9	vidět	3	1,66	4,66	13,75
10	k	7	2,51	4,26	28,80
11	svůj	4	1,87	3,97	14,75
12	a	40	5,91	3,92	171,32
13	na	16	3,73	3,89	60,10
14	tak	3	1,61	3,87	10,61
15	z	6	2,28	3,83	21,21
16	co	3	1,60	3,75	10,15
17	ona	4	1,85	3,72	13,47
18	do	6	2,26	3,67	19,94
19	v	12	3,18	3,62	40,22
20	on	10	2,89	3,53	32,05
21	ale	4	1,83	3,52	12,44
22	se	26	4,60	3,36	84,18
23	s	7	2,39	3,35	20,51
24	mít	3	1,56	3,33	8,57
25	-	3	1,54	3,15	7,89
26	jako	3	1,50	2,92	7,04
27	,	39	5,36	2,82	108,99
28	.	24	4,16	2,73	57,62
29	"	5	1,77	2,25	7,93

30	být	10	2,49	2,24	16,26
31	ten	4	1,51	2,04	5,39

Tabulka 22c: Kolokace lemmatu *oko* v (sub)korpusu *Vigilie* (1928), řazeno podle MI-score.

<i>Vigilie</i> (1928)					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	a	40	5,91	3,92	171,32
2	,	39	5,36	2,82	108,99
3	se	26	4,60	3,36	84,18
4	na	16	3,73	3,89	60,10
5	.	24	4,16	2,73	57,62
6	slza	6	2,44	7,71	52,90
7	v	12	3,18	3,62	40,22
8	zavřít	4	1,99	8,07	37,18
9	Rozárka	4	1,99	7,36	33,21
10	její	6	2,39	5,30	32,97
11	on	10	2,89	3,53	32,05
12	obrátit	4	1,99	7,09	31,66
13	zpod	3	1,73	8,92	31,43
14	k	7	2,51	4,26	28,80
15	bílý	4	1,97	6,02	25,77
16	klobouk	3	1,72	7,27	24,45
17	z	6	2,28	3,83	21,21
18	s	7	2,39	3,35	20,51
19	do	6	2,26	3,67	19,94
20	být	10	2,49	2,24	16,26
21	svůj	4	1,87	3,97	14,75
22	vidět	3	1,66	4,66	13,75
23	ona	4	1,85	3,72	13,47
24	ale	4	1,83	3,52	12,44
25	tak	3	1,61	3,87	10,61
26	co	3	1,60	3,75	10,15
27	mít	3	1,56	3,33	8,57
28	"	5	1,77	2,25	7,93
29	-	3	1,54	3,15	7,89
30	jako	3	1,50	2,92	7,04
31	ten	4	1,51	2,04	5,39

Tabulka 22d: Kolokace lemmatu *oko* v (sub)korpusu *Vigilie* (1928), řazeno podle log-likelihood.

<i>Zeměžluč – oddíl</i> (1931)					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	nově	3	1,73	11,54	44,42
2	planoucí	3	1,73	11,28	42,83
3	vytřeštěný	4	2,00	10,47	51,49
4	albína	4	2,00	10,19	49,69
5	průsvitný	3	1,73	9,47	33,93

6	odvrátit	3	1,73	8,54	29,85
7	obrácený	3	1,73	8,47	29,55
8	poprvé	3	1,73	7,91	27,14
9	šátek	3	1,72	7,84	26,86
10	pohlédnout	3	1,72	7,40	25,01
11	slza	6	2,43	7,07	47,38
12	spánek	3	1,72	7,05	23,55
13	zavřít	3	1,72	7,02	23,39
14	bledý	3	1,71	6,61	21,71
15	úzkost	3	1,71	6,48	21,15
16	hrůza	3	1,71	6,32	20,47
17	obrátit	3	1,71	6,03	19,29
18	ústa	4	1,97	5,98	25,50
19	lidský	3	1,70	5,97	19,03
20	vlas	3	1,70	5,92	18,84
21	ret	3	1,70	5,83	18,49
22	hledět	4	1,96	5,62	23,49
23	její	10	3,09	5,40	56,32
24	dva	5	2,18	5,36	27,69
25	jejich	4	1,95	5,32	21,86
26	nový	3	1,69	5,27	16,16
27	stín	3	1,69	5,26	16,15
28	dítě	3	1,68	5,10	15,50
29	žena	3	1,68	5,07	15,37
30	ona	14	3,62	4,89	69,76
31	tvář	5	2,16	4,84	24,16
32	najednou	4	1,92	4,70	18,51
33	země	3	1,66	4,67	13,76
34	ruka	5	2,14	4,55	22,18
35	hlava	4	1,91	4,45	17,19
36	do	16	3,82	4,44	70,43
37	i	4	1,91	4,44	17,13
38	vidět	4	1,91	4,43	17,10
39	s	22	4,46	4,36	95,77
40	oko	6	2,32	4,26	24,38
41	jenž	4	1,88	4,09	15,31
42	před	4	1,88	4,08	15,23
43	v	22	4,37	3,86	81,12
44	a	52	6,64	3,66	196,41
45	po	5	2,06	3,64	16,26
46	když	5	2,06	3,62	16,14
47	u	3	1,59	3,61	9,59
48	od	3	1,59	3,58	9,49
49	-	6	2,23	3,51	18,55
50	!	4	1,82	3,44	11,95
51	svůj	4	1,80	3,33	11,38
52	on	13	3,23	3,27	37,00

53	z	6	2,18	3,19	16,14
54	jeho	3	1,54	3,14	7,81
55	se	34	5,15	3,10	96,48
56	,	74	7,60	3,10	249,98
57	na	14	3,29	3,05	36,17
58	oni	3	1,52	3,03	7,40
59	jako	5	1,96	3,02	12,33
60	ale	3	1,42	2,47	5,40
61	který	4	1,62	2,40	6,95
62	k	3	1,40	2,40	5,17
63	ten	8	2,29	2,39	14,02
64	"	7	2,03	2,09	9,87
65	.	21	3,35	1,90	27,04
66	být	10	2,11	1,59	9,18

Tabulka 22e: Kolokace lemmatu *oko* v (sub)korpusu *Zeměžluč – oddíl* (1931), řazeno podle MI-score.

<i>Zeměžluč – oddíl (1931)</i>					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	<i>log-likelihood</i>
1	,	74	7,60	3,10	249,98
2	a	52	6,64	3,66	196,41
3	se	34	5,15	3,10	96,48
4	s	22	4,46	4,36	95,77
5	v	22	4,37	3,86	81,12
6	do	16	3,82	4,44	70,43
7	ona	14	3,62	4,89	69,76
8	její	10	3,09	5,40	56,32
9	vytřeštěný	4	2,00	10,47	51,49
10	albína	4	2,00	10,19	49,69
11	slza	6	2,43	7,07	47,38
12	nově	3	1,73	11,54	44,42
13	planoucí	3	1,73	11,28	42,83
14	on	13	3,23	3,27	37,00
15	na	14	3,29	3,05	36,17
16	průsvitný	3	1,73	9,47	33,93
17	odvrátit	3	1,73	8,54	29,85
18	obrácený	3	1,73	8,47	29,55
19	dva	5	2,18	5,36	27,69
20	poprvé	3	1,73	7,91	27,14
21	.	21	3,35	1,90	27,04
22	šátek	3	1,72	7,84	26,86
23	ústa	4	1,97	5,98	25,50
24	pohlédnout	3	1,72	7,40	25,01
25	oko	6	2,32	4,26	24,38
26	tvář	5	2,16	4,84	24,16
27	spánek	3	1,72	7,05	23,55
28	hledět	4	1,96	5,62	23,49

29	zavřít	3	1,72	7,02	23,39
30	ruka	5	2,14	4,55	22,18
31	jejich	4	1,95	5,32	21,86
32	bledý	3	1,71	6,61	21,71
33	úzkost	3	1,71	6,48	21,15
34	hrůza	3	1,71	6,32	20,47
35	obrátit	3	1,71	6,03	19,29
36	lidský	3	1,70	5,97	19,03
37	vlas	3	1,70	5,92	18,84
38	-	6	2,23	3,51	18,55
39	najednou	4	1,92	4,70	18,51
40	ret	3	1,70	5,83	18,49
41	hlava	4	1,91	4,45	17,19
42	i	4	1,91	4,44	17,13
43	vidět	4	1,91	4,43	17,10
44	po	5	2,06	3,64	16,26
45	nový	3	1,69	5,27	16,16
46	stín	3	1,69	5,26	16,15
47	když	5	2,06	3,62	16,14
48	z	6	2,18	3,19	16,14
49	dítě	3	1,68	5,10	15,50
50	žena	3	1,68	5,07	15,37
51	jenž	4	1,88	4,09	15,31
52	před	4	1,88	4,08	15,23
53	ten	8	2,29	2,39	14,02
54	země	3	1,66	4,67	13,76
55	jako	5	1,96	3,02	12,33
56	!	4	1,82	3,44	11,95
57	svůj	4	1,80	3,33	11,38
58	"	7	2,03	2,09	9,87
59	u	3	1,59	3,61	9,59
60	od	3	1,59	3,58	9,49
61	být	10	2,11	1,59	9,18
62	jeho	3	1,54	3,14	7,81
63	oni	3	1,52	3,03	7,40
64	který	4	1,62	2,40	6,95
65	ale	3	1,42	2,47	5,40
66	k	3	1,40	2,40	5,17

Tabulka 22f: Kolokace lemmatu *oko* v (sub)korpusu *Zeměžluč – oddíl* (1931), řazeno podle log-likelihood.

V (sub)korpusu *Dvojit domov* (1926) je frekvenčně příznakově zatížené lemma *oko* kontextualizováno lexikálním okolím, které implikuje výraznou expresivitu (*vytřeštit, bledý, hrůza, hledět, veliký, slza*). Ve *Vigiliích* (1928) se naopak zdá, že je tato expresivita poněkud oslabena na úkor možné deskriptivity (*zpod, zavřít, klobouk, bílý, vidět*), i když anulována není (*slza*). Naopak v oddíle *Zeměžluč* (1931) je expresivita opět výrazně přítomná

(*nově, planoucí, vytřeštěný, průsvitný, odvrátit, obrácený, slza*) s tím (určitým) rozdílem, že v tomto případě se do seznamu dostávají hodnotící a relační adjektiva (*planoucí, průsvitný, bledý, lidský*), která by mohla signalizovat například změnu narativní strategie.²⁰¹

Analogickou situaci pozorujeme také v kontextech lemmat *jít* a *cesta*, jež vytvářejí komplementární sémantickou dvojici určenou přirozenou vazbou mezi sémantickou kategorií pohybu a prostoru. Všimněme si, že zatímco lemma *jít* má v (sub)korpusu *Dvojí domov* (1926) významné frekvenční postavení, v (sub)korpusu *Vigilie* (1928), kde se sice rovněž umísťuje na první pozici autosémantik nad *h-bodem*, již nevytváří tak výrazný přechod mezi ostatními lemmaty. Křivka průběhu relativních frekvencí je zde vyrovnanější než v (sub)korpusu *Dvojí domov* (1926). Konečně v (sub)korpusu oddílu *Zeměžluč* (1931) se lemma *jít* již nedostává nad hranici *h-bodu*.

<i>Dvojí domov</i> (1926)					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	<i>log-likelihood</i>
1	nutný	4	2,00	8,95	42,21
2	rovnou	3	1,73	8,69	30,52
3	Rudolf	4	2,00	8,61	40,25
4	švec	3	1,73	8,57	29,99
5	kam	6	2,44	7,78	53,38
6	vstávat	3	1,72	7,69	26,24
7	kráva	3	1,72	7,42	25,10
8	Jeník	3	1,72	7,18	24,08
9	domů	5	2,22	7,12	39,77
10	blízko	3	1,72	6,79	22,45
11	pořád	3	1,71	6,61	21,70
12	Jakub	4	1,98	6,41	27,84
13	však	7	2,61	6,02	45,17
14	náves	3	1,71	5,99	19,13
15	daleko	4	1,97	5,91	25,12
16	otevřít	3	1,70	5,91	18,80
17	les	4	1,97	5,84	24,70
18	nikdo	4	1,96	5,82	24,58
19	muset	7	2,60	5,72	42,33
20	kolem	3	1,70	5,60	17,52
21	jít	12	3,39	5,47	68,87
22	přece	3	1,69	5,41	16,74
23	otec	4	1,95	5,40	22,31

201 Toto a jemu podobná tvrzení směřující k uměleckým postupům a strategiím je v tuto chvíli nezbytné chápat jako prozatímní hypotézy stanovené na základě výsledků kvantitativně-korpusové analýzy. Postupně je budeme rozšiřovat o další typy formálních analýz. V závěrečné části se je pokusíme funkčně propojit s konkrétní literárněvědnou interpretací zaměřenou na kategorii vypravěče, času a prostoru. Jak bylo nejednou zdůrazněno, samotná kvantitativně-korpusová analýza nemůže suplovat literárněvědný výklad. Proto na tuto skutečnost upozorňujeme s tím, že interpretace výstupních korpusových informací a hodnot probíhá s ohledem na jejich potencialitu v struktuře uměleckého narativu.

24	svět	4	1,95	5,30	21,77
25	přes	4	1,95	5,20	21,21
26	rád	3	1,68	5,11	15,53
27	večer	3	1,68	5,07	15,37
28	my	4	1,94	5,05	20,43
29	dítě	3	1,68	5,02	15,17
30	pak	5	2,16	4,96	24,93
31	cesta	5	2,16	4,85	24,23
32	aby	5	2,16	4,83	24,08
33	po	12	3,34	4,82	58,41
34	pro	3	1,67	4,73	13,97
35	za	11	3,17	4,54	49,27
36	k	14	3,58	4,54	63,09
37	okno	3	1,65	4,48	12,98
38	den	4	1,90	4,38	16,85
39	!	8	2,69	4,36	33,70
40	:	6	2,33	4,35	25,06
41	?	6	2,33	4,34	25,03
42	moci	5	2,13	4,33	20,72
43	chtít	3	1,65	4,31	12,32
44	vědět	4	1,89	4,19	15,82
45	s	20	4,22	4,14	80,61
46	jeden	3	1,63	4,05	11,29
47	ale	9	2,81	3,97	33,37
48	svůj	6	2,28	3,83	21,01
49	do	11	3,08	3,82	38,81
50	"	24	4,54	3,79	86,51

Tabulka 23a: Kolokace lemmatu *jít* v (sub)korpusu *Dvojí domov* (1926), řazeno podle MI-score.

<i>Dvojí domov</i> (1926)					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	.	65	7,32	3,44	235,59
2	,	73	7,48	3,00	230,39
3	a	45	6,06	3,37	147,36
4	"	24	4,54	3,79	86,51
5	s	20	4,22	4,14	80,61
6	být	28	4,63	3,00	73,58
7	se	29	4,61	2,79	69,10
8	jít	12	3,39	5,47	68,87
9	k	14	3,58	4,54	63,09
10	po	12	3,34	4,82	58,41
11	kam	6	2,44	7,78	53,38
12	za	11	3,17	4,54	49,27
13	však	7	2,61	6,02	45,17
14	muset	7	2,60	5,72	42,33
15	nutný	4	2,00	8,95	42,21

16	Rudolf	4	2,00	8,61	40,25
17	domů	5	2,22	7,12	39,77
18	na	15	3,41	3,07	39,11
19	do	11	3,08	3,82	38,81
20	!	8	2,69	4,36	33,70
21	ale	9	2,81	3,97	33,37
22	on	12	3,05	3,07	31,03
23	rovnou	3	1,73	8,69	30,52
24	švec	3	1,73	8,57	29,99
25	Jakub	4	1,98	6,41	27,84
26	vstávat	3	1,72	7,69	26,24
27	daleko	4	1,97	5,91	25,12
28	kráva	3	1,72	7,42	25,10
29	:	6	2,33	4,35	25,06
30	?	6	2,33	4,34	25,03
31	pak	5	2,16	4,96	24,93
32	les	4	1,97	5,84	24,70
33	nikdo	4	1,96	5,82	24,58
34	cesta	5	2,16	4,85	24,23
35	aby	5	2,16	4,83	24,08
36	Jeník	3	1,72	7,18	24,08
37	blízko	3	1,72	6,79	22,45
38	otec	4	1,95	5,40	22,31
39	svět	4	1,95	5,30	21,77
40	pořád	3	1,71	6,61	21,70
41	přes	4	1,95	5,20	21,21
42	svůj	6	2,28	3,83	21,01
43	jako	7	2,40	3,42	20,88
44	moci	5	2,13	4,33	20,72
45	my	4	1,94	5,05	20,43
46	náves	3	1,71	5,99	19,13
47	otevřít	3	1,70	5,91	18,80
48	kolem	3	1,70	5,60	17,52
49	den	4	1,90	4,38	16,85
50	přece	3	1,69	5,41	16,74

Tabulka 23b: Kolokace lemmatu *jít* v (sub)korpusu *Dvojí domov* (1926), řazeno podle log-likelihood.

<i>Vigilie</i> (1928)					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	log-likelihood
1	Francek	3	1,73	11,02	40,95
2	mimo	3	1,72	7,82	26,75
3	zároveň	3	1,72	7,14	23,92
4	vstát	3	1,72	7,01	23,37
5	rameno	3	1,71	6,62	21,75
6	jednou	4	1,97	6,18	26,66
7	jen	3	1,71	6,17	19,91

8	ty	4	1,96	5,76	24,34
9	dveře	3	1,70	5,58	17,47
10	cesta	5	2,18	5,41	28,11
11	!	10	3,08	5,24	54,56
12	za	12	3,37	5,22	65,56
13	pak	4	1,95	5,19	21,25
14	jít	6	2,37	5,02	30,65
15	po	8	2,73	4,80	38,63
16	?	5	2,15	4,64	22,88
17	když	7	2,54	4,59	31,73
18	i	3	1,66	4,50	13,10
19	ona	7	2,52	4,37	29,68
20	jak	4	1,89	4,20	15,91
21	k	7	2,49	4,10	27,20
22	do	9	2,82	4,09	35,09
23	už	6	2,30	4,03	22,70
24	ale	6	2,29	3,94	22,03
25	:	3	1,62	3,90	10,75
26	a	39	5,77	3,72	151,74
27	tak	3	1,60	3,70	9,96
28	mít	4	1,83	3,58	12,72
29	já	4	1,83	3,57	12,66
30	-	4	1,81	3,40	11,81
31	jako	4	1,78	3,17	10,65
32	,	55	6,58	3,15	192,03
33	z	4	1,76	3,08	10,21
34	na	10	2,78	3,04	25,72
35	který	4	1,73	2,88	9,23
36	v	8	2,44	2,87	18,73
37	s	5	1,89	2,70	10,49
38	.	26	4,30	2,68	60,50
39	být	15	3,26	2,66	32,36
40	se	15	3,14	2,40	27,82
41	že	3	1,39	2,35	5,02
42	"	6	1,97	2,35	10,20
43	ten	5	1,75	2,19	7,60
44	on	3	1,17	1,63	2,77

Tabulka 23c: Kolokace lemmatu *jít* v (sub)korpusu *Vigilie* (1928), řazeno podle MI-score.

<i>Vigilie</i> (1928)					
<i>r</i>	lemma	<i>Af</i>	<i>t</i> -score	MI	log-likelihood
1	,	55	6,58	3,15	192,03
2	a	39	5,77	3,72	151,74
3	za	12	3,37	5,22	65,56
4	.	26	4,30	2,68	60,50
5	!	10	3,08	5,24	54,56

6	Francek	3	1,73	11,02	40,95
7	po	8	2,73	4,80	38,63
8	do	9	2,82	4,09	35,09
9	být	15	3,26	2,66	32,36
10	když	7	2,54	4,59	31,73
11	jít	6	2,37	5,02	30,65
12	ona	7	2,52	4,37	29,68
13	cesta	5	2,18	5,41	28,11
14	se	15	3,14	2,40	27,82
15	k	7	2,49	4,10	27,20
16	mimo	3	1,72	7,82	26,75
17	jednou	4	1,97	6,18	26,66
18	na	10	2,78	3,04	25,72
19	ty	4	1,96	5,76	24,34
20	zároveň	3	1,72	7,14	23,92
21	vstát	3	1,72	7,01	23,37
22	?	5	2,15	4,64	22,88
23	už	6	2,30	4,03	22,70
24	ale	6	2,29	3,94	22,03
25	rameno	3	1,71	6,62	21,75
26	pak	4	1,95	5,19	21,25
27	jen	3	1,71	6,17	19,91
28	v	8	2,44	2,87	18,73
29	dveře	3	1,70	5,58	17,47
30	jak	4	1,89	4,20	15,91
31	i	3	1,66	4,50	13,10
32	mít	4	1,83	3,58	12,72
33	já	4	1,83	3,57	12,66
34	-	4	1,81	3,40	11,81
35	:	3	1,62	3,90	10,75
36	jako	4	1,78	3,17	10,65
37	s	5	1,89	2,70	10,49
38	z	4	1,76	3,08	10,21
39	"	6	1,97	2,35	10,20
40	tak	3	1,60	3,70	9,96
41	který	4	1,73	2,88	9,23
42	ten	5	1,75	2,19	7,60
43	že	3	1,39	2,35	5,02
44	on	3	1,17	1,63	2,77

Tabulka 23d: Kolokace lemmatu *jít* v (sub)korpusu *Vigilie* (1928), řazeno podle log-likelihood.

Zeměžluč – oddíl (1931)					
r	lemma	Af	t-score	MI	log-likelihood
1	tento	5	2,21	6,65	37,17
2	matka	3	1,72	6,64	22,03
3	cesta	4	1,98	6,49	28,70

4	po	5	2,19	5,52	29,40
5	!	4	1,95	5,32	22,28
6	s	8	2,73	4,78	39,96
7	jako	4	1,92	4,57	18,25
8	už	3	1,65	4,43	13,01
9	-	3	1,65	4,39	12,85
10	který	4	1,90	4,28	16,71
11	ten	6	2,28	3,86	22,14
12	on	5	2,07	3,77	17,68
13	.	24	4,42	3,36	96,85
14	být	8	2,51	3,15	22,79
15	a	9	2,63	3,01	24,35
16	.	12	3,02	2,97	33,26
17	"	3	1,47	2,75	6,57
18	na	3	1,47	2,71	6,42
19	se	7	2,24	2,70	15,80

Tabulka 23e: Kolokace lemmatu *jít* v (sub)korpusu *Zeměžluč – oddíl (1931)*, řazeno podle MI-score.

<i>Zeměžluč – oddíl (1931)</i>					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	<i>log-likelihood</i>
1	.	24	4,42	3,36	96,85
2	s	8	2,73	4,78	39,96
3	tento	5	2,21	6,65	37,17
4	.	12	3,02	2,97	33,26
5	po	5	2,19	5,52	29,40
6	cesta	4	1,98	6,49	28,70
7	a	9	2,63	3,01	24,35
8	být	8	2,51	3,15	22,79
9	!	4	1,95	5,32	22,28
10	ten	6	2,28	3,86	22,14
11	matka	3	1,72	6,64	22,03
12	jako	4	1,92	4,57	18,25
13	on	5	2,07	3,77	17,68
14	který	4	1,90	4,28	16,71
15	se	7	2,24	2,70	15,80
16	už	3	1,65	4,43	13,01
17	-	3	1,65	4,39	12,85
18	"	3	1,47	2,75	6,57
19	na	3	1,47	2,71	6,42

Tabulka 23f: Kolokace lemmatu *jít* v (sub)korpusu *Zeměžluč – oddíl (1931)*, řazeno podle log-likelihood.

Rozdíly, ke kterým dochází v kontextovém okolí lemmatu *jít* u jednotlivých (sub)korpusů, jsou zjevné. Pro jejich přehlednější schematizaci je uspořádáme do kategorií.

1. V (sub)korpusu *Dvojí domov* (1926) se realizují tyto základní typologické kombinace sémantických kategorií:

a) modalita

jít: nutný, rovnou, muset, pořád, poslušně, moci, přece, pořád (*Konečně však přece nutno jít... / ... nutno jít poslušně, kam nás vedou... / ... jak je možno pořád jít dál a dál...*)

b) třída aktantů

jít: Rudolf, švec, kráva, Jakub, otec (*Rudolf jde zamyšlen za branami...*)

c) prostorová orientace²⁰²

jít: rovnou, kam, domů, blízko, daleko, les, náves, cesta, dům (*Doma říkali, aby šel pořád rovnou... / Nevěděl ještě, kam půjde... / Pojď blíž, dítě, pojď blíž, přitiskni se ke mně...*)

2. V (sub)korpusu *Vigilie* (1928) se realizují tyto základní typologické kombinace sémantických kategorií:

a) třída aktantů

jít: Francek, ona („Nu, však už jdu,“ odpovídá Francek...)

b) prostorová orientace

jít: mimo, cesta, dveře (... *šel mimo švec Hejl... / ... šel po opuštěné lesní cestě... / ... šel ke dveřím...*)

c) způsob

jít: zároveň, rameno (... *jdeš zároveň s krasavcem Frýbem... / Kdokoliv šel mimo a byl sám... / ... jda s vidlemi na rameně...*)

d) následnost

jít: vstát (*Když vstal a chtěl jít nazpět... / ... vstav od stolu šel chystat pluh a brány...*)

e) čas

jít: jednou, pak, po (*Ale pak jsem jednou vstal a šel jsem do dědiny za kopcem...*)

3. V (sub)korpusu *Zeměžluč – oddíl* (1931) se realizují tyto základní typologické kombinace sémantických kategorií:

a) aktanty

jít: matka (*Tato cesta, po které šel Cyril s matkou...*)

²⁰² Slovesa *jít* je však užito i v jiném než výhradně prostorově orientačním významu: *Oči šly blíž a blíž...*

b) prostorová orientace

jít: cesta, po (... cesta jeho dětství, byla cestou... / ... vrací se a jde zase kupředu, jako by chtěla jít dnes zároveň všemi cestami... / ... naskytne-li se mi někdy jít touto cestou, udělám velikou okliku...)

Kontextové okolí lemmatu *jít* je v (sub)korpusu *Vigilie* (1928) z hlediska kombinací se sémantickými kategoriemi bohatší, naopak v (sub)korpusu *Zeměžluč – oddíl* (1931) dochází opět k redukci kombinačních typů. Podíváme-li se pozorněji na skladbu sémantických kategorií v jednotlivých (sub)korpusech v rámci výběru autosémantik nad *h-bodem*, potom zjistíme, že je ve *Vigiliích* (1928) významněji obohacena o kategorii prostoru reprezentovanou lemmatem *cesta*, které nejenže vytváří korelační opozici k lemmatu *jít*, ale dokonce se může zdát, že jej bude částečně suplovat, jak vyplývá z oslabenějších frekvenčních rozdílů mezi oběma lemmaty.

Vygenerujeme-li z korpusu kolokace lemmatu *cesta* (viz Tab. 24a–24f), vidíme, jaké významové a strukturní analogie vytváří k základním významovým okruhům, které jsme určili v případě lemmat *jít* (pohyb, subjekt) a *oko* (subjekt).

<i>Dvojí domov (1926)</i>					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	<i>log-likelihood</i>
1	vrátit	3	1,72	7,06	23,65
2	pole	5	2,22	6,89	38,48
3	uprostřed	3	1,72	6,67	22,05
4	dva	4	1,98	6,34	27,64
5	mezi	3	1,70	5,75	18,23
6	jít	5	2,19	5,58	29,48
7	od	4	1,95	5,29	21,94
8	po	6	2,38	5,20	32,43
9	k	8	2,75	5,11	42,67
10	za	4	1,91	4,46	17,44
11	!	3	1,65	4,32	12,47
12	když	3	1,64	4,18	11,92
13	na	11	3,11	4,00	43,16
14	z	4	1,87	3,90	14,50
15	s	6	2,27	3,78	21,09
16	který	4	1,85	3,70	13,47
17	a	21	4,22	3,65	78,76
18	se	14	3,31	3,12	40,02
19	být	11	2,91	3,03	29,34
20	,	28	4,63	3,00	87,95
21	.	18	3,70	2,97	49,90

Tabulka 24a: Kolokace lemmatu *cesta* v (sub)korpusu *Dvojí domov* (1926), řazeno podle *MI-score*.

<i>Dvojí domov (1926)</i>					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	,	28	4,63	3,00	87,95
2	a	21	4,22	3,65	78,76
3	.	18	3,70	2,97	49,90
4	na	11	3,11	4,00	43,16
5	k	8	2,75	5,11	42,67
6	se	14	3,31	3,12	40,02
7	pole	5	2,22	6,89	38,48
8	po	6	2,38	5,20	32,43
9	jít	5	2,19	5,58	29,48
10	být	11	2,91	3,03	29,34
11	dva	4	1,98	6,34	27,64
12	vrátit	3	1,72	7,06	23,65
13	uprostřed	3	1,72	6,67	22,05
14	od	4	1,95	5,29	21,94
15	s	6	2,27	3,78	21,09
16	mezi	3	1,70	5,75	18,23
17	za	4	1,91	4,46	17,44
18	z	4	1,87	3,90	14,50
19	který	4	1,85	3,70	13,47
20	!	3	1,65	4,32	12,47
21	když	3	1,64	4,18	11,92

Tabulka 24b: Kolokace lemmatu *cesta* v (sub)korpusu *Dvojí domov* (1926), řazeno podle log-likelihood.

<i>Vigilie (1928)</i>					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	kvítí	3	1,73	9,19	32,60
2	polní	5	2,23	8,78	51,60
3	zastavit	5	2,23	8,18	47,36
4	potok	3	1,72	6,91	22,98
5	vedle	3	1,71	6,31	20,49
6	jaký	3	1,71	6,19	20,00
7	cesta	6	2,41	5,95	38,35
8	po	12	3,40	5,66	73,13
9	nad	4	1,95	5,24	21,53
10	jít	5	2,17	5,04	25,65
11	jenž	4	1,93	4,85	19,42
12	!	5	2,14	4,52	22,15
13	za	5	2,12	4,24	20,30
14	-	5	2,10	4,00	18,75
15	k	5	2,09	3,89	18,02
16	do	6	2,27	3,78	20,86
17	a	32	5,23	3,72	124,10
18	na	13	3,33	3,70	45,27
19	se	29	4,95	3,63	107,21

20	jako	4	1,82	3,45	12,08
21	já	3	1,57	3,43	8,96
22	být	17	3,65	3,12	47,55
23	z	3	1,51	2,95	7,14
24	on	6	2,12	2,91	14,26
25	,	36	5,15	2,82	100,48
26	který	3	1,47	2,75	6,42
27	.	21	3,85	2,65	47,95
28	v	5	1,83	2,47	9,24
29	ten	4	1,55	2,15	5,89

Tabulka 24c: Kolokace lemmatu *cesta* v (sub)korpusu *Vigilie* (1928), řazeno podle MI-score.

<i>Vigilie</i> (1928)					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	a	32	5,23	3,72	124,10
2	se	29	4,95	3,63	107,21
3	,	36	5,15	2,82	100,48
4	po	12	3,40	5,66	73,13
5	polní	5	2,23	8,78	51,60
6	.	21	3,85	2,65	47,95
7	být	17	3,65	3,12	47,55
8	zastavit	5	2,23	8,18	47,36
9	na	13	3,33	3,70	45,27
10	cesta	6	2,41	5,95	38,35
11	kvítí	3	1,73	9,19	32,60
12	jit	5	2,17	5,04	25,65
13	potok	3	1,72	6,91	22,98
14	!	5	2,14	4,52	22,15
15	nad	4	1,95	5,24	21,53
16	do	6	2,27	3,78	20,86
17	vedle	3	1,71	6,31	20,49
18	za	5	2,12	4,24	20,30
19	jaký	3	1,71	6,19	20,00
20	jenž	4	1,93	4,85	19,42
21	-	5	2,10	4,00	18,75
22	k	5	2,09	3,89	18,02
23	on	6	2,12	2,91	14,26
24	jako	4	1,82	3,45	12,08
25	v	5	1,83	2,47	9,24
26	já	3	1,57	3,43	8,96
27	z	3	1,51	2,95	7,14
28	který	3	1,47	2,75	6,42
29	ten	4	1,55	2,15	5,89

Tabulka 24d: Kolokace lemmatu *cesta* v (sub)korpusu *Vigilie* (1928), řazeno podle log-likelihood.

<i>Zeměžluč – oddíl (1931)</i>					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	Cyril	3	1,73	10,09	36,47
2	polní	3	1,73	8,81	31,00
3	dětství	3	1,73	8,68	30,45
4	duše	3	1,72	7,14	24,02
5	vedle	3	1,72	7,07	23,73
6	dnes	3	1,72	7,07	23,73
7	poslední	3	1,72	6,70	22,19
8	všechen	6	2,42	6,41	42,54
9	cesta	4	1,97	6,13	26,57
10	teď	3	1,71	6,10	19,72
11	po	9	2,95	6,01	59,64
12	tento	4	1,97	5,97	25,69
13	jít	5	2,20	5,80	31,11
14	svůj	5	2,17	5,17	26,79
15	jeho	4	1,94	5,07	20,80
16	oni	3	1,66	4,54	13,39
17	na	11	3,14	4,22	46,95
18	který	4	1,87	3,92	14,68
19	,	27	4,62	3,17	95,10
20	s	3	1,52	3,00	7,43
21	být	9	2,61	2,96	23,14
22	se	10	2,73	2,86	24,70
23	a	10	2,71	2,80	23,99
24	ten	3	1,43	2,50	5,62
25	.	10	2,54	2,34	18,47

Tabulka 24e: Kolokace lemmatu *cesta* v (sub)korpusu *Zeměžluč – oddíl (1931)*, řazeno podle MI-score.

<i>Zeměžluč – oddíl (1931)</i>					
<i>r</i>	lemma	<i>Af</i>	t-score	MI	log-likelihood
1	,	27	4,62	3,17	95,10
2	po	9	2,95	6,01	59,64
3	na	11	3,14	4,22	46,95
4	všechen	6	2,42	6,41	42,54
5	Cyril	3	1,73	10,09	36,47
6	jít	5	2,20	5,80	31,11
7	polní	3	1,73	8,81	31,00
8	dětství	3	1,73	8,68	30,45
9	svůj	5	2,17	5,17	26,79
10	cesta	4	1,97	6,13	26,57
11	tento	4	1,97	5,97	25,69
12	se	10	2,73	2,86	24,70
13	duše	3	1,72	7,14	24,02
14	a	10	2,71	2,80	23,99
15	dnes	3	1,72	7,07	23,73

16	vedle	3	1,72	7,07	23,73
17	být	9	2,61	2,96	23,14
18	poslední	3	1,72	6,70	22,19
19	jeho	4	1,94	5,07	20,80
20	teď	3	1,71	6,10	19,72
21	.	10	2,54	2,34	18,47
22	kteřý	4	1,87	3,92	14,68
23	oni	3	1,66	4,54	13,39
24	s	3	1,52	3,00	7,43
25	ten	3	1,43	2,50	5,62

Tabulka 24f: Kolokace lemmatu *cesta* v (sub)korpusu *Zeměžluč – oddíl* (1931), řazeno podle log-likelihood.

Pro srovnání, jak se kontextové okolí liší ve zcela jiném autorském (sub)korpusu, uvádíme výpis kolokací lemmatu *jít* (viz Tab. 25a, 25b) v Čapkově (sub)korpusu *Povídky z jedné a druhé kapsy* (1929).

<i>Povídky z jedné a druhé kapsy (1929)</i>					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	log-likelihood
1	severozápadní	4	2,00	12,11	62,09
2	dlážděný	4	2,00	11,45	57,23
3	příručí	4	2,00	11,01	54,25
4	kontor	4	2,00	10,59	51,63
5	pozpátku	4	2,00	10,33	50,02
6	strážmistr	3	1,73	9,69	34,67
7	Janík	7	2,64	9,61	80,15
8	Rybka	3	1,73	9,54	34,02
9	udat	6	2,45	9,31	66,04
10	kudy	7	2,64	9,23	76,26
11	Hirsch	3	1,73	9,14	32,27
12	vyspat	3	1,73	9,07	31,98
13	Záruba	4	2,00	8,98	42,13
14	oč	7	2,64	8,70	70,94
15	šlépěj	4	2,00	8,63	40,11
16	panička	3	1,73	8,57	29,82
17	komisař	7	2,64	8,39	67,88
18	nikam	3	1,73	8,37	28,98
19	teda	8	2,82	8,34	77,08
20	skrz	7	2,64	8,24	66,37
21	tudy	4	1,99	8,19	37,61
22	jestli	4	1,99	7,82	35,53
23	šofér	3	1,72	7,77	26,45
24	někam	6	2,44	7,73	52,61
25	rovnou	6	2,44	7,72	52,55
26	lehnout	3	1,72	7,72	26,23
27	krám	4	1,99	7,64	34,53

28	nádraží	5	2,23	7,60	42,91
29	pošta	5	2,22	7,58	42,76
30	zeptat	6	2,44	7,47	50,38
31	hospoda	5	2,22	7,45	41,90
32	podívat	22	4,66	7,29	180,38
33	četník	4	1,99	7,26	32,44
34	nadávat	3	1,72	7,25	24,27
35	domů	10	3,14	7,14	79,51
36	vražda	3	1,72	7,11	23,71
37	zavolat	3	1,72	7,06	23,49
38	nahoru	5	2,22	7,04	39,02
39	sebrat	3	1,72	6,90	22,81
40	zpátky	3	1,72	6,88	22,75
41	chlap	4	1,98	6,88	30,32
42	pryč	4	1,98	6,75	29,59
43	policie	3	1,72	6,72	22,08
44	byt	3	1,71	6,57	21,43
45	Juraj	3	1,71	6,48	21,06
46	vždyť	13	3,57	6,46	91,22
47	kam	7	2,62	6,44	48,89
48	ženská	3	1,71	6,35	20,53
49	odtud	3	1,71	6,29	20,30
50	služba	3	1,71	6,29	20,28

Tabulka 25a: Kolokace lemmatu *jít* v (sub)korpusu *Povídky z jedné a druhé kapsy* (1929), řazeno podle MI-score.

<i>Povídky z jedné a druhé kapsy (1929)</i>					
<i>r</i>	lemma	<i>Af</i>	<i>t-score</i>	<i>MI</i>	<i>log-likelihood</i>
1	.	301	15,33	3,11	1026,59
2	.	214	13,24	3,40	738,20
3	ten	159	11,83	4,02	651,87
4	a	133	10,36	3,30	409,61
5	já	78	8,53	4,89	392,62
6	"	90	8,79	3,77	322,37
7	na	76	8,20	4,07	299,86
8	se	99	8,94	3,30	296,53
9	k	53	7,05	4,99	270,63
10	pan	47	6,68	5,28	257,84
11	tak	50	6,82	4,81	243,08
12	;	69	7,58	3,52	220,82
13	do	44	6,38	4,72	207,71
14	_být	53	6,74	3,74	182,94
15	podívat	22	4,66	7,29	180,38
16	vy	36	5,77	4,68	167,47
17	řící	30	5,29	4,88	147,04
18	za	32	5,43	4,64	146,83

19	on	40	5,86	3,75	137,47
20	?	32	5,32	4,06	122,24
21	o	30	5,17	4,13	117,39
22	ale	31	5,15	3,73	105,17
23	!	26	4,74	3,83	91,24
24	vždyť	13	3,57	6,46	91,22
25	-	30	4,94	3,35	87,05
26	už	22	4,43	4,16	86,52
27	Janík	7	2,64	9,61	80,15
28	domů	10	3,14	7,14	79,51
29	svůj	24	4,52	3,68	79,42
30	teda	8	2,82	8,34	77,08
31	kudy	7	2,64	9,23	76,26
32	prosit	13	3,53	5,55	75,11
33	vědět	18	4,03	4,34	74,74
34	s	25	4,53	3,42	74,40
35	že	30	4,77	2,96	72,49
36	po	18	4,01	4,20	71,45
37	oč	7	2,64	8,70	70,94
38	komisař	7	2,64	8,39	67,88
39	_by	24	4,40	3,29	67,57
40	v	33	4,82	2,63	67,03
41	skrz	7	2,64	8,24	66,37
42	udat	6	2,45	9,31	66,04
43	když	16	3,80	4,30	65,55
44	tady	12	3,38	5,28	64,89
45	pak	13	3,48	4,83	62,35
46	severozápadní	4	2,00	12,11	62,09
47	být	44	5,04	2,06	61,71
48	dlážděný	4	2,00	11,45	57,23
49	tam	13	3,44	4,48	56,34
50	moci	17	3,80	3,68	56,04

Tabulka 25b: Kolokace lemmatu *jít* v (sub)korpusu *Povídky z jedné a druhé kapsy* (1929), řazeno podle log-likelihood.

Lemma *cesta* vytváří ve *Dvojím domově* (1926) omezené kontexty (*vrátit, pole, uprostřed, dva, mezi, jít...*), které nejen implikují užší strukturní vazby ke kategorii subjektu, ale současně se zdá, že sémantický okruh, který v souboru *Dvojí domov* (1926) produkuje lemma *jít*, souvisí s hlediskem postavy. Jedná se totiž o soubor modálních determinací chůze, základních směrových relací, jež se nacházejí v bezprostředním dosahu subjektu (*jít: blízko, daleko, les, náves, cesta*). Z hlediska sémantických kategorií pestřejší kontextová pole lemmatu *cesta* v (sub)korpusu *Vigilie* (1928) oproti (sub)korpusu *Dvojí domov* (1926) zakládají domněnku, že motiv cesty bude mít ve *Vigiliích* modifikovaný význam.

Ve *Dvojím domově* (1926) se vytváří příznaková tematická struktura „subjekt–jít“, přičemž subjekt je realizován hyponymně-hyperonymními vztahy, které předpokládají umělecký princip synekdochy. Ve *Vigiliích* (1928) je naopak zdůrazněn motiv cesty ve spojení s kategorií subjektu, který je rovněž řešen podobnou strategií jako v předchozím případě, tudíž v profilu jeho umělecké reprezentace významně dominují hyponymně-hyperonymní vazby. Ovšem kategorie subjektu (na korpusové úrovni reprezentovaná lemmatem *oko*) nevykazuje tak významně expresivní kontextové okolí, jako je tomu ve *Dvojím domově* (1926) a následně v oddíle *Zeměžluč* (1931).

Shrneme-li předběžné závěry, je možné na základě kvantitativně-korpusové analýzy tematických oblastí sledovaných (sub)korpusů prohlásit, že se od prvního (sub)korpusu, kterým je *Dvojí domov* (1926), pozvolna proměňuje vazba „subjekt–jít“ (cesta/prostor) v závislosti na míře expresivity subjektu (postavy). V oddíle *Zeměžluč* se potom expresivně projektovaný profil kategorie subjektu ocitá ve strukturní valenci s motivem cesty. Ačkoli se dané lemma nevyskytuje vyloženě nad *h-bodem*, jeho frekvenční pozice je sama o sobě vysoká a této hranici se přibližuje. Lemma *cesta* má modifikované kontextové okolí (*cesta: duše, dětství, poslední, cesta, polní*) a opět, avšak modifikovaněji než ve *Vigiliích* (1928) vybízí k tomu, abychom chápali motiv cesty v metaforickém významu. Na pozadí těchto strukturních analýz prováděných pomocí kvantitativně-korpusové metody se objevuje hypotéza o synekdochickém vztahu mezi základními kategoriemi „subjekt–prostor“ s tím, že tato vazba, která tvoří základní (fundamentální) tematický rámec, bude v každém ze sledovaných Čepových povídkových souborů realizována jinou uměleckou technikou; přitom však půjde o společné tematické paradigma. Přirozeně tak s touto úvahou souvisí i otázka směřující například k narativním strategiím, formám či narativnímu hledisku.

3.4 Strukturace sémantických tříd substantiv a jejich vztahy, základní situace adjektiv a sloves

Z analýz tematických oblastí jednotlivých (sub)korpusů je zjevné, že zde převládají substantiva s výjimkou vysoce frekventovaného slovesa *jít*. V rámci kvantitativně-korpusové analýzy se nelze spokojit výhradně s izolovaným výskytem lexému či slovního druhu, ale je nezbytné se rovněž zaměřit na typ lexému, jeho sémantickou třídu a kontextové pole, jež je utvářeno tzv. kolokačním paradigmatickým příslušného lexému. Tyto atributy spolu s frekvenčním příznakem a rozložením lexémů pomohou specifikovat obecnější tendence textu, z nichž lze modelovat jejich strukturní vektor. Stále však musíme mít na paměti, že analýza je empiricky prováděna na jazykové rovině; teprve až indukci je

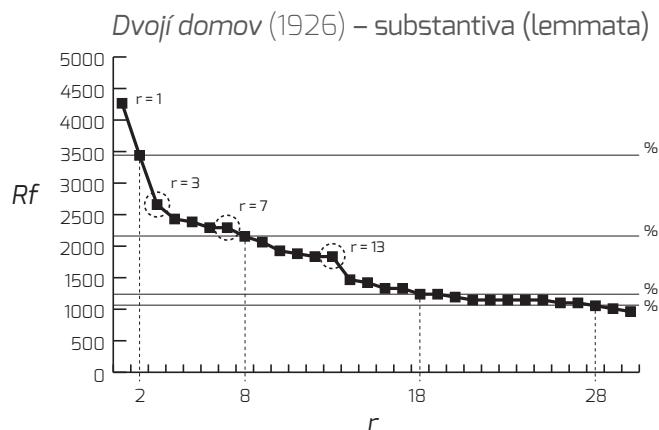
v rámci metodologicky koherentního postupu (strukturálního) extrapolována do vyšších strukturně-významových vrstev fikčního narativu (viz kap. 3.5). Proto je důležité, aby formální analýza (sub)korpusů dokázala postihnout co nejširší oblasti strukturního pole textu, které lze ve shodě s formálními metodami kvantitativně-korpusového přístupu objektivně analyzovat a taxonomizovat. Dílčí analýzy by tak měly být kombinovány, abychom se neomezovali na jedinou a izolovanou sondu do struktury uměleckého textu. Jestliže se východiskem stala analýza tematických oblastí, resp. kvantitativní stanovení centrálních tematických pásem, která jsme dále pomocí korpusových nástrojů analyzovali (kolokace), pak nyní přistoupíme k návaznému kroku, který spočívá v rozboru sémantické třídy substantiv. Ve sledovaných (sub)korpusech nás budou zajímat nejen frekvenční průběhy substantiv a jejich kontexty, ale také distribuce do sémantických tříd a vzájemné poměry.

Následující pracovní postup spočívá ve vygenerování frekvenčního seznamu substantiv z každého (sub)korpusu (viz Tab. 26, 28, 33, 39, 41) a v návazné formální taxonomizaci do sémantických tříd (viz Schéma 16–18).

<i>r</i>	lemma	<i>Af</i>	%	<i>Rf</i>
1	oko	93	1.9	4265,47
2	hlava	75	1.5	3439,89
3	ruka	58	1.2	2660,18
4	matka	53	1.1	2430,86
5	okno	52	1.1	2384,99
6	otec	50	1.0	2293,26
7	den	50	1.0	2293,26
8	Rudolf	47	1.0	2155,67
9	Ludvík	45	0.9	2063,94
10	cesta	42	0.9	1926,34
11	žena	41	0.8	1880,48
12	člověk	40	0.8	1834,61
13	Jeník	40	0.8	1834,61
14	tvář	32	0.7	1467,69
15	pole	31	0.6	1421,82
16	svět	29	0.6	1330,09
17	noha	29	0.6	1330,09
18	kráva	27	0.5	1238,36
19	chvíle	27	0.5	1238,36
20	země	26	0.5	1192,50
21	čas	25	0.5	1146,63
22	vůz	25	0.5	1146,63
23	voda	25	0.5	1146,63
24	tráva	25	0.5	1146,63
25	hlas	25	0.5	1146,63
26	silnice	24	0.5	1100,77

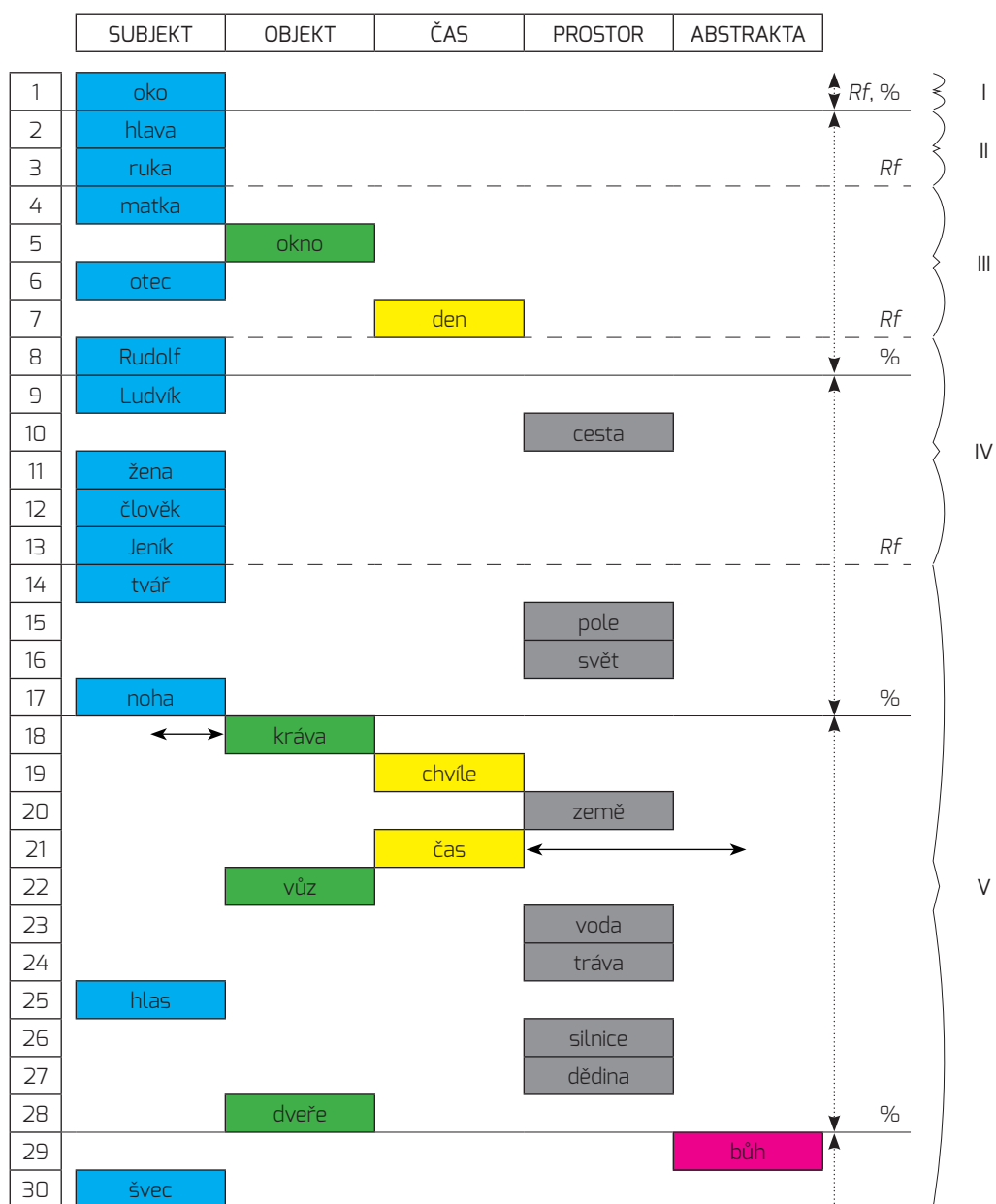
27	dědina	24	0.5	1100,77
28	dveře	23	0.5	1054,90
29	bůh	22	0.4	1009,04
30	švec	21	0.4	963,17

Tabulka 26: Prvních třicet substantiv (lemmat) v (sub)korpusu *Dvojí domov* (1926). Procento určuje zastoupení příslušného lemmatu ve třídě substantiv, hodnota relativní frekvence (*Rf*) je potom normalizovanou hodnotou vypočítanou vzhledem k velikosti (sub)korpusu ($N = 21\,803$).



Graf 7: Frekvenční průběh prvních třiceti substantiv (lemmat) dle relativní frekvence v (sub)korpusu *Dvojí domov* (1926).

Podle průběhu frekvencí můžeme vysledovat určité přechody, které tvoří výraznější předěly mezi frekvenčními hodnotami jednotlivých substantiv a jejich procentuální distribuci v množině všech substantiv (viz Graf 7). Tyto dvě tendence nám ukazují jednak kvantifikační hledisko definované hodnotou relativní frekvence, jednak procentuální zastoupení lemmatu v korpusu substantiv, kdy jsme rozlišovací hranici mezi jednotlivými bloky stanovili rozsahem 0,5 %. Vzájemné průniky těchto hodnot nám určují vztahy mezi frekvenční distribucí a vlastním vnitřním rozložením typu substantiv uvnitř slovnědruhové skupiny. U (sub)korpusu *Dvojí domov* (1926) sledujeme, jak hranice frekvencí protínají oblasti distribuce jednotlivých substantivních lemmat, což *de facto* hierarchizuje uspořádání těchto substantiv. Jinými slovy, oblasti segmentované procentuálně jsou dále parcelovány dle frekvenčního symptomu (*Rf*), který je hodnotou vztahující se k celému (sub)korpusu, tj. v rámci struktury všech lemmat. Uspořádáme-li navíc substantiva do sémantických tříd a vyznačíme-li zde uvedené hranice, získáme následující tabulku (viz Tab. 27), která kromě rozložení substantiv do sémantických tříd ukazuje další vnitřní členění podle frekvence a procentuálního zastoupení lemmatu v třídě substantiv. V případě (sub)korpusu *Dvojí domov* (1926) tak vzniká jakýsi kaskádovitý efekt, jenž výrazně hierarchizuje skupiny substantiv.



Tabulka 27: Distribuce substantivních lemmat do sémantických tříd v (sub)korpusu *Dvojí domov* (1926). Vodorovná plná čára vyznačuje procentuální distribuci v rámci kategorie substantiv, vodorovná přerušovaná čára je hranicí výraznějších frekvenčních přechodů (zlomů). Lemma *kráva* lze na základě analogie (metafory) vztahovat ke kategorii subjektu. Osud zvířete v povídce *Dvojí domov* tvoří analogii k osudu ústřední ženské postavy (srov. Kubiček 2014, s. 149).

Na uvedených distribucích vidíme, že jednotlivé zlomy (hranice) segmentují substantiva podle jejich příslušnosti k sémantické třídě. V (sub)korpusu *Dvojí domov* (1926) přirozeně sledujeme jednoznačnou převahu třídy +HUM (subjekt), ke které se postupně

připojují další sémantické třídy s atributy času a prostoru. Teprve od pátého segmentu (V) dochází k pozvolnému zaplňování sémantických tříd. Výrazná hierarchie vztahu „subjekt/objekt–čas–prostor“ je zcela zřejmá s tím, že příznakovou kategorií je jednoznačně subjekt. Uvedenou distribuci lze podle daného kritéria dále modelovat (viz Schéma 16) s ohledem na kontext (kolokaci). Struktura schématu vyplývá ze vztahů mezi sémantickými třídami, které vyplývají z formální kvantitativně-korpusové analýzy.

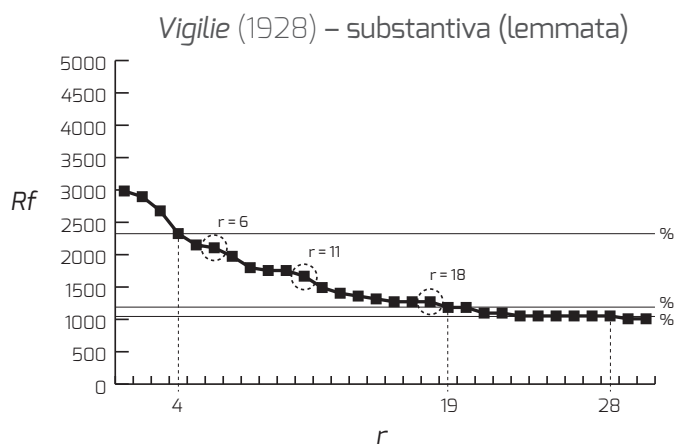
Kontexty nám ukazují, jak jsou jednotlivé kategorie propojeny. Lemma *oko* souvisí nejen se subjektem, ale přirozeně i s časem a prostorem, lemma *cesta* zase se subjektem a pohybem (viz lemma *jít*), *pole* s lemmatem *cesta* a pohybem atd. Navíc uspořádání modelu je *de facto* hierarchickou transformací konkrétně textově realizovaných hyponymně-hyperonymních vztahů. V modelu (viz Schéma 16) jsou sémantické kategorie vyznačeny tak, aby z kategorie nadřazené vyplývaly kategorie závislé, zatímco povrchová textová struktura *Dvojího domova* (1926) realizuje opačnou strategii. Tato struktura signalizuje určitý umělecký postup (záměr), jehož nutnou a nedílnou součástí jsou i vyšší strukturně-významové vrstvy fikčního narativu, které se k uvedené struktuře funkčně vztahují. S ohledem na výraznou dominanci pásma kategorie +HUM (subjekt) a celkovou stratifikaci ostatních lemmat a sémantických tříd se lze domnívat, že se budou nejspíš nacházet v bezprostředním modálním rádiu subjektu (postavy). Jinými slovy, nejenže subjekt bude centrální tematickou osou výše vyznačeného základního vztahu „subjekt–čas–prostor“, což ostatně platí pro všechny zde analyzované Čepovy knihy, ale především způsob uměleckého řešení této struktury bude mít zcela specifické parametry, které ve funkčním propojení s kategoriemi vyšších rovin budou výše uvedený základní vztah modelovat umělecky jinak, než jak je tomu u ostatních souborů. S touto problematikou úzce souvisí otázky narativní strategie, vyprávěcího hlediska, kompozice, ale též pojetí času a prostoru.

Stejný postup, který jsme provedli u (sub)korpusu *Dvojí domov* (1926), budeme aplikovat i na ostatní (sub)korpusy.

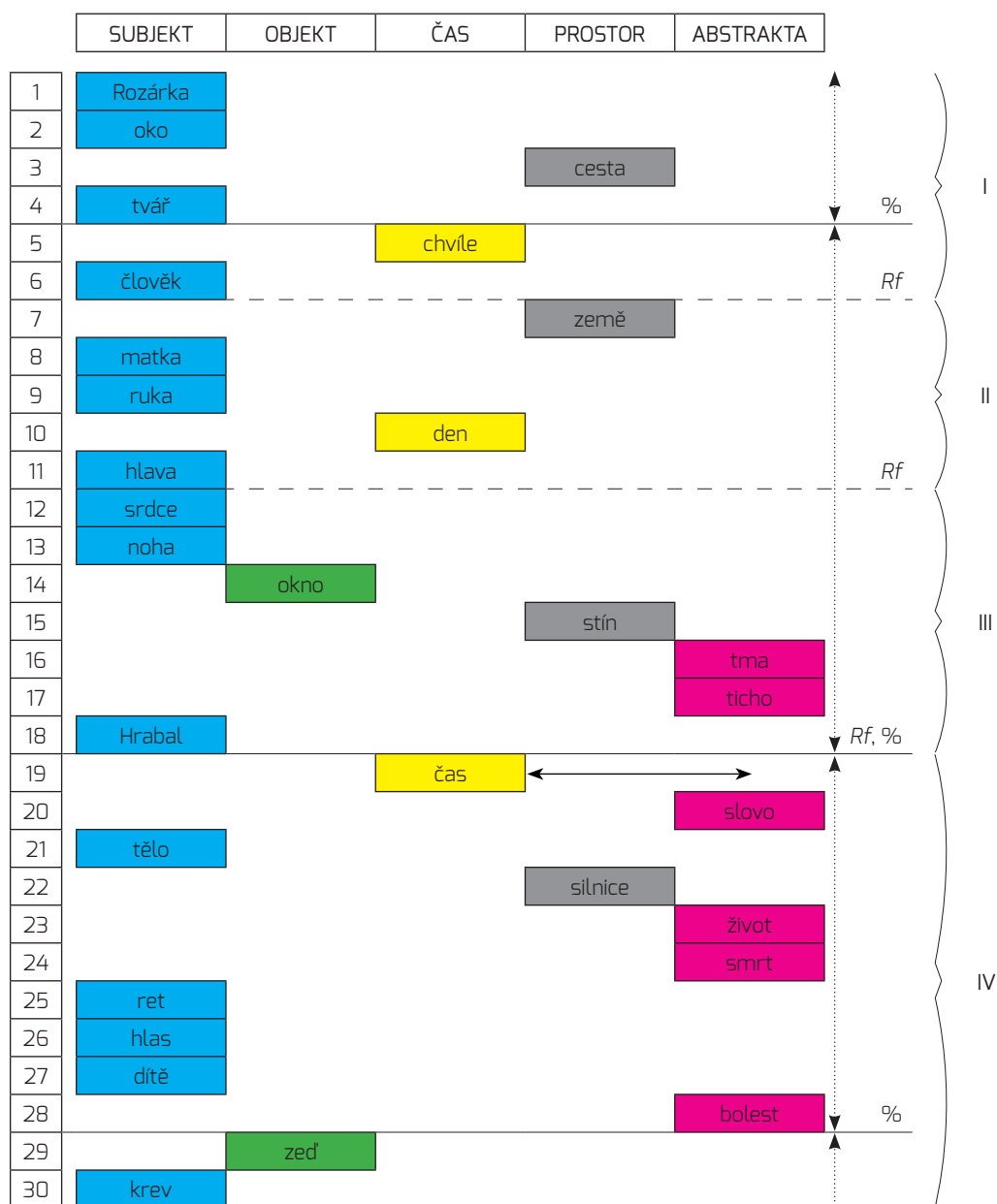
<i>r</i>	lemma	<i>Af</i>	%	<i>Rf</i>
1	Rozárka	68	1.3	2984,03
2	oko	66	1.3	2896,26
3	cesta	61	1.2	2676,85
4	tvář	53	1.0	2325,79
5	chvilé	49	0.9	2150,25
6	člověk	48	0.9	2106,37
7	země	45	0.9	1974,72
8	matka	41	0.8	1799,19
9	ruka	40	0.8	1755,31
10	den	40	0.8	1755,31

11	hlava	38	0.7	1667,54
12	srdce	34	0.6	1492,01
13	noha	32	0.6	1404,25
14	okno	31	0.6	1360,37
15	stín	30	0.6	1316,48
16	tma	29	0.6	1272,60
17	ticho	29	0.6	1272,60
18	Hrabal	29	0.6	1272,60
19	čas	27	0.5	1184,83
20	slovo	27	0.5	1184,83
21	tělo	25	0.5	1097,07
22	silnice	25	0.5	1097,07
23	život	24	0.5	1053,19
24	smrt	24	0.5	1053,19
25	ret	24	0.5	1053,19
26	hlas	24	0.5	1053,19
27	dítě	24	0.5	1053,19
28	bolest	24	0.5	1053,19
29	zeď	23	0.4	1009,30
30	krev	23	0.4	1009,30

Tabulka 28: Prvních třicet substantiv (lemmat) v (sub)korpusu *Vigilie* (1928). Procento určuje zastoupení příslušného lemmatu ve třídě substantiv, hodnota relativní frekvence (*Rf*) je potom normalizovanou hodnotou vypočítanou vzhledem k velikosti (sub)korpusu ($N = 22\,788$).



Graf 8: Frekvencní průběh prvních třiceti substantiv (lemmat) dle relativní frekvence v (sub)korpusu *Vigilie* (1928).



Tabulka 29: Distribuce substantivních lemat do sémantických tříd v (sub)korpusu *Vigilie* (1928). Vodorovná plná čára vyznačuje procentuální distribuci v rámci kategorie substantiv, vodorovná přerušovaná čára je hranicí výraznějších frekvenčních přechodů (zlomů).

Oproti distribuci sémantických kategorií v (sub)korpusu *Dvojí domov* (1926) zde sledujeme dvě základní tendence. Ačkoli vlastní kategorie subjektu je i v tomto případě dominantní, poněkud se oslabuje její výsadní postavení zvýrazněním kategorií času a prostoru. Dochází tak k jakémusi „zředění“ kategorie subjektu. Oproti předchozímu

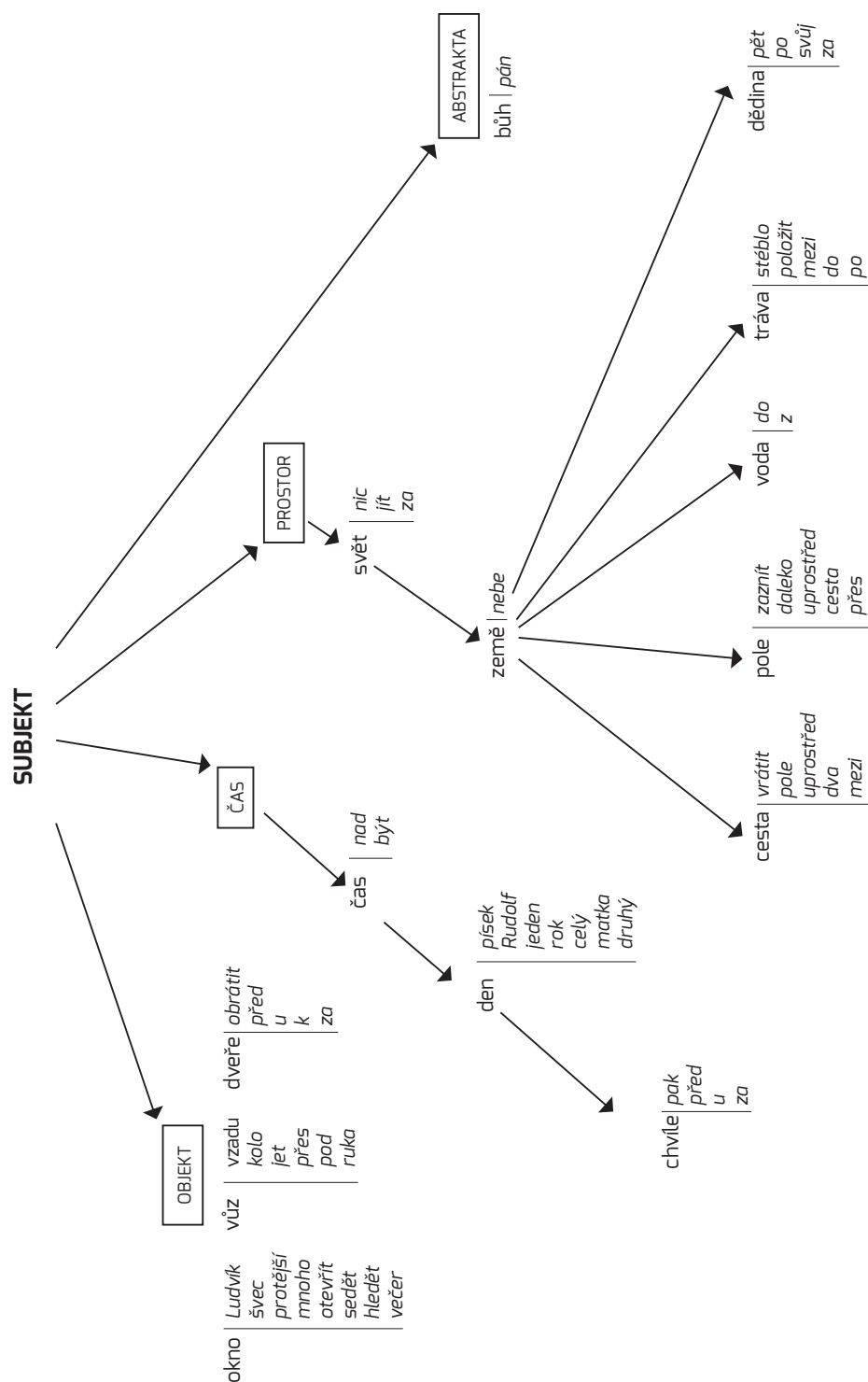


Schéma 16: Hierarchie a vztahy sémantických tříd nejfrekventovanějších třiceti substantiv v (sub)korpusu *Dvojí domov* (1926). Kurzivou jsou vyznačeny nejčastější kontexty (kolokace) příslušného lemmatu.

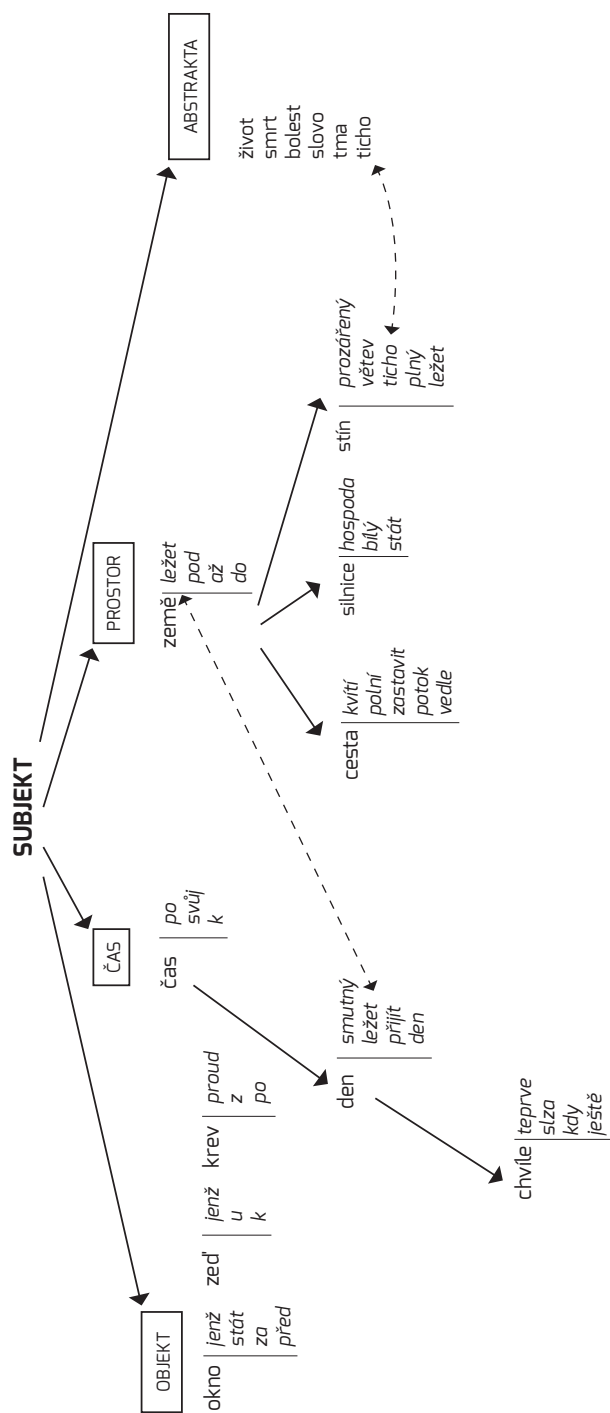


Schéma 17: Hierarchie a vztahy sémantických tříd nejfrekventovanějších třiceti substantiv v (sub)korpusu *Vigilie* (1928). Kurzivou jsou vyznačeny nejčastější kolokace (kontexty).

(sub)korpusu se navíc výrazněji do popředí dostává kategorie abstrakt, která se pojí s kategorií prostoru. Pochopitelně je přítomná i v (sub)korpusu *Dvojí domov* (1926), ale v aktuálním (sub)korpusu se dostává do významnější oblasti, kterou podmiňuje frekvenční aspekt. Kromě toho pozorujeme, že distribuční hranice relativních frekvencí a procent tvoří mnohem pozvolnější přechody, než tomu bylo u předchozího (sub)korpusu. S ohledem na stratifikaci sémantických tříd lze usuzovat na proměnu invariantního paradigmatu „subjekt–čas–prostor“, s čímž opět bude souviset funkčně-sémantická modifikace literárněvědných kategorií na vyšších strukturních úrovních. Na uvedených kolokacích lemmatu *cesta* (viz Schéma 17) konkrétně vidíme, jak se proměňuje zapojení uvedeného lemmatu do kontextu, což se odráží v jeho sémantice (srov. Schéma 16). Kontextové okolí lexému *země* vyznačené konkordančním výpisem (viz Tab. 30) ukazuje, že jeho sémantika se v (sub)korpusu *Vigilie* (1928) jemně posunuje do kategorie abstrakt. *Země* tak vedle svého nociónálního významu výrazněji odkazuje k symbolické rovině.

1	byla plna prozářeného stínu , stíny ležely na jitrní orosené	zemi	a na měkké trávě , zachvívajice se hlubokým štěstím -
2	uvidět svůj osud a sebe sama , jak trčí ze	země	podobn zmrzačenému trupu , poset ranami
3	Cesty tajemně přikované právě k tomuto okrsku	země	marně nalhávaly svou linií , ubíhající k dálkám , že
4	ale pak se kočka znovu shýbla k tomu černému na	zemi	a dala se znovu do žrádla , jeho si už
5	Stařec sundal syna s žebříku a položil ho na	zem	doprostřed dvorku . Brada se mu odchlípla a zuby
6	Šest lidí , schýlených nad hroudou a rvoucích ze	země	žádostivými prsty její plod , nikdy nenasycených
7	najednou zahlédl . Seděla shrbená u hranice dříví , na	zemi	před ní leželo cosi černého a až k němu bylo
8	tiše stálo v modři a zlatě schylujícího se slunce a	země	se prostírala do dálky v sladké únavě zářivého odpoledne
9	zestárlá výrazem skrytého utrpení , jež užírá kdesi pod	zemi	chlapcova života a žhne bolestně ve velikých černých očích ,
10	chodím dokola jako vrtolavý a hledám znovu a znovu ,	země	mlčí , neotvírá se . . . Rozárka Lukášová Jeroným
11	„ učití najednou stařec , a jak vstává se	země	„ jako by mu sklouzly po rameni čísi roztažené prsty
12	vdechla zhluboka nejomamnější vůni této	země	a zalomivši rukama záupěla tak , že se Rozárka
13	nových sousedů a seděla záhy pevně , zabořivši se do	země	svou vlastní tíhou . Cihelné průčelí bývalého gruntu , jehož
14	postava v bílých šatech jako by se náhle položila na	zem	„ Jiří měl sotva kdy , aby se vzpamatoval
15	vysoko vylítna slepice do vzduchu a jak padá obloukem k	zemi	„ až její tělo udeří hluchou ranou do prachu
16	bizarní balvan , nesmírně tvrdý a těžký , vražený do	země	vzdorovitou silou , a rezavá korouhvička tesklivě skřípe
17	mají všichni lidé spojit dohromady a potom bude dobře na	zemi	„ Svoboda , rovnost , bratrství , jak řekl jeden
18	ani rukou a nelze ji nijak strážat a přimáčknot k	zemi	„ Nejlíp je dívat se kolem sebe a jmenovat známé
19	nehýbe , stojí s rukama v kapsách a hledí do	země	„ a změří - li občas očima splající hlouček ,
20	tesklivě skřípe nad okny zatahujícími se zoufalou slepotou .	Země	kupodivu nevládná a lysá táhne se k obzoru jako vymřelý
21	zarýval jsem prsty v zoufalství do	země	„ Co já tu sám , co já tu sám
22	Tiše a tesklivě položil se přede mne na holou	zemi	můj nekonečný stín . Pod vrátky poslední chalupy protáhl se
23	průduchy , neočekávaně na okamžik pootevřené , znova stáhly a	země	se zdála pevná a bezpečná pod nohama starce dosud marně
24	rakev , bez jediné ozdůbky , a vltká díra v	zemi	zavřela v sobě hrozný hřích , obrátili se pozůstalí bázlivě
25	probudila se náhle jakoby něčím dotekem . Vyskočila se	země	„ ve světnici byla už úplná tma , jen bílá
26	nesmírného prostoru nad hlavou , ohýbá se až do prachu	země	„ ale plazí se dále , vzpírá se urputně
27	útek , tma je bez směru a bez konce .	země	se vyšínula ze své dráhy a padá , padá jako
28	Podzim se však záhy přesunul přes období melancholické tklivosti .	země	stydla a stromy co den obnaženější trčely do prostoru stále
29	před tajemným výbuchem , který doutná hluboko pod	zemi	„ a nad vším hemžením , nad bezhlavým mumrajem ,
30	a její oči , strašné otvory uprostřed křabaté tváře barvy	země	„ otvory , jimiž hledí stařenina úzkost , příšerný hlad
31	lidmi , a podél cesty trčí odrané stromy a leží	země	proniklá smrtí do nejhlubších útrob . A tu pojednou na
32	s ranní modlitbou posílali za mnou své vzpomínky do dalekých	zemi	„ netušili , že na jejich prahu nezchladí ještě dotek
33	Petr Kleofáš zpola klečel , zpola ležel na vychládající	zemi	„ oči měl zavřeny a tiše počítal zvolňující se tepot
34	ti pomáhám ! „ Rozárka se domodlíla , vstala se	země	a ohlédnuvši se spatřila vůdce procesí , opřeného o hůl

35	a hrůzu loučení na neshledanou , a nejvíc příšerné mlčení	země	i nebe na všecko lidské úpění a všecku úzkost ,
36	který by ho skryl . Věděl , že je teď	země	vylidněná a nikoho se nelze dovolat , opuštěné zdi trčí
37	už chvíli nehybně podle těla a bič mu sklouzl na	zem	. Najednou sebou trhl : jako by se byla mihla
38	v houštině do mechu a ležel jsem tam tvář k	zemí	až do večera . Pak přišla zima a bylo by
39	neodvažoval ohlédnout se za sebe , ale tiskna se k	zemí	toužil mít v té chvíli srdce probodeno . Toužil cedit
40	zidky stojí opuštěný , vyhlodaný brus , hluboko zabořený do	země	. Nemám odvahy přiskočit potok a jít na dvůr .
41	- a žížeň po krvi je stále větší , pukliny	země	se otvírají šíř a šíře , srdce jsou vyprahlá a
42	zachvácen pocitem strašné samoty uprostřed tmy a vychladlé	země	rozprostřené do nedozírna , zapomenut ode všech
43	ložem , a cítě se bez ochrany mezi nebem a	zemí	, marně hledal kout , který by ho skryl .
44	jež ho rdousila , cítě se najednou přirostlý k	zemí	všemi klouby svého těla . Zápas byl dlouhý , pařez
45	za zvyk každodenní práce , ale pod nohama se mu	země	dosud povážlivě kolísá a v hrdle cítí stále sevření oné

Tabulka 30: Konkordance lemmatu *země* v (sub)korpusu *Vigilie* (1928).

1	Také je krásné rozorávat pluhem	zemí	nebo být myslivcem a chodit s puškou po lese ,
2	ale je mu líto mrtvých . Mají barvu jako	země	a jsou věčné s ní . Mrzí ho , když
3	a šel k domovu , urputně zabodává podpatky do vzdorovitě	země	- - - Eliášův grunt je dnes zamklý . Čeleď
4	poslední mandel zůstal na poli , nebe se řítilo k	zemí	těžkými proudy deště . " Stoupněte na kolo ! "
5	vrabci , kolena mu podklesávají a pivo se vylilo na	zem	. Stařečku , není pomoci , musí se jít domů
6	jako svět , větvicky s měkkým listím visely až k	zemí	a zelené ticho spouštělo tu křídla jako snící pták .
7	mez a díval se , jak brouci lezou z pórů	země	a šplhají po stéblech trávy . " Proháněj se mezi
8	vůz byl už uprostřed polí , krajina se tiskla k	zemí	jako koroptev , ženy utíkaly bosa , s hlavami sehnutými
9	je , dáváje jim sladká jména . Položil se na	zem	a líbal trávu , tuto vonící a měkkou věc .
10	mléka octla se ve hnoji , zdvihla se Amalie se	země	a belhající a nadávající , tloukla kozu po hlavě
11	Frantik zůstal sám s mrtvým tělem otcovým , ležícím na	zemí	tváří dolů . Vzpoura Sedlák Eliáš byl zle postižen svým
12	poutníku , který jdeš do světa kolem chalup hluboko do	země	zařatých , viz kapky krve na kamení ! A neposmívej
13	spadly na kolena a otec se svezl po fůře k	zemí	, zatímco matka se chytla klance , bledá jako smrt
14	zpozorovav , že zapadl až na Studýnky , kde	země	je vždycky černá a kvícaly pišti " pit , pit
15	Cesta se strachem přikrčila , krásy svěřily hlavy až k	zemí	a vůz přestal klapat , avšak otec si přitlačil slaměný
16	zahrady za hromadu hlíny . Kusem starého železa vyhloubil do	země	důlek a pochová do něho obtížný dar . Pak se
17	chránily , aby nezapašily ticho , táhnoucí s šumotem nad	zemí	. Cilka odcházela z domu , až všechno spalo ,
18	a nejmíň stokrát se pohádali a Frantik jednou praštil o	zem	kusem dřeva a zvolal : " Fiks mordie , že
19	jako krev , prosakující Ludvíkovou košilí . Elegie Chaloupka u	země	, křívá , za chaloupkou potok . Dětský ráj ,
20	A nevěda si rady , klekl na	zem	a drže ji na rukou , zdvihal ji vzhůru
21	kdežto lány obilí modlí se kleče , naklánějící se k	zemí	hlavami klasů . Snášejí tíže deště i slunce i blesky
22	chvěje v černém větru , neboť černé je vše ,	země	, nebe . . . Jeník vzhlédl , otcova bílá
23	vždy dříve , než se přihodilo neštěstí . Přitisknuta k	zemí	spínala ruce a věřila , že Bůh vidí vše .
24	se nezakl sladkou pokorou , jež ho tiskla k rozmoklé	zemí	a zároveň ho zbavovala tíže . Hledal očima otce ,
25	že mě Bůh ranil slepotou ? Není jí ,	země	je pustá . Potok bublá uprostřed holých břehů , zrcadle
26	se a ohromné tělo otcovo vznášelo se mezi nebem a	zemí	a každou chvíli se zdálo , že se utrhne

Tabulka 31: Konkordance lemmatu *země* v (sub)korpusu *Dvoji domov* (1926).

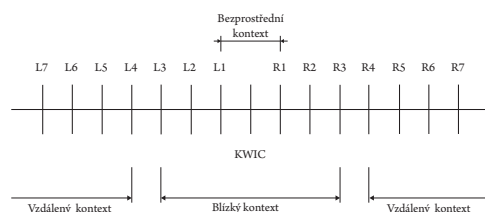
Z uvedených konkordancí je patrné, že lexém *země* se objevuje v kontextech, které utvářejí jeho nociónální a symbolický význam. Vidíme, že je tomu tak v obou (sub)korpusech. Co se však mění, je, že lemma *země* je v (sub)korpusu *Vigilie* (1928) nejen frekvenčně zatíženější oproti (sub)korpusu *Dvoji domov* (1926), ale že se ocitá v modifikovaném strukturním poli sémantických tříd (viz Tab. 29). Význam pro funkční modifikaci motivu země má i skutečnost, že kategorie subjektu v (sub)korpusu *Vigilie* (1928) nepřesahuje tak výrazně jiné sémantické třídy. Zdá se, že mezi motivem chůze – viz lemma *jít* v (sub)korpusu *Vigilie* (1928) –, kategorií subjektu a prostoru bude ve jmenovaném

(sub)korpusu vztah, který mnohem více zdůrazní symbolický aspekt země a prostoru ve vztahu k motivu cesty.

Zajímavé jsou i blízké či bezprostřední kontexty²⁰³ lemmatu *den* v (sub)korpusu *Vigilie* (1928), které významněji specifikují význam temporální statickosti (*ležet, celý, následující, uprostřed dní*) nebo následnost (*přijít, den ode dne, každým dnem*).

1	žebrákům a pocestným a pro postavení kříže nebo kapličky .	Dni	byly tiché a tklivé , jak bývají po žních .
2	prostor jejich obydlí je špinavý a smutný , den ze	dne	tyž , a včerejší nešvár kvasí na dně kbelíku .
3	k jehož slávě umírali mučedníci a hořela města . Druhého	dne	kvečeru klečela Rozárka před obrazem Matky sedmiboletné
4	jí tekly po tvářích bez jediného hlesnutí . Ve vzduchu	dní	, které přišly potom , bylo už cítit vanutí smrti
5	obklopuje ho vanutí nekonečné prohlubně . V následujících	dnech	se tajemné průduchy , neočekávaně na okamžik pootevřené
6	vdá , bude zas místa dost , " řekla jednoho	dne	macecha , stojíc s vyhrnutými rukávý
7	zvracení oškívá , se srstí řídkou a spečenou , ve	dne	ležela obyčejně na sluníčku a lysající břicho se jí pohybovalo
8	se nepřítis časné . Klempířské nůžky a kleště ležely celé	dny	netknuty a po podlaze válely se kusy plechu se zoufale
9	dcerku na čelo a odkvapil ven . Uplynulo ještě několik	dnů	a oblouk Rozárcina utrpení , přitahovaný tajemným středem
10	v tomto živobytí bez oddechu a bez zastavení , uprostřed	dní	plných špíny a starostí úžících dech , uprostřed nocí
11	takže mu rázem uvěřila . Za okny modral se	den	, mužova postel byla prázdná - Rozárka zavřela oči
12	a zděsila se pomyšlení , že by měla žít o	den	děle . Leč na dvoře už nakládali na vůz její
13	Anděl Páně " . Ale má trýzeň byla den ode	dne	větší . Byl jsem krutě trestán za to , že
14	a po němž teď spěchají schoulené postavy , strávivší	den	v úzkém prostoru kanceláří , ubožáků zmrzačených na těle
15	Tak šel i za rakví synovou , když ho třetího	dne	na úsvitě vezli v deskovém voze k farnímu hřbitovu .
16	vzrušením , jež se podobalo opojení . Ale potom přišly	dny	plné starostí , pod nimiž lze sotva vydechnout , malý
17	děje . Zdálo se jí docela jisto , že ten	den	se neskončí večerem a nocí jako jiné dny . Bylo
18	neboť nebudeš mít podobných dvou sluhů do konce svých	dní	! Nevím už ani , který z nich odešel dřív
19	Na jakou šňůru navlékají ti lidé kalné perly svých	dní	? Čas se jim sune ztěžka od pondělka do soboty
20	drůbeží , zatímco matka leží v hrobě . V prvních	dnech	horké slzy tekly nepřetržitě , ale někdy bolest jako by
21	Bože můj , nenechávej mne v tomto hrozném mlčení !	Dni	do pohřbu uplynuly v bolestném strnutí , jímž se co
22	žhoucí a vytepává ho jako plíšek zlata . Toho	dne	byla už chvíle soumraku s rudými skvrnami na stěnách
23	a Rozárka se skácela do vozu jako bezduchá . -	Dni	v městečku , kam ji muž zavezl , mýjely jí
24	jejich " Anděl Páně " . Ale má trýzeň byla	den	ode dne větší . Byl jsem krutě trestán za to
25	nabývá rysů tvrdších a bolestnějších . Každým	dnem	se množí nářky , chrastí lucerny a rozbitá okna ,
26	období melancholické tklivosti , země stydla a stromy co	den	obnaženější trčely do prostoru stále prázdnějšího .
27	se v něm měkce zachvívaly . Rozárka ležela tiše celé	dny	a byla sama jako smutný a tklivý stín , jako
28	hlíny . se vši svou přichýlností k drobným věcem všedních	dní	, se vši zarytostí a tuhostí své lidské přirozenosti ,
29	duši . . . A na těch cestách , když	den	se nacyloval a stíny ožívaly , zmocňoval se mne pocit
30	Tráva už začala vyrážet , housata dorůstala a přišel	den	, kdy starý Vrána musel sundat s hřebíku svůj dlouhanský
31	s houštím kopřiv , po jejichž omítce vídala se sklonkem	dne	stoupatí stín , pěšinka v zahradě mezi dvěma hranicemi dříví

203 Bezprostřední kontext „označují takové jednotky, které stojí na pozici přímo předcházející klíčovému slovu nebo za ním přímo následující. V korpusové notaci jsou to pozice L1 a R1“ (Cvrček 2013, s. 25). Blízký kontext „jsou [...] maximálně tři pozice od hledaného klíčového slova. Blízký kontext tak zahrnuje kontext bezprostřední i jednotky o dvě pozice vzdálenější. V korpusové notaci jsou to pozice L3, L2, L1 a R1, R2, R3“ (Ibidem, s. 25).



(Ibidem, s. 25)

32	Vždyť prostor jejich obydlí je špinavý a smutný .	den	ze dne týž , a včerejší nešvár kvasí na dně
33	jiní se při těch slovech křičovali . Husy se toho	dne	potulovaly svobodně po návsi , porůznu kejhající
34	otcova zdála se jí táž od prvního dětství až do	dne	její smrti : rozkazovala bez odvolání . Tíha starostí
35	zvlhl a oteplil se . Pahorky a svahy krajiny přes	den	přísné a zachmuřené jako by ožily v temnu a staly
36	jichž nelze oběsit ani zardousit a jež leží tři	dny	uprostřed dvora a nemohou pojit . Nuže , tento zdechající
37	že ten den se neskončí večerem a nocí jako jiné	dny	. Bylo jí , jako by ji měl kdosi vzít
38	do žrádla , jeho si už nevšímajíc . - Druhého	dne	před polednem ji zase potkal , jda pěšinkou pod stromy
39	jsem bloudil noc co noc okolo našich plotů a za	dne	jsem se skrýval po roklínách . Zpočátku jsem se styděl
40	tak utýrané , tak pokojné , a zvláště na sklonku	dne	tak důvěrně se vinoucí k duši . . .

Tabulka 32: Konkordance lemmatu *den* v (sub)korpusu *Vigilie* (1928).

Rozdíly oproti kontextovému okolí uvedených lemmat v (sub)korpusu *Dvojí domov* (1926) spolu s celkovými tendencemi ve strukturaci a stratifikaci sémantických tříd naznačují, že je s významy času a prostoru ve *Vigiliích* (1928) zacházeno poněkud jinak, než tomu bylo ve *Dvojím domově* (1926). Současně se ukazuje, že tato změna není radikální tematickou změnou, jako spíše proměnnou v rámci invariantní tematiky. Lze se domnívat, že Čep ve *Vigiliích* nemění téma, ale volí jinou uměleckou strategii.

V souvislosti s analýzou tematických oblastí jsme konstatovali, že v (sub)korpusu *Vigilie* (1928) dochází k oslabení frekvenčně koncentrované kategorie subjektu a že spolu s tím se uvolňuje i její dostředivý sémantický rádius. Jak uvidíme v následující kapitole, tyto jevy souvisí s proměnou narativní strategie a promítají se do pojetí času i prostoru. Rovněž jsme konstatovali, že sjednocujícím principem se ve *Dvojím domově* (1926) stává synekdocha. Synekdochický princip sledujeme také v korpusové struktuře *Vigilií* (1928), i když v modifikované podobě. Tato srovnání nás upozorňují na určité posuny a změny v strukturních organizacích těchto textů, které souvisí s komplexními významovými posuny. Na uvedených srovnáních si lze rovněž povšimnout, jak princip synekdochy doslova „prorůstá“ jednotlivé kategorie a jak se manifestuje jejich hierarchickým uspořádáním, jehož formálním projevem jsou vedle frekvence hyponymně-hyperonymní vztahy. Na základě dosavadních zjištění můžeme dokonce usuzovat, že princip synekdochy tvoří hloubkovou strukturu textů, což zakládá domněnku, která odkazuje k axiologické soustavě Čepových fikčních světů. Tyto základní kvantitativně-korpusové analýzy musejí ovšem probíhat v přímé interakci s konkrétními texty. Mnohé z toho, co bylo nejnověji konstatováno při strukturálně-sémiotickém rozboru Čepova *Dvojího domova* Tomášem Kubíčkem (2014), potvrzují také naše zjištění. Jestliže Kubíček například upozorňuje na povahu času, jehož součástí je kromě vrstvení (2014, s. 33) i cykličnost (Ibidem, s. 34), dokládá tuto základní skutečnost i kvantitativně-korpusová analýza.²⁰⁴ Vysoce

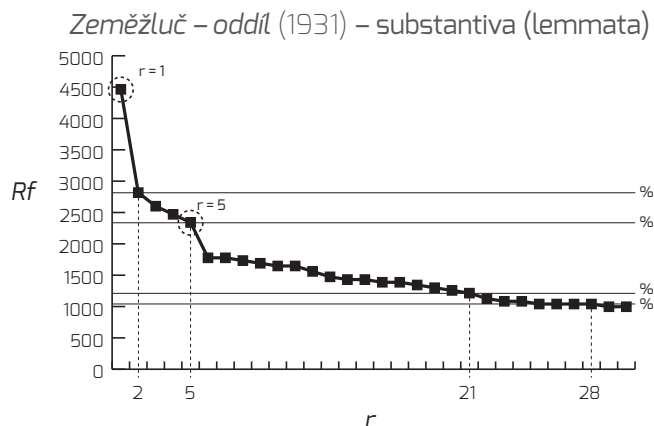
²⁰⁴ Základní poznatky, které o Čepově *Dvojím domově* přinesl ve své knize Kubíček, lze ověřit na pozadí kvantitativně-korpusové analýzy. Tato shoda je *de facto* potvrzením správnosti obou postupů a dokladem funkční koherence strukturálně-sémioticky orientované literárněvědné analýzy a kvantitativně-korpusové metody.

frekventované lemma *den* je nejen strukturně propojeno s kategorií subjektu (den: *Rudolf, matka*), ale jeho okolí rovněž napovídá, že paradigmaticky tenduje k základnímu konstrukčnímu principu synekdochy (den: *jeden, rok, celý*), která je stále soustředěna kolem kategorie subjektu a její perspektivy, zatímco v (sub)korpusu *Vigilie* (1928) nám okolí tohoto lemmatu ukazuje významovou modifikaci spočívající v oslabení kategorie subjektu. Uvedené skutečnosti je ovšem nezbytné konkrétně interpretovat v kontextu celých narativů. Významová modifikace se týká rovněž ústředního motivu cesty, který – jak jsme konstatovali – ve spojení s motivem země a kategorií prostoru má tendenci inklinovat k symbolické rovině.

Zbývá nám provést totéž u posledního (sub)korpusu, kterým je samostatný oddíl *Zeměluč* (1931) stejnojmenného povídkového souboru.

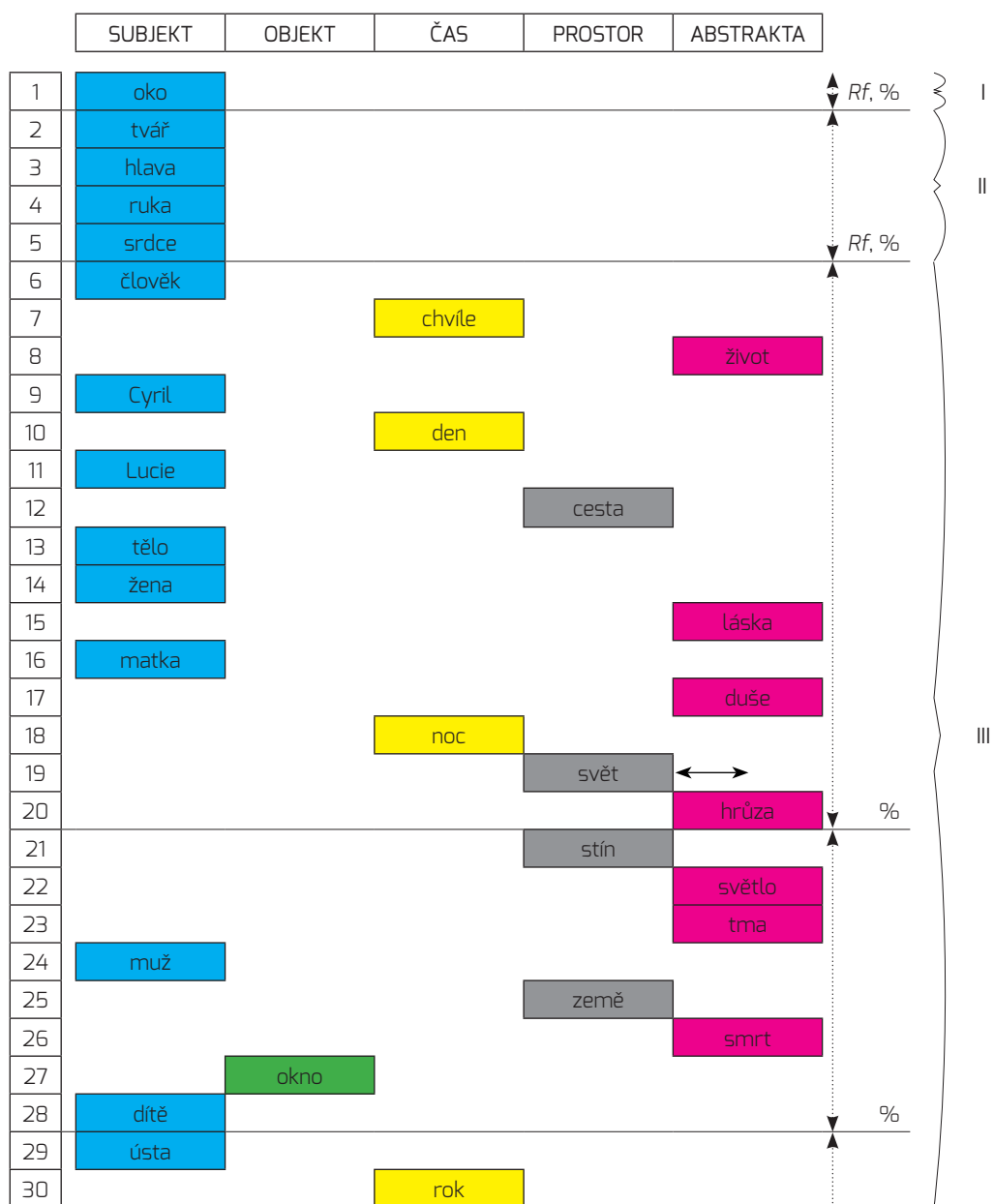
<i>r</i>	lemma	<i>Af</i>	%	<i>Rf</i>
1	oko	103	2,0	4464,29
2	tvář	65	1,3	2817,27
3	hlava	60	1,2	2600,55
4	ruka	57	1,1	2470,53
5	srdce	54	1,0	2340,50
6	člověk	41	0,8	1777,05
7	chvíle	41	0,8	1777,05
8	život	40	0,8	1733,70
9	Cyril	39	0,8	1690,36
10	den	38	0,7	1647,02
11	Lucie	38	0,7	1647,02
12	cesta	36	0,7	1560,33
13	tělo	34	0,7	1473,65
14	žena	33	0,6	1430,31
15	láska	33	0,6	1430,31
16	matka	32	0,6	1386,96
17	duše	32	0,6	1386,96
18	noc	31	0,6	1343,62
19	svět	30	0,6	1300,28
20	hrůza	29	0,6	1256,93
21	stín	28	0,5	1213,59
22	světlo	26	0,5	1126,91
23	tma	25	0,5	1083,56
24	muž	25	0,5	1083,56
25	země	24	0,5	1040,22
26	smrt	24	0,5	1040,22
27	okno	24	0,5	1040,22
28	dítě	24	0,5	1040,22
29	ústa	23	0,4	996,88
30	rok	23	0,4	996,88

Tabulka 33: Prvních třicet substantiv (lemmat) v (sub)korpusu *Zeměžluč – oddíl* (1931). Procento určuje zastoupení příslušného lemmatu ve třídě substantiv, hodnota relativní frekvence (*Rf*) je potom normalizovanou hodnotou vypočítanou vzhledem k velikosti (sub)korpusu ($N = 23\,072$).



Graf 9: Frekvenční průběh prvních třiceti substantiv (lemmat) dle relativní frekvence v (sub)korpusu *Zeměžluč – oddíl* (1931).

Podobně jako v (sub)korpusu *Dvojí domov* (1926) je i v tomto případě exponovaná kategorie subjektu. Co se však proměňuje, je distribuce ostatních kategorií (viz Tab. 34). Z frekvenčního hlediska nabývají abstrakta výraznějšího postavení. Kontextová pole nám ukazují, že propojenost jednotlivých sémantických tříd je zde symptomatictější a explicitnější než u předchozích (sub)korpusů, zejména se to týká relací „subjekt–abstrakta–prostor“ (viz Schéma 18). Nedochozí opět k zásadní proměně tematiky, ale vztahových relací. Korpusové a frekvenční hodnoty nás upozorňují na další modifikaci invariantního paradigmatu.



Tabulka 34: Distribuce substantivních lemmat do sémantických tříd v (sub)korpusu *Zeměžluč* – oddíl (1931). Vodorovná plná čára vyznačuje procentuální distribuci v rámci kategorie substantiv, vodorovná přerušovaná čára je hranicí výraznějších frekvenčních přechodů (zlomů).

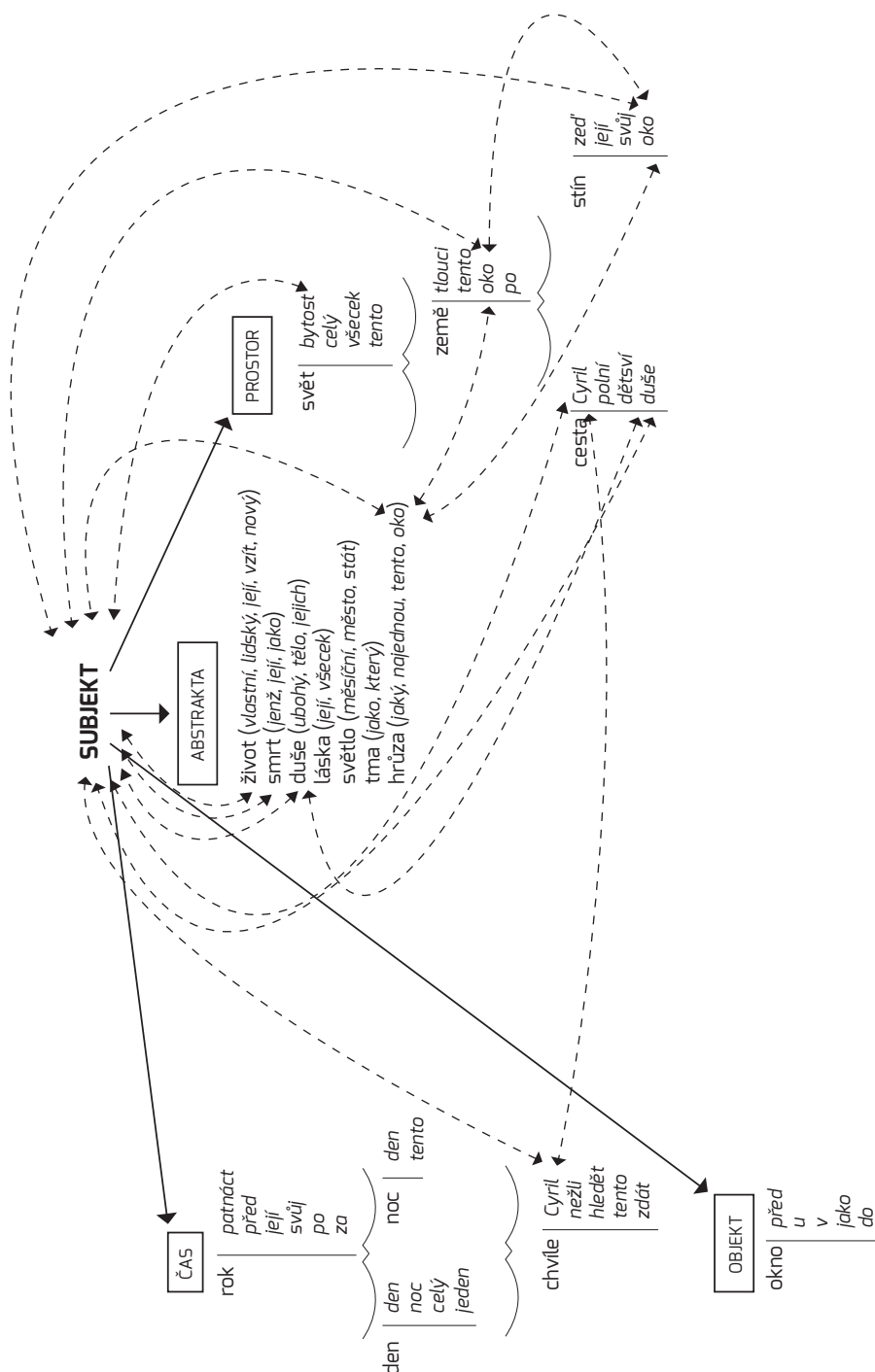


Schéma 18: Hierarchie a vztahy sémantických tříd třiceti nejfrekventovanějších substantiv v (sub)korpusu *Zeměžluč – oddíl* (1931). Kurzivou jsou vyznačeny nejčastější kolokace (kontexty).

Analýza třídy substantiv se nabízí ze dvou základních důvodů:

1. Konkrétní výběr substantivních lemmat tvoří z hlediska tematické oblasti (sub)korpusů příznakovou skupinu. Bylo proto na místě zjišťovat další vztahy uvnitř této třídy, přičemž analýza byla provedena s ohledem na širší sémantickou kategorizaci.
2. Substantiva se zásadně podílejí na vytváření nociónálního rámce fikčního světa.

Na místě je analyzovat také další slovnědruhé kategorie autosémantik, především třídu sloves a adjektiv.

	<i>Dvojí domov</i> (1926)	<i>Vigilie</i> (1928)	<i>Zeměžluč – oddíl</i> (1931)
1	starý	starý	mrtvý
2	celý	bílý	celý
3	nový	černý	poslední
4	rád	plný	bílý
5	otcův	celý	starý
6	malý	mrtvý	lidský
7	veliký	známý	černý
8	bílý	nový	vlastní
9	černý	malý	tichý
10	plný	lidský	nový
11	modrý	poslední	ubohý
12	mladý	jiný	mladý
13	krásný	jediný	tajemný
14	dlouhý	živý	známý
15	bosý	veliký	sladký
16	zelený	tajemný	bledý
17	poslední	strašný	živý
18	jiný	Rozárčín	nesmírný
19	protější	sladký	jiný
20	mrtvý	hluboký	jediný

Tabulka 35: Prvních dvacet nejfrekventovanějších adjektiv (lemmat).

	<i>Dvojí domov</i> (1926)	<i>Vigilie</i> (1928)	<i>Zeměžluč – oddíl</i> (1931)
1	být	být	být
2	jít	jít	vidět
3	slyšet	cítit	tisknout
4	vidět	vidět	jít
5	zůstat	slyšet	řici
6	spát	mít	spát
7	sedět	vzít	spatřit
8	ležet	stát	pohlédnout
9	stát	rozloučit	chtít
10	mít	mluvit	udělat
11	jíst	dělat	stát
12	držet	dívat	sedět
13	chodit	řici	pochopit
14	znát	číst	očekávat
15	vrátit	tisknout	obrátit
16	sednout	přeskočit	objímat
17	plakat	ohlédnout	mluvit
18	očekávat	odejít	klečet
19	moci	modlit	hovořit
20	zakrývat	domnívat	dostat

Tabulka 36: Prvních dvacet nejfrekventovanějších sloves (lemmat).

U prvních dvaceti nejfrekventovanějších adjektiv a sloves nepozorujeme výraznější změny z hlediska jejich typologie. Rozdíly se však místy objevují ve frekvenční distribuci některých slov. Viděli jsme, že například sloveso *jít* má v (sub)korpusu *Vigilie* (1928) mnohem nižší frekvenci než v (sub)korpusu *Dvojí domov* (1926). Naopak v (sub)korpusu *Zeměžluč – oddíl* (1931) se neobjevuje nad *h-bodem*. Všimněme si dále, že v (sub)korpusech *Vigilie* (1928) a *Zeměžluč – oddíl* (1931) přibývá verb dicendi. Z adjektiv je zajímavá distribuce barev; adjektiva *modrý*, popř. *zelený* jsou typičtější pro (sub)korpus *Dvojí domov* (1926), zatímco adjektiva *bílý*, popř. *černý* jsou častější v (sub)korpusech *Vigilie* (1928) a *Zeměžluč – oddíl* (1931), přičemž nejfrekventovanější lemma *bílý* vytváří nejvýraznější frekvenční rozdíly (viz křivku průběhu *Rf* u lemmatu *bílý* – Graf 10) v korpusu fikčních narativů. Otázkou interpretace zůstává, nakolik tyto barvy korespondují s tendencemi, které jsme u (sub)korpusů sledovali.

	NÁZEV (SUB)KORPUSU	Af	Rf
1	<i>Letnice</i> (1932)	27	860,00
2	časopisecké prózy 30. let	5	781,60
3	<i>Tvář pod pavučinou</i> (1941)	38	749,30
4	časopisecké prózy 20. let	52	675,00
5	<i>Dvojí domov</i> (1926)	12	550,40
6	<i>Modrá a zlatá</i> (1938)	28	544,10
7	<i>Děravý plášť</i> (1934)	18	505,70
8	<i>Polní tráva</i> (1946)	28	439,70
9	<i>Cikáni</i> (1953+)	8	339,90
10	<i>Hranice stínu</i> (1935)	19	335,10
11	<i>Vigilie</i> (1928)	7	307,20
12	<i>Zeměžluč – oddíl</i> (1931)	5	216,70

Tabulka 37a: Frekvenční distribuce lemmatu *modrý* v Čepově korpusu fikčních narativů (řazeno podle *Rf*).

	NÁZEV (SUB)KORPUSU	Af	Rf
1	časopisecké prózy 20. let	56	726,90
2	<i>Dvojí domov</i> (1926)	10	458,70
3	<i>Vigilie</i> (1928)	9	394,90
4	<i>Modrá a zlatá</i> (1938)	18	349,80
5	<i>Letnice</i> (1932)	10	318,50
6	<i>Hranice stínu</i> (1935)	11	194,00
7	<i>Děravý plášť</i> (1934)	6	168,60
8	časopisecké prózy 30. let	1	156,30
9	<i>Polní tráva</i> (1946)	9	141,30
10	<i>Cikáni</i> (1953+)	3	127,50
11	<i>Tvář pod pavučinou</i> (1941)	5	98,60
12	<i>Zeměžluč – oddíl</i> (1931)	2	86,70

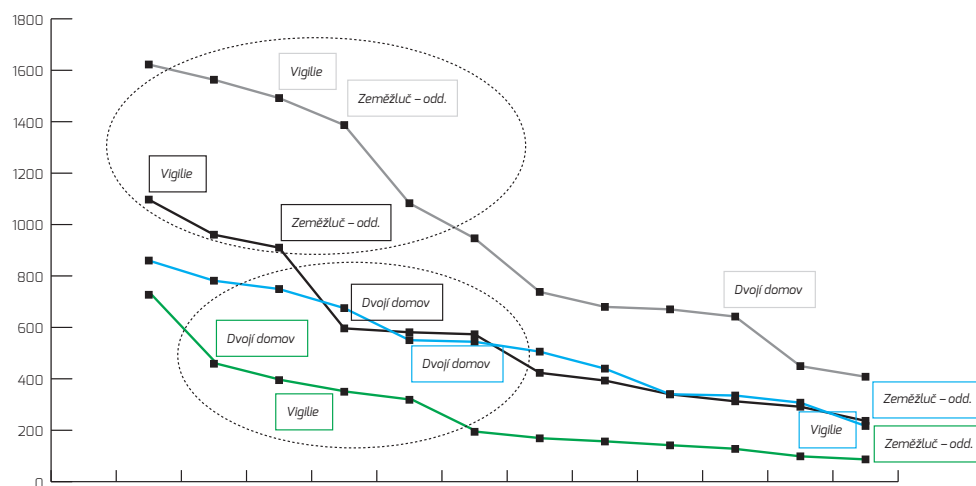
Tabulka 37b: Frekvenční distribuce lemmatu *zelený* v Čepově korpusu fikčních narativů (řazeno podle *Rf*).

	NÁZEV (SUB)KORPUSU	Af	Rf
1	časopisecké prózy 20. let	125	1622,50
2	časopisecké prózy 30. let	10	1563,20
3	<i>Vigilie</i> (1928)	34	1492,00
4	<i>Zeměžluč – oddíl</i> (1931)	32	1387,00
5	<i>Letnice</i> (1932)	34	1083,00
6	<i>Tvář pod pavučinou</i> (1941)	48	946,50
7	<i>Polní tráva</i> (1946)	47	738,10
8	<i>Cikáni</i> (1953+)	16	679,80
9	<i>Hranice stínu</i> (1935)	38	670,20
10	<i>Dvojí domov</i> (1926)	14	642,10
11	<i>Děravý plášť</i> (1934)	16	449,50
12	<i>Modrá a zlatá</i> (1938)	21	408,10

Tabulka 37c: Frekvenční distribuce lemmatu *bílý* v Čepově korpusu fikčních narativů (řazeno podle *Rf*).

	NÁZEV (SUB)KORPUSU	Af	Rf
1	<i>Vigilie</i> (1928)	25	1097,10
2	časopisecké prózy 20. let	74	960,50
3	<i>Zeměžluč – oddíl</i> (1931)	21	910,20
4	<i>Dvojí domov</i> (1926)	13	596,20
5	<i>Polní tráva</i> (1946)	37	581,10
6	<i>Letnice</i> (1932)	18	573,30
7	<i>Hranice stínu</i> (1935)	24	423,30
8	<i>Děravý plášť</i> (1934)	14	393,30
9	<i>Cikáni</i> (1953+)	8	339,90
10	časopisecké prózy 30. let	2	312,60
11	<i>Modrá a zlatá</i> (1938)	15	291,50
12	<i>Tvář pod pavučinou</i> (1941)	12	236,60

Tabulka 37d: Frekvenční distribuce lemmatu *černý* v Čepově korpusu fikčních narativů (řazeno podle *Rf*).

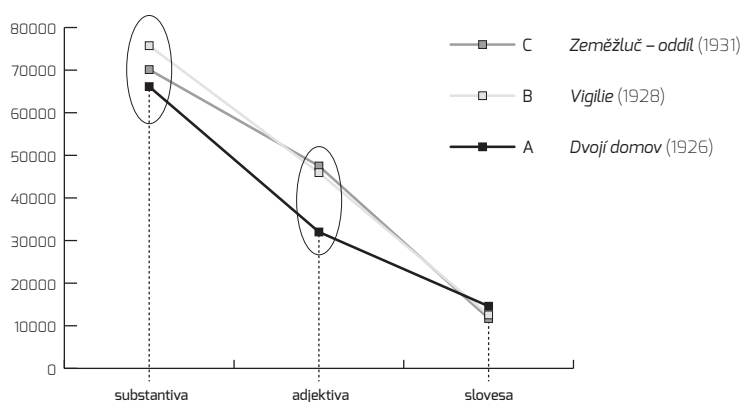


Graf 10: Průběhy relativních frekvencí vybraných adjektiv v korpusu fikčních narativů.

Srovnání frekvenčních hodnot (*Rf*) slovních druhů u tří sledovaných (sub)korpusů poskytuje následující tabulka (viz Tab. 38), ze které je patrné, že oproti (sub)korpusu *Dvojí domov* (1926) dochází u (sub)korpusů *Vigilie* (1928) a oddílu *Zeměžluč* (1931) k nárůstu bohatosti slovníku substantiv a adjektiv, zatímco v případě sloves je tomu naopak. Jak uvidíme posléze při srovnání (sub)korpusů *Dvojí domov* (1931), *Vigilie* (1931) a *Zeměžluč – oddíl* (1931), situace se změní (viz vyznačené oblasti v Grafu 11 a Grafu 14).

	substantiva	adjektiva	slovesa
<i>Dvojí domov</i> (1926)	66137,69	32013,94	14585,15
<i>Vigilie</i> (1928)	75741,62	45945,23	12550,47
<i>Zeměžluč – oddíl</i> (1931)	70128,29	47503,47	11702,50

Tabulka 38: Srovnávací frekvenční hodnoty (Rf) počtu různých substantiv, adjektiv a sloves u (sub)korpusů *Dvojí domov* (1926), *Vigilie* (1928) a *Zeměžluč – oddíl* (1931).



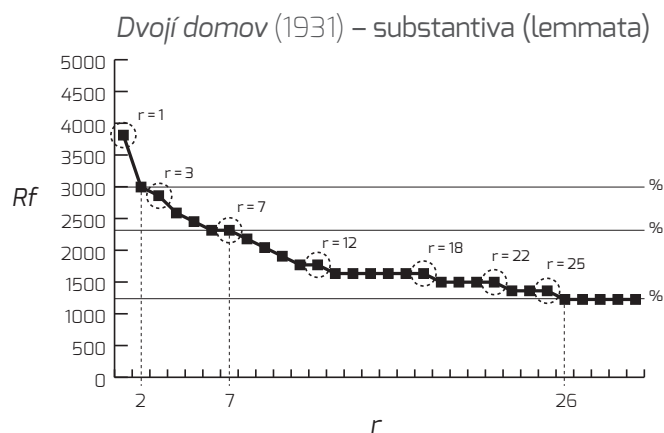
Graf 11: Průběh relativních frekvencí substantiv, adjektiv a sloves u (sub)korpusů *Dvojí domov* (1926), *Vigilie* (1928) a *Zeměžluč – oddíl* (1931).

Lze tedy konstatovat, že jestliže se na nejvyšších frekvenčních pozicích příliš nemění typ adjektiv a sloves s výjimkami, o kterých jsme hovořili výše, bývá potom bohatost slovníku jmenovaných slovních druhů výrazně jiná. (Sub)korpus *Vigilie* (1928) je z hlediska své celkové velikosti (N) větší o 4,52 % oproti (sub)korpusu *Dvojí domov* (1926). Počtem substantiv však činí tento rozdíl 14,52 %, u adjektiv 43,52 %. *Vigilie* jsou tak bohatší co do počtu různých substantiv a adjektiv. Naopak u sloves se přibližují hodnoty (sub)korpusů *Vigilie* (1928) a oddílu *Zeměžluč* (1931), zatímco (sub)korpus *Dvojí domov* (1926) má o 13,65 % bohatší lexikon sloves. Opět sledujeme rozdíl mezi prvním (sub)korpusem a následujícími dvěma (sub)korpusemi. Z takto získaných hodnot vyplývá hypotéza o možném nárůstu epichnosti v (sub)korpusech *Vigilie* (1928) a *Zeměžluč – oddíl* (1931), což by mohl signalizovat právě poměr vztahů mezi substantivy a adjektivy na straně jedné a slovesy na straně druhé, stejně jako srovnání mezi průběhy frekvenčních hodnot sledovaných slovních druhů u všech (sub)korpusů.

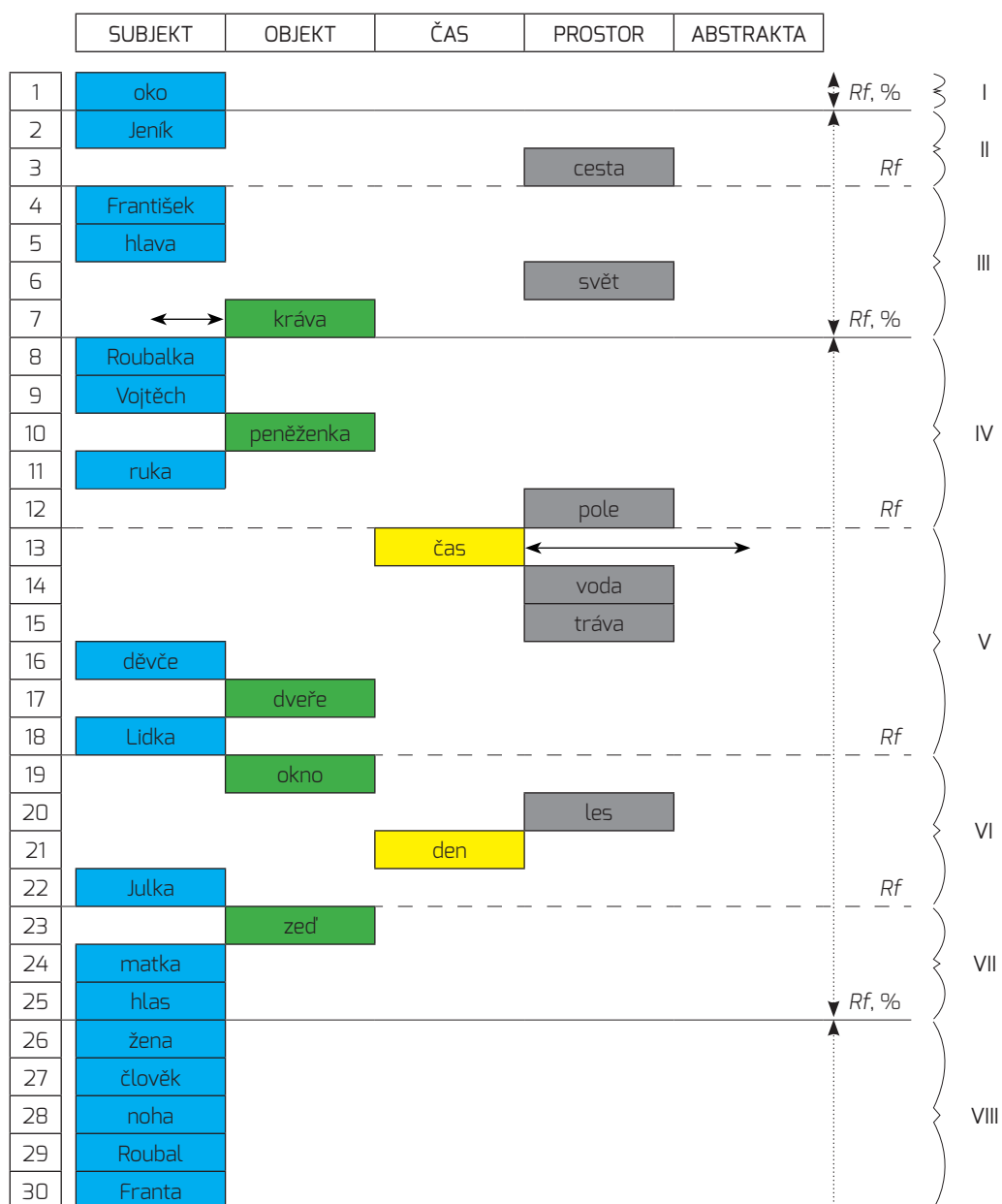
Dříve nežli tyto poznatky ověříme jiným statistickým parametrem, provedeme pro úplnost předcházející sled kvantitativně-korpusových analýz rovněž u (sub)korpusů *Dvojí domov* (1931) a *Vigilie* (1931).

<i>r</i>	lemma	<i>Af</i>	%	<i>Rf</i>
1	oko	28	1.7	3812,64
2	Jeník	22	1.3	2995,64
3	cesta	21	1.2	2859,48
4	František	19	1.1	2587,15
5	hlava	18	1.1	2450,98
6	svět	17	1.0	2314,81
7	kráva	17	1.0	2314,81
8	Roubalka	16	0.9	2178,65
9	Vojtěch	15	0.9	2042,48
10	peněženka	14	0.8	1906,32
11	ruka	13	0.8	1770,15
12	pole	13	0.8	1770,15
13	čas	12	0.7	1633,99
14	voda	12	0.7	1633,99
15	tráva	12	0.7	1633,99
16	děvče	12	0.7	1633,99
17	dveře	12	0.7	1633,99
18	Lidka	12	0.7	1633,99
19	okno	11	0.7	1497,82
20	les	11	0.7	1497,82
21	den	11	0.7	1497,82
22	Julka	11	0.7	1497,82
23	zed'	10	0.6	1361,66
24	matka	10	0.6	1361,66
25	hlas	10	0.6	1361,66
26	žena	9	0.5	1225,49
27	člověk	9	0.5	1225,49
28	noha	9	0.5	1225,49
29	Roubal	9	0.5	1225,49
30	Franta	9	0.5	1225,49

Tabulka 39: Prvních třicet substantiv (lemmat) v (sub)korpusu *Dvojí domov* (1931). Procento určuje zastoupení příslušného lemmatu ve třídě substantiv, hodnota relativní frekvence (*Rf*) je potom normalizovanou hodnotou vypočítanou vzhledem k velikosti (sub)korpusu ($N = 7344$).



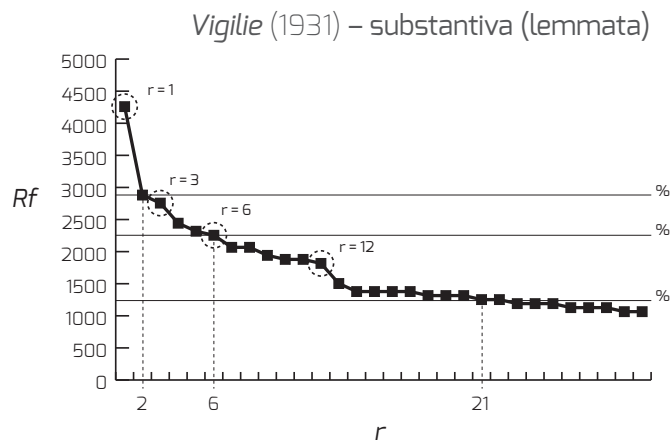
Graf 12: Frekvenční průběh prvních třiceti substantiv (lemmat) dle relativní frekvence v (sub)korpusu *Dvojí domov* (1931).



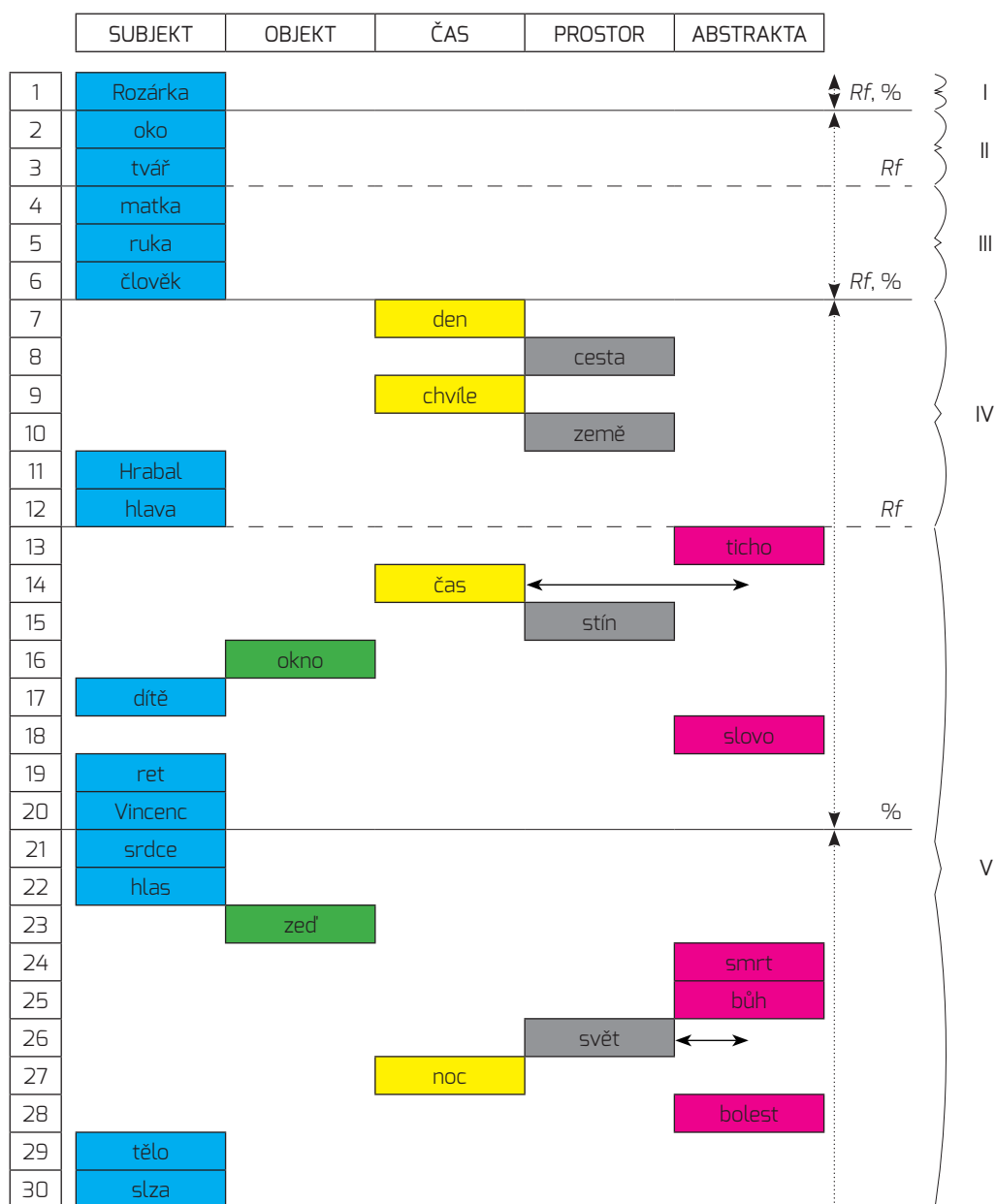
Tabulka 40: Distribuce substantivních lemmat do sémantických tříd v (sub)korpusu *Dvojí domov* (1931). Vodorovná plná čára vyznačuje procentuální distribuci v rámci kategorie substantiv, vodorovná přerušovaná čára je hranicí výraznějších frekvenčních přechodů (zlomů).

<i>r</i>	lemma	<i>Af</i>	%	<i>Rf</i>
1	Rozárka	68	1.8	4259,05
2	oko	46	1.2	2881,12
3	tvář	44	1.2	2755,86
4	matka	39	1.1	2442,69
5	ruka	37	1.0	2317,42
6	člověk	36	1.0	2254,79
7	den	33	0.9	2066,89
8	cesta	33	0.9	2066,89
9	chvíle	31	0.8	1941,63
10	země	30	0.8	1878,99
11	Hrabal	30	0.8	1878,99
12	hlava	29	0.8	1816,36
13	ticho	24	0.6	1503,19
14	čas	22	0.6	1377,93
15	stín	22	0.6	1377,93
16	okno	22	0.6	1377,93
17	dítě	22	0.6	1377,93
18	slovo	21	0.6	1315,30
19	ret	21	0.6	1315,30
20	Vincenc	21	0.6	1315,30
21	srdce	20	0.5	1252,66
22	hlas	20	0.5	1252,66
23	zed'	19	0.5	1190,03
24	smrt	19	0.5	1190,03
25	bůh	19	0.5	1190,03
26	svět	18	0.5	1127,40
27	noc	18	0.5	1127,40
28	bolest	18	0.5	1127,40
29	tělo	17	0.5	1064,76
30	slza	17	0.5	1064,76

Tabulka 41: Prvních třicet substantiv (lemmat) v (sub)korpusu *Vigilie* (1931). Procento určuje zastoupení příslušného lemmatu ve třídě substantiv, hodnota relativní frekvence (*Rf*) je potom normalizovanou hodnotou vypočítanou vzhledem k velikosti (sub)korpusu ($N = 15\,966$).



Graf 13: Frekvenční průběh prvních třiceti substantiv (lemmat) dle relativní frekvence v (sub)korpusu *Vigilie* (1931).



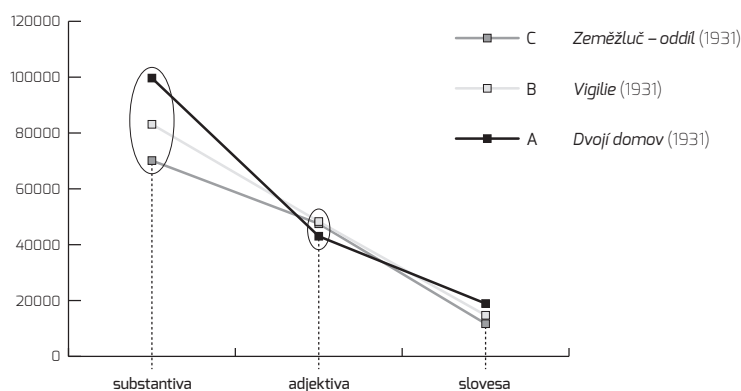
Tabulka 42: Distribuce substantivních lemmat do sémantických tříd v (sub)korpusu *Vigilie* (1931). Vodorovná plná čára vyznačuje procentuální distribuci v rámci kategorie substantiv, vodorovná přerušovaná čára je hranicí výraznějších frekvenčních přechodů (zlomů). Lemmata *čas* a *svět* tendují k sémantické třídě abstrakt.

(Sub)korpus *Dvojí domov* (1931) vykazuje tendenci, kterou sledujeme u analýzy (sub)korpusu *Vigilie* (1928), to znamená, že je oslabena frekvenční distinkce mezi prvními substantivy, zatímco druhá verze *Vigilií* (1931) ve skladbě a rozložení substantiv odkazuje k tendenci oddílu *Zeměžluč* (1931) s tím, že oproti prvnímu vydání, kde je frekvenční průběh substantiv plynulý, zaznamenáváme ve druhé verzi výrazný frekvenční skok mezi prvním a následujícím substantivním lemmatem. Současně i frekvenční struktura sémantických kategorií (viz Tab. 42) inklinuje k analogičnosti s oddílem *Zeměžluč* (1931). K této tendenci se připojuje i druhá verze *Dvojího domova* (1931), neboť frekvenčně zvýznamňuje kategorii prostoru, zatímco u *Vigilií* (1931) dochází ke zdůraznění kategorie subjektu. Srovnáme-li uvedené frekvenční stratifikace sémantických tříd mezi (sub)korpusy *Dvojí domov* (1931), *Vigilie* (1931) a *Zeměžluč – oddíl* (1931), sledujeme obecnou tendenci k vyrovnávání původních rozdílů. To se ovšem neděje jednoduše pod vlivem posledního ze souborů, kterým je oddíl *Zeměžluč* (1931), ale doslova napříč mezi všemi třemi oddíly. Zdá se, že druhá verze *Vigilií* (1931) je modifikována pod vlivem oddílu *Zeměžluč* (1931), zatímco druhá verze *Dvojího domova* (1931) zase pod působením původních *Vigilií* (1928). Tuto tendenci manifestuje nejen frekvenční stratifikace sémantických tříd (srov. Tab. 34, Tab. 40 a Tab. 42), ale též průběhy *Rf* tematických slov (viz Graf 6a, 6b), které u (sub)korpusu *Vigilie* (1931) mají tendenci k analogičnosti se (sub)korpusem *Zeměžluč – oddíl* (1931). V oblasti prvních tří Čepových povídkových souborů tak dochází k propojení, které neprobíhá lineárně (chronologicky), ale je ukázkou složitějších strukturních vazeb. Uvedené knižní soubory sice fakticky vznikaly v návazném časovém sledu, avšak finální soubor *Zeměžluč* (1931), který obsahuje upravené verze prvních dvou povídkových sbírek, je projevem vzájemného působení strukturních silových polí.

V následující tabulce a grafu (viz Tab. 43, Graf 14) zpřehledňujeme rozložení substantiv, adjektiv a sloves ve druhých verzích *Dvojího domova* (1931) a *Vigilií* (1931) ve srovnání s oddílem *Zeměžluč* (1931). Z celkových průběhů relativních frekvencí uvedených slovních druhů je zjevné, že může platit výše konstatovaná tendence k vyrovnávání některých původních strukturních rozdílů mezi jednotlivými (sub)korpusy (srov. Tab. 38 a Graf 11).

	substantiva	adjektiva	slovesa
<i>Dvojí domov</i> (1931)	99673,20	43028,32	18927,02
<i>Vigilie</i> (1931)	83114,12	48290,12	14718,78
<i>Zeměžluč – oddíl</i> (1931)	70128,29	47503,47	11702,50

Tabulka 43: Srovnávací hodnoty (*Rf*) počtu různých substantiv, adjektiv a sloves u (sub)korpusů *Dvojí domov* (1931), *Vigilie* (1931) a oddílu *Zeměžluč* (1931).



Graf 14: Průběh relativních frekvencí vybraných slovních druhů ve sledovaných (sub)korpusech.

Radikálním zásahem do první verze *Dvojího domova* (1926) dochází zejména k celkovému posílení kategorie substantiv. Kategorie adjektiv rovněž posiluje, ovšem na rozdíl od substantiv směřuje spíše k vyrovnávání původních rozdílů. Bez zajímavosti nejsou ani hodnoty tzv. aktivity a deskriptivity textu (hodnota *Q*), se kterými pracují autoři publikace *Metody kvantitativní analýzy (nejen) básnických textů* (2014).²⁰⁵

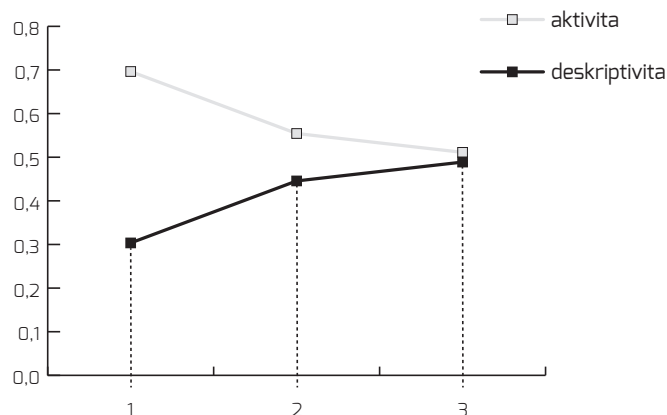
		entropie	deskriptivita	aktivita
1	<i>Dvojí domov</i> (1926)	9,646012	0,303522	0,696478
2	<i>Vigilie</i> (1928)	9,777621	0,445625	0,554375
3	<i>Zeměžluč – oddíl</i> (1931)	9,690760	0,488854	0,511146

1	<i>Dvojí domov</i> (1931)	9,300418	0,313289	0,686711
2	<i>Vigilie</i> (1931)	9,588441	0,433894	0,566106
3	<i>Zeměžluč – oddíl</i> (1931)	9,953811	0,443379	0,556621

Tabulka 44: Hodnoty entropie, deskriptivity a aktivity měřených (sub)korpusů.

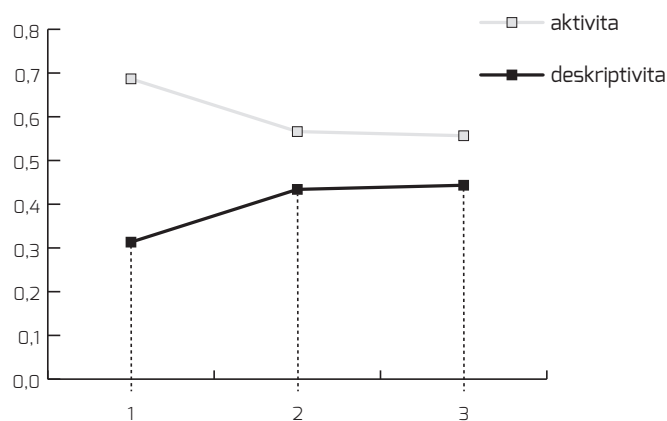
²⁰⁵ Viz Čech – Popescu – Altmann 2014, s. 52–73. Pro výpočet hodnot jsme použili software *Quantitative Index Text Analyser* (QUITA), který je volně přístupný na adrese WWW: <<https://code.google.com/p/oltk/>>.

Dvojí domov (1926) – *Vigilie* (1928) – *Zeměžluč* – oddíl (1931)



Graf 15a: Vztah deskriptivity a aktivity u jednotlivých (sub)korpusů.

Dvojí domov (1931) – *Vigilie* (1931) – *Zeměžluč* – oddíl (1931)



Graf 15b: Vztah deskriptivity a aktivity u jednotlivých (sub)korpusů.

Vztahy mezi kategoriemi deskriptivity a aktivity u prvního vydání *Dvojího domova*, *Vigilií* a oddílu *Zeměžluč* ukazují jednoznačnou tendenci ke snižování aktivity a zvyšování deskriptivity s tím, že výraznější předěl sledujeme opět mezi *Dvojím domovem* (1926) a *Vigiliemi* (1928). V oddíle *Zeměžluč* jsou hodnoty vyrovnanější. Naopak mezi částmi celku knihy *Zeměžluč*, tzn. mezi druhými verzemi *Dvojího domova* a *Vigilií* spolu s oddílem *Zeměžluč* (viz Graf 15b) pozorujeme tendenci k vyrovnávání rozdílů měřených hodnot.

Vezmeme-li v úvahu charakter veškerých provedených analýz spolu s dosaženými výsledky, lze konstatovat společný jev, který je možné označit za modifikaci centrální invariantní tematické struktury, jejímž hlavním organizujícím principem se stává

synekdochičnost. Na pozadí této invariantnosti se ukazují jednak vývojové tendence (viz distribuce a vztahy sémantických kategorií třídy substantiv) směřující od *Dvojího domova* (1926) k oddílu *Zeměžluč* (1931), jednak návratné tendence spočívající v modifikaci druhé verze *Dvojího domova* (1931) a *Vigilií* (1931). Synekdochičnost se projevuje rovněž v kontextových vazbách (kolokacích), které od *Vigilií* směrem k oddílu *Zeměžluč* (1931) realizují propojenější (manifestovanější) vazby mezi základními sémantickými kategoriemi substantiv, čímž posilují základní vztahový tematický model „subjekt–čas–prostor“, resp. činí jej explicitnějším. Synekdochičnost může také vyjadřovat celkový vztah mezi jednotlivými (sub)korpusy – částmi celku knihy *Zeměžluč*.

Otázkou samozřejmě zůstává, jaké další strukturní síly způsobují tyto modifikace. Zcela jistě je nelze vykládat pouze jako prosté rozhodnutí autora, neboť výsledky našich analýz jednoznačně ukazují na působení společného invariantního základu. Působí zde tedy dva základní aspekty, které jsou komplementární: textová pole a autorská intence neboli invariantní strukturní rádius propojující se s autorskou intencí. Tím se kategorie autora zapojuje do sítě strukturních vztahů prvních tří knih a vlastním způsobem, tj. jako textově vnější element, avšak v dosahu strukturní „radiace“ textového pole uměleckých próz, se podílí na organizaci (sestavení) výsledné podoby knihy *Zeměžluč*.

3.5 Literárněvědné aspekty (vypravěč, tematika, motivika, čas a prostor)

Jak konstatoval Tomáš Kubíček, povídky *Dvojího domova* (1926) jsou výrazně organizovány personickým hlediskem, které je zprostředkováváno formou narativní výpovědi. Fikční svět, který je konfigurován touto narativní strategií, se na jedné straně jeví jako projekce personického modu, který přirozeně disponuje určitým kognitivním rozsahem, okruhem toho, co lze smyslově vnímat a pozorovat, na straně druhé právě díky formálně-narativní zprostředkovanosti daného hlediska se stává jakýmsi axiologickým úběžníkem fikčních světů povídek *Dvojího domova* (1926), jejich validitou. Zkušenost postav osvojovaná v konfrontaci s okolním světem, zejména pak základní zkušenost dětí s poprvé si uvědomovanou rozlehlostí a nezměrností okolního světa, není díky realizaci polopřímých řečí a vnitřních monologů s minimálním využitím přímých řečí pouze nahodilým projevem těkavého stavu postavy, ale hodnotou – existenciální zkušeností, kterou je naplněn a doslova prostoupen fikční svět, resp. fikční světy povídek. Centrem *Dvojího domova* se stává narativní výpověď o fikčním světě, jehož celková povaha je významně determinována personickým modelem. Subjekt skrze své kognitivní schopnosti, jež jsou velmi často transformované do narativního pásma, filtruje okolní svět.

Důležitou úlohu zde mají výrazy spojené s okruhem jednání subjektu-postavy. Kvantitativně-korpusová analýza nám ukázala systém jejich odstupňování. Ve *Dvojím domově* dominuje lemma *oko*, které nerealizuje pouhé nociónální významy, ale nabývá i expresivního rozměru. Korelační lexém *dívat se* spojený se sémantickým polem lemmatu *oko* centralizuje významy ve fikčních světech povídek *Dvojího domova*, které se úzce napojují na vlastní akt vnímání a uvědomování si.²⁰⁶ Tematický význam dle korpusové analýzy mají konkrétně označující části lidského těla, kategorie rodinných příslušníků (otec, matka), vlastní jména nebo základní apelativa, která se ocitají v centru tematického plánu a často se stávají katalyzátorem dění, jež je procesem uvědomování si změny situace, náhlého pochopení a prozření. Právě zacílení na tuto kategorii konkrétně vytváří funkční a významové impulzy v těchto prózách.

p. Bouře (*Dvojí domov*):

Jeník vidí zdola otcovu postavu proti černému nebi, ani kabátu nemá na sobě, matka hýbe potichu rty a kravky kývají hlavami, přimhuřující oči, kdykoliv zahřmí. (Čep 1926, s. 12)

Otec stojí nahoře bez klobouku a po ruce mu teče krev, maminka chudák v modré sukni se napíná a nemůže dosáhnout, zrní a vousy jí padají do očí, do úst, za košili, pot jí teče po tváři a nemá kdy si jej utřít, černá prohlubeň visí už nad samou hlavou a praští to v ní, až nohy podklesávají, Bože, ještě jeden mandel! (Ibidem, s. 13)

Jeník vzhlédl, otcova bílá tvář zůstala přibita k mračnům, strašná svítivá hloubka se za ním rozevřela a krávy sražené děsným úderem spadly na kolena a otec se svezl po fůře k zemi, zatímco matka se chytila klanice, bleďá jako smrt. Jeník plakal a spíнал ruce, ale nemohl si vzpomenout na žádnou modlitbu. (Ibidem, s. 13)

p. Smrt ševce Nerušila (*Dvojí domov*):

Švec dlouho neusíná a ráno ho mučí stejná nejistota. Z protějščí zídky naň civí kamenné tváře a vyhrožují mu: „Umřeš!“ Švec prolelkuje den a procitne-li, marně se pokouší zapudit hlas prudkými ranami kladívka. Brzo se unaví a tváře civí dál. (Ibidem, s. 17)

Ševcovo okno dívalo se do kamenné zídky. Zídka byla neomítnutá, ne však jenom kámen na kameni, nýbrž kamení časem a dlouhým hleděním z okna proměněné v hromadu na sobě ležících tváří, obrostlých mechem. Vždycky to bylo stejné: kdykoliv se švec trochu déle zadíval, z každého kamene naň zírala tvář. Mezi oknem a zídkou cesta lezla po břiše: dvě koleje hluboko vryté a znovu vytepané okutými koly. (Ibidem, s. 15)

²⁰⁶ S tím velice úzce souvisí i další frekventované lemma *okno*, které se dostává do korelačního spojení s lemmatem *oko* (srov. např. povídku *Smrt ševce Nerušila*). Tohoto aspektu si přirozeně všimá i Kubiček ve své analýze povídek *Dvojího domova*: „Okna se u Čepových stavení dívají, pozorují člověka a stávají se pohledem ne-přítomných příslušníků rodu. Tato silná antropomorfní figura se často vyskytuje i ve fikčních světech Čepových próz a Čep nám v tuto chvíli nabízí její funkční význam.“ (Kubiček 2014, s. 68)

Ačkoli se jedná o *er-formu*, hledisko vypravěče je transpozicí hlediska postavy. To se projevuje nejvýrazněji v přirovnání hromady kamení k lidským tvářím, což je výsledkem subjektivní projekce postavy. Jiný aspekt, který propojuje *nocionální entity* fikčního světa se subjektem (postavou) a vytváří mezi nimi podmíněný vztah, je aspekt kompoziční. Poslední citovaná ukázka je incipitem povídky *Smrt ševce Nerušila*. Pro většinu povídek *Dvojího domova* platí, že je v nich výrazně oslaben epický potenciál. Okno ve vyprávěcí intenci není pouze obyčejným objektem, částí domu, ale rovněž jakýmsi zástupným symbolem, jenž navazuje metaforicko-synekdochický vztah se subjektem (*okno se dívá*). Jednotlivé základní elementy, ze kterých je konstruován fikční svět, se zde spojují společně se způsobem narativního projevu v úzce funkčně provázané významové celky, které určuje nejen typ vztahu, ale i kvalita vztahujících se částí. Tím je budována již zmíněná struktura (metaforicko-)synekdochického principu, jehož bazálními konstituenty jsou konkrétní disponující elementární věcnou referencí.

Ne vždy se však musí jednat o transpozici hlediska postavy do narativní roviny. Přestože by tomu mnohé odpovídalo, v následující ukázce je na místě zvažovat spíše funkčně-sémantickou konfuzi mezi objektivním hlediskem *narátora* a hlediskem postavy.

p. *Smrt ševce Nerušila (Dvojí domov)*:

Okno je otevřeno, a když se stmívá, obecní sluha napájí býky, kteří zdlouha bučí a tlukou řetězy o koryta. Ševcova žena hekajíc skládá na záspi nůši, chrastí a kouty světnice, už už se pohružující do stínů, rychle se probírají z dřímoty. Švec prudce zamává kladívkem. Chalupa rázem ožije, hlas ševcové řinčí na dvoře a kolem plotny a švec zoufalými údery přibíjí výčitky příliš dotěrné. (Ibidem, s. 15)

Jedná se o deskripci prostředí, která nutně nepředpokládá transformované hledisko postav tak, jako tomu bylo v případě přirovnání kamenů k lidským tvářím, ale dovoluje chápat narativní klasifikaci jednotlivých komponentů deskripce (*zdlouha bučící býci, hekající žena, světnice pohroužena do stínů* apod.) jako vlastní vypravěčský (latentně interpretující) záměr, který má rámovat a náležitě evokovat ševcovu životní situaci (*švec zoufalými údery přibíjí výčitky příliš dotěrné*).

Podobně jako zde i na jiných místech dochází na základě hyponymně-hyperonymních poměrů a vypravěčovy garance k realizaci vztahu mezi subjektem a transcendentně projektovaným prostorem. Na pozadí této valence je latentně evokován vztah člověka k nezměrnosti okolního světa, jež však překračuje rozlehlost topografickou směrem k symbolické rovině. Čepovy postavy se naopak pohybují na ploše poměrně úzkého radiusu, který je vymezen dosahem jejich činností (chůzí, rodinnými vztahy apod.). Nezměrnost je tedy orientována nikoli horizontálně jako nezměrnost topografická, ale v prvé řadě vertikálně jako nezměrnost zakoušená a prožívaná subjektem; její přirozenou

součástí se ve světě Čepových próz *Dvojího domova* stává pocíťovaná úzkost, náhlé prozření a pochopení nesamozřejmosti prosté reálné situovanosti člověka do konkrétního prostředí a prostoru. Symbolický význam, který se projevuje ve vztahu člověka a jeho situovanosti v prostoru světa, je dán součinností vypravěče, jeho personického modu spolu s organizací lexikální vrstvy na ose hyponymně-hyperonymních vazeb. Spolu s tematikou vytvářejí tyto základní komponenty strukturně-sémantické paradigma, jehož výrazem je synekdocha realizující se jako vztah člověka k času a prostoru.

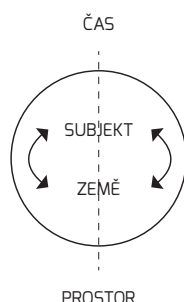


Schéma 19: Vertikalita jako synekdochický princip.

Koncept této vertikality je realizován skrze výraznou tematizaci kategorie subjektu, zdůrazněním jeho modálního rádiusu (*být v dosahu subjektu*), potlačením epičnosti, s čímž souvisí kompoziční postup přímého uvedení do děje, a personickým hlediskem vypravěče. Význam prostoru jako ve výsledku transcendentální kategorie není explicitně formulován, ale skrze uvedené umělecké postupy a techniky sugerován. Je rovněž zajímavé, že propojení tohoto významu s křesťanskou symbolikou není tak explicitní, jak by se u Čepa dalo očekávat.

lemma	<i>Af</i>	<i>Rf</i>
Bůh	22	1009,04
kříž	17	779,71
modlitba	4	183,46
hřích	3	137,60
víra	1	45,87
odpuštění	1	45,87
křesťanský	0	0,00
vzkříšení	0	0,00

Tabulka 45: Výběr lemmat pojících se s náboženskou tematikou v (sub)korpusu *Dvojí domov* (1926).

V (sub)korpusu *Vigilie* (1928) korpusové hodnoty signalizují určitou proměnu v dosavadní koncepci a strukturaci základních vztahů mezi kategoriemi subjekt, čas

a prostor. Tuto proměnu jistě nelze označit za radikální, spíše lze hovořit o významovém posunu a modifikaci invariantního paradigmatu tematického a časoprostorového rámce. Jak jsme viděli při analýze tematických slov, oslabuje se distinkce mezi významně (frekvenčně) akcentovými lemmaty a dochází k zapojení širšího spektra sémantických tříd (čas, prostor a postupně i abstrakta). Reorganizuje se i rozložení sémantických tříd (viz Tab. 29) a jejich kontextové okolí (viz Schéma 17). Další krok v tomto směru představuje oddíl *Zeměžluč* (1931), kde výrazněji nabývají na významu abstrakta. Pro všechny (sub)korpusy samozřejmě platí, že jejich centrální osou je sémantická kategorie subjektu. Všechny sledované (sub)korpusy vykazují základní strukturu vztahů subjektu k ostatním sémantickým kategoriím (subjekt–objekt–čas–prostor–abstrakta), s tím, že strukturace každé z dílčích kategorií je založena na principu hyperonymně-hyponymních vztahů, což směřuje k předpokladu chápat tyto vazby nejen uvnitř sémantických tříd, ale i mezi sémantickými třídami navzájem. V každém ze (sub)korpusů je ovšem tato struktura realizována specificky, což se odráží v celkovém významovém plánu. V prvním vydání *Dvojího domova* (1926) je ústředním momentem personické hledisko er-formového vypravěče, které modeluje fikční svět jako na jedné straně koncepci subjektivní projekce (modální složka), na straně druhé formální složka vyprávěcí techniky (er-forma) zajišťuje věcný rámec fikčního světa v jeho intersubjektivní platnosti. Synekdocha je realizována na ose subjektivního vnímání a zakoušení tušeného, avšak skrytého tajemství řádu věcí, které přesahuje racionální poznávací schopnosti a směřuje k transcendentálnímu rozměru, který však není explicitně pojmenován, ale umělecky sugerován na základě uvedené strategie. Právě subjekt a jeho akt vnímání (reflexe) a prožívání je centrem tohoto významového dění, jemuž se podřizuje i strategie vypravěče. Na straně druhé konkrétní zkušenosti přesahují postavu a nabývají (či dávají tušit) platnost univerzální. Malý svět zabydlený postavou (vesnice, pole, dětský domek), relativně úzký modální rádius subjektu (jeho jednání a dosah jeho pozorovacích schopností) vstupuje do korelačního vztahu s nezměrností času a prostoru, které v rámci synekdochy implikují vyšší axiologické (nadsmyslové) hodnoty propojující se se smyslovým světem postav; ve výsledku však díky er-formě překračují čisté subjektivní rámec a přerůstají do oblasti nadindividuální. Tento význam není ve *Dvojím domově* explikován, ale evokován.

(Sub)korpus *Vigilie* (1928) signalizuje, že dochází k určité změně v této koncepci. Na rovině vyprávění se sice pracuje s personickým hlediskem, avšak není tak výrazně kondenzováno jako v povídkách *Dvojího domova*. Proměnou prošel zejména ústřední motiv cesty, jehož sémantický okruh je ve *Dvojím domově* (1926) spojen především s faktickým pohybem postav (*Domek, Bouře, Do města*). Teprve ve druhém plánu se motiv cesty realizuje v symbolické rovině. Nicméně tento význam zde není natolik explicitní jako ve *Vigiliích* a posléze v *Zeměžluči*. Právě tuto skutečnost signalizuje na úrovni

(sub)korpusu *Vigilie* (1928) nápadně oslabená frekvence lemmatu *jít* a významová potencialita lemmatu *cesta*, které se dostává do oblasti nad *h-bodem*. Kromě toho dochází k tomu, co jsme na jiném místě označili jako „zředění“ původně výrazně tematizované kategorie subjektu. Projevuje se to tím, že se do exponované oblasti nad *h-bodem* dostávají i lemmata reprezentující jiné sémantické kategorie. Proměňují se i poměry mezi sémantickými třídami subjekt, čas a prostor. Mění se rovněž strategie vypravěče. Jestliže prostor a čas byl ve *Dvojím domově* úzce spojen s prožíváním subjektu, ve *Vigiliích* nabývají tyto kategorie výrazně lineárního a topografického rozměru. S tím souvisí i větší tendence k epičnosti (analepsím) a vypravěčské fabulaci.

p. Rozárka Lukášová (*Vigilie*):

Po silnicích však zatím hřmotily automobily a sedláci v hospodě se obírali nejnovější politikou, čtvero ročních počasí se několikrát vystřídalo a Rozárčino bledé panenství stálo uprostřed rudého plání léta a divokého stesku vesnických muzik s neobyčejně bolestným úlekem v modrých očích, prohlubujících se až k pomezí jiných světů. Cesty tajemně přikované právě k tomuto okrsku země marně nalhávaly svou linií, ubíhající k dálkám, že je možno čemusi uniknout. [...] Houfy poutníků prodlením času prořídly a procesí, beroucí se s křížkem, s odkrytými hlavami a s uzlíky na zádech po zaprášených cestách, mimo vesnice a Boží muka uprostřed polí, stalo se znamením nevkusy a terčem hanby. Dávno minula doba, kdy lidé v úzkosti srdce, v zármutku, pro něž není pozemské útěchy, v touze po nevýslovném oddání se spínali ruce k nebi a odtrhující se na čas od rodného kraje, od rodiny a přátel, od denních starostí, dávali se proniknout pocitem věčného míjení, v němž všechny věci dostávají novou drahou tvář, a svěřovali se docela do moci Bohu, nastavující mu svá okoralá srdce, aby je zkravil láskou k sobě. (Čep 1928, s. 34, 36)

V souboru *Dvojí domov* se postavy skutečně pohybují po několika málo trajektoriích, které vymezují rozsah fikčního světa. Tematizace pohybu člověka, jeho chůze v intencích jeho modalit (*muset jít, rovnou jít*) spolu se základní prostorovou orientací je ve *Dvojím domově* nositelkou skrytého významu dočasnosti postav v reálném časoprostoru, na jehož místě se objevuje (postavou) tušený a umělecky evokovaný časoprostor Čepova jiného transcendentálního domova. Motiv chůze, cesty vázaný na mnohdy expresivní projevy subjektu (postavy) je motivem, který významně spoluvytváří téma neukotvenosti člověka v reálném světě; pohyb člověka je projevem jeho přirozené aktivity a je současně spojen s momentem přechodu k náhlému zjištění, které způsobuje existenciální znejistění. V povídce *Domek* se uvedené skutečnosti pojí s postavou malého Jeníka.

p. Domek (*Dvojí domov*):

Jeník zatím, proměniv se, procházel se mezi stébly trav, jež rostly alejemi mezi lavičkami a stolem, lysá pěšinka vcházela sem jako do dvora a přívětivě zvala, jako zvou cesty venku, jenže nikdo po ní nemohl jít, prostor byl uzavřen jako svět, větvičky s měkkým listím visely až k zemi a zelené ticho spouštělo tu křídla jako snící pták. Ale znenadání si ticho postesкло,

bůhví proč. Právě v té chvíli, když si Jeník pomyslel, jak dobře by tu bylo zůstat a nikam odtud nechodit, tesknota si usedla do listů a domeček si zívnu.

V Jeníkovi se něco sevřelo, div mu slzy nevstoupily do očí. Vždyť to nejde, vždyť venku je svět a o tom by pak nic nevědě! Už chtěl rozhrnout větvičky a stoupnout na hlinu, ale ještě jednou si prohlédl „domek“, aleje travných stébel a pěšinku pro brouky, již on nemůže jít, malý svět musí zůstat opuštěný a marně krásný, nikdo se o něm nedozví... (Čep 1926, s. 8–9)

Jinde, jako v povídce *Do města*, je s chůzí (cestou) hlavní postavy do města a zpět spojen prožitek transcendentální úzkosti, jejímž výsledkem je smrt hlavní postavy. Jako základní životní zkušenost se motiv konkrétní (dílcí) cesty objevuje také v povídce *Bouře*, ve které otec, matka a syn vzdorují silnému přírodnímu živlu. V předchozích kapitolách jsme viděli, že významným prvkem v (sub)korpusu je motiv cesty. Ve *Dvojím domově* (1926) je korpusově zastoupen silně frekventovaným lemmatem *jít*, zatímco ve *Vigiliích* (1928) rovněž lemmatem *cesta*. Sémantika postav i fikčního světa se u Čepa zásadním způsobem realizuje skrze spojení postavy a motivu cesty, který je ovšem ve *Dvojím domově* (1926) řešen přece jen poněkud odlišně než ve *Vigiliích* (1928). Pohyb postavy je v první povídkové knize těsně propojen s postavou, která velice často vykonává určitý reálný pohyb prostorem; jedná se obvykle o chůzi. Motiv cesty je v intenci personického modu, hlediska postavy, nepřekračuje ji, což platí rovněž o vyprávěcí strategii.

Ve *Vigiliích* (1928) se motiv cesty modifikuje z významu aktuálně realizované cesty, momentálního pohybu postavy ve fikčním světě, do symbolického významu cesty jako životní pouti. Proměňuje se – jak bylo řečeno – i vyprávěčská strategie, která je na rozdíl od *Dvojího domova* (1926) schopna generovat větší míru epičnosti, čemuž v korpusu odpovídá zvýrazněný akcent sémantické kategorie prostoru (reprezentované zejména lemmatem *cesta*). Příznačný je z tohoto hlediska již vstupní text souboru *Vigilie* (1928), jakýsi makrokompoziční incipit celé knihy, který v prvním vydání (ve druhé verzi Čep tento text vypustil) nese název *Na rozchodnou*.

p. Na rozchodnou (Vigilie):

Vzduch vymáchaný letními bouřemi pomalu vychládá, čekanka a čapí nůsek vedle cesty kvetou melancholicky a švestkové stromy na travnatých březích těžknou už modrajícími se chumáči plodů hebců ojíněných.

Nuž, vydejme se s holemi v rukou, s kapsou sotva tuze obtíženou, ale s čerstvým vánkem kolem čela, zanechme roviny a jejích cest příliš střízlivých, pustnoucích lánů hnědě přeora-
ných, střídajících se s širou a únavnou zelení cukrové řepy, zanechme této rozsáhlé roviny, jež nás zrodila a jež hlídá naše prázdninové mravy okem tak nepřejím, a obraťme své kroky k zeleným pahorkům, skrývajícím mezi svahy rozmilé kroučící se pěšiny, šumot potoků průhledných v chladné trávě a hladkém kamení, hloučky střech, kupících se vzhůru do strání, zvoucích hospody na zatáčkách bílých silnic, tam, kde se cesty rozbíhají všemi směry a kde stojí pod dubem na kraji lesa železný křížek na památku hajného Josefa Nováka, kterého tu zabili

pytláci L. P. 1841. Nuž, vyjděme rychle ze dvora a nedopřejme času svému svědomí, aby nám nevhod připomnělo tíhu úkolů, jež jsme nechali doma nedotčeny, a místo po návsi jděme raději pěšinou za humny, neboť pomluva, šelestící za pootevřenými dveřmi, mohla by otrávit naši dočasnou svobodu hned v začátku! Jen co mineme pastoušku s vůní suchého roští a chalupu na konci s bílým koněm na vratech, jen co vystoupíme k osamělé borovici na svahu pod sečemi, hle, už se vine v modru bílá silnice v boku krotkého kopce a dál pod zelenem stromů třepetá už křídly první hospůdka jako bílá holubice! Leč chvíle ta je ještě příliš daleko a ty, Pindo, hospodo moje na světě nejmilejší, mi odpustíš, když se tu teď zastavíme na jedinou skleničku slivovice! Odtud pak už jen poslední pohled do roviny, na jejímž okraji leží moje dědina, schoulená v svém stromoví tak stydlivě a střílející do mého srdce tenký šíp mateřsky něžné, ale přece jen nezvykle bolestné výčitky. Oh, zrádný otvore, mohl by ses stát průlomem, kterým by se provalila všecka má skrytá bolest a všecko mé zoufalství, všecka hrůza lidského loučení a všecek žal, hrdlo se mi svírá – však ne, nejsem sám a ti druzí by uhádli, že nejde jen o maličkost, jen o jednodenní výlet. Nuže, zatočme kloboukem a zazpívejme si! Před námi modrý vzduch a bílá silnice, útlá a čistá silnice horského kraje, a přes ni kladou se s obou stran vlhké stíny... En avant! (Čep 1928, s. 9–10)

Text je zajímavý tím, že jej nelze jednoznačně žánrově specifikovat. Osciluje mezi vzpomínkou, reportáží, esejem a povídkou, avšak na rozdíl od ostatních textů *Vigilií* výrazně postrádá dějovost. Hlas vypravěče v podstatě nevypráví, ale sugeruje vzpomínku na rodný kraj (podobně i v povídce *Vigilie*) a postavy s ním spojené. Popisuje a tematizuje se zde čas a prostor jako uplývající hodnoty zanikajícího světa.

p. Na rozchodnou (*Vigilie*):

Je nás sedm, kteří zpíváme písničku. Je nás sedm, kteří jsme byli ještě před chvílí dětmi – můžeš číst tu historii v nalévajících se svalech, ve tvářích tak proměnlivých, že se jediným škubnutím, jedinou vzpomínkou ještě vrátí o deset let nazpět. Před chvílí ještě sen a dobrodružství, důvěrné obcování s druhým světem, který se už už navždy uzavírá – konec šťastným nesouvislostem, tu je vážný věk, přísahá se na zděděné bludy, babylon se začíná. Ale co si počnu já, který obcházím s utajeným dechem kol věcí, na jejichž místě vidí ostatní jen vzduch a prázdny prostor? (Ibidem, s. 13)

Text otevírající soubor *Vigilie* odkazuje ke způsobu, jakým bude s tématem času a prostoru zacházeno. Prostor a čas se na rozdíl od *Dvojího domova* stávají mnohem výrazněji objektivizovanými kategoriemi, které demonstrují ulpívání prostoru a míjení času. Na tomto základním strukturním pozadí se realizuje konkrétní prostředí (fikční topografie) ve významu palimpsestu²⁰⁷ – překrývajících se časoprostorů. Ve *Vigiliích* se ústředním tématem stává vztah k předchozím generacím, vztah člověka (postavy) k místu, k jeho objektivní historii, a zejména k historii individuální, rodové. S tím souvisí

207 „Oh, kdybyste uměli číst spodní písmo palimpsestu! Leč tak je líp, byli byste pak nenávratně ztraceni pro své kariéry.“ (Ibidem, s. 16)

i zmíněný motiv cesty, který se vedle významu, jenž spočívá v překročení hranice mezi jevovým a nadsmyslovým světem (*Zbloudilý*), v náhlém poznání dvojdomosti zdánlivě důvěrného a známého prostoru (*Husopas*, *Vigilie*), realizuje rovněž jako symbolický motiv pouti životem (*Rozárka Lukášová*, *Vigilie*²⁰⁸).

Jestliže ve *Vigiliích* dochází převážně k explicitní tematizaci časoprostoru, například k jeho narativní anticipaci, a tím *de facto* k exemplifikaci,²⁰⁹ ve *Dvojím domově* (1926) převládá spíše jeho implicitní tematizace, a to zejména pomocí narativní strategie, která se týká personického hlediska.

p. *Dvojí domov* (*Dvojí domov*):

Měsíc svítíl a bylo jako ve dne. Roubalka dlouho nemohla usnout, viděla chalupu v horách pod lesem, jeden došek trčí ze střechy a měsíc se jím na chvilku zastíní. Ale cosi navěky odešlo, otec i matka jsou mrtvi, jinde je skutečnost... (Čep 1926, s. 34)

p. *Kozlovice* (*Dvojí domov*):

Obě děvčata sedí pod dubem, dědina je dole zabořena v zahradách, muziku slyšet až sem a oběma je teskno. (Ibidem, s. 36)

Hospoda stojí ve tmě a do tančírny se jde přes dvůr a musí se dát pozor, neboť náspa jsou úzká a hnojiště je hned přede dveřmi. Tančírna je těsná sklenice, do níž bys nacpal chroustů a pak tím silně zatřepal. Jak se to hemží, bzučí a leze po sobě! (Ibidem, s. 36)

p. *Smrt ševce Nerušila* (*Dvojí domov*):

Okno je otevřeno, a když se stmívá, obecní sluha napájí býky, kteří zdlouha bučí a tlukou řetězy o koryta. Ševcova žena hekajíc skládá na záspi nůši chrastí a kouty světnice, už už se pohružující do stínů, rychle se probírají z dřimoty. (Ibidem, s. 15)

V uvedených případech sledujeme, jak vypravěč přejímá hledisko postavy. Prostorové relace jsou modifikovány pod jejím zorným úhlem. V následujících dvou incipitech, první je z povídky *Kozlovice* ze souboru *Dvojí domov* (1926), druhý z povídky *Zbloudilý*, která je součástí *Vigilií* (1928), je prostřednictvím narativního hlasu evokován rámcový časoprostor ve smyslu topografie fikčního světa povídky.

208 „Cesty jsem si zamiloval pro ně samy, pro jejich poutnickou podobu, pro jejich pokoru a trpělivost. Co lidí po nich přejde, co šlépějí je poznamená, oh, kdo by chtěl luštit to písmo, vyryté do cest šlépějemi lidských nohou! Jaký obraz lidského bloudění, tápání slepců a posedlých, a cíl je mýj, cíl je mýj [...] Vy všichni, kteří jste mne předešli, zde stojím a bojím se odclonit oči [...] Nevím, jak dlouho se ještě budu potloukat po vyprahlých cestách, rozdrčených pod lomoživou tíhou lidských břemen. Ale neslyším toho hluku, mé ucho je napjato do ticha pod břeskňými hlasy, a když stanu za šera na křižovatce ve stínu lesního koutu, naslouchám bázlívě, chvějí se a přece toužím, a zdá se mi, že ticho se staví okolo mne jako zeď a už už jako by se čímsi rozpukávalo. A když se ohlédnu, jako by se kmitl nad tmavou cestou cíp bílého roucha [...]“ (Ibidem, s. 82–83, 86)

209 Jinou z narativních technik, která se podílí na tematizaci časoprostoru jako nadosobního řádu spojeného s konkrétním místem, je prodlužování času vyprávění (časté jsou zde deskripce), kterým se zpomaluje čas příběhu (u Čepa vzniká zajímavý efekt spočívající téměř v jeho zastavení) a do popředí se dostává význam prostorovosti.

p. Kozlovice (Dvojí domov):

Chalupy sedí kolem románské kaple s barokovým křížem a moudří lidé praví, že se dějí ve světě veliké věci a že se teď musí všechno zjinačit, protože to už nejde jako zastara. „Tak, tak“, přisvědčují sedláci a chalupníci, ohledavše před spaním zámky u vrat a dobytek v chlévě. Rozpukané dlaně bolí a pole od cesty k úvozu nutno zítra přeorat, hrudy roztlouct a ještě na lůžku cítíš, jak se ti zažírání do duše rez starostí. Ale ze zídky za Spurných zahradou vyrazilo dnes do rána mlíčí, žáby v rybníčku na návsi spouštějí do snů zeleně zářící drahokamy a celý svět plný vůně hrne se otevřeným oknem do světničky a chce po tobě bůhví co. „Musíme mít nejmíň deset muzikantů, ale řekneme savínským, naši jsou moc velcí páni!“ Bedřich Spurných, prosloviv tuto řeč, potahuje kalhoty a bude se bít, neprosadí-li svou. (Ibidem, s. 35)

p. Zbloudilý (Vigilie):

Stavení toho odpoledne zpusťlo, hlasy v sousedních světnicích ztichly, na dvorku zůstal převrhnutý dětský vozík a pootevřenými vrátky nakukovala čísi koza. Nebe se zatáhlo, pahorky uzavřely kolem dokola krajinu náhle vymřelou a mstivou, cesty nešly nikam a hladina rybníka se leskla ocelově a nepřátelsky. Petr Kleofáš vyšel z domu a pustil se cestou, která se mu namanula, a nikde nepotkal živé duše. Na močálovitých lukách s uschlou travou stály nad černými tůněmi pahýly starých vrb. Zamračené lesíky, provanuté dechem staroby a smrti, zahryzávaly se do chudých strání s ponurou psí nevráživostí. Na druhé straně za rybníkem krčily se pod tísnivou šedivostí oblohy hrbaté střechy nuzné vsi. (Čep 1928, s. 17)

V první ukázce ve vyprávěcím segmentu opět dominuje hledisko postavy, což je signalizováno jednak antropomorfizací (*chalupy sedí*), jednak intenzivnější identifikací mezi vyprávěčem a postavou (*Rozpukané dlaně bolí a pole od cesty k úvozu nutno zítra přeorat, hrudy roztlouct a ještě na lůžku cítíš, jak se ti zažírání do duše rez starostí*). Zde je narativní výpověď významově obohacena nejen o zorný úhel postavy, která něco vidí nebo sleduje, ale rovněž o intenzivní emocionální prožitek – bolest dlaní, úkoly, které má člověk před sebou, starosti zažírající se do duše a vyvolávající vnitřní neklid. Výrazně subjektivní modalitu evokuje rovněž rychlá změna pohledů: úvodní část je krátce zaměřena na celek (*chalupy sedí kolem románské kaple*), posléze je vystřídána intenzitou vnitřního prožitku a tematizací ryze individuální zkušenosti. Následně opět dochází k přesunutí do oblasti rámcového časoprostoru (topografie), který je reflektován z hlediska postavy, avšak formálně realizován vyprávěcím modelem *er-formy* (*Ale ze zídky za Spurných zahradou vyrazilo dnes do rána mlíčí, žáby v rybníčku na návsi spouštějí do snů zeleně zářící drahokamy a celý svět plný vůně hrne se otevřeným oknem do světničky a chce po tobě bůhví co*). Tyto narativní úseky jsou navíc přerušovány přímou řečí, která vytváří významný kontrast posilující aspekt subjektivního hlediska a vnímání okolního prostoru postavou. Podíváme-li se na druhou povídku, pozorujeme, že zde sice rovněž můžeme mluvit o jisté infiltraci subjektivního hlediska – krajina je vymřelá a mstivá, lesíky jsou zamračené, hrbaté střechy vesnice se krčí apod. –, avšak intenzita, s jakou se zde tyto prvky realizují, není ani zdaleka tak výrazná jako v předešlém případě. Dochází zde k mnohem pozvolnějšímu,

plynulejšímu budování prostorového fikčnětopografického rámce, který mnohem více odpovídá modální intenci er-formového vypravěče.

Následující ukázka z povídky *Čí bude vítězství* dokládá další z narativních „figur“, která směřuje na rozdíl od *Dvojího domova* k objektivizaci vyprávěného časoprostoru, čímž zakládá jeho epictější šíři. Časoprostor se tak stává epicky stabilizovanější, což umožňuje realizovat širší příběhové linie postav.

p. Čí bude vítězství (Vigilie):

Nuzná chatrč se zahrádkou podél zídky, zarostlou kopřivami a rmenem, s okny plnými prachu a tesknoty neobydlených stavení, s opuštěným příkopem podle plotu, v němž se kdysi válávaly loukotě zlamaných kol a odraný dětský kočárek s vyhrzlými peřinami. Od dob, co se z ní vystěhovali Hrabalovi na dolní konec vesnice, nesnesla tato chalupa, nasáklá potem skrblické rodiny, jediného trvalého obyvatele. Od té doby, co kolář Vincenc Hrabal, potahuje v pasu kalhoty vyztužené staletou špínou, naložil svůj ponk na vůz tažený spřežením strakatých krav, zatímco jeho žena, pobíhající kolem nábytku převráceného vzhůru nohama, rozmnožovala jen zmatek posledních okamžiků, od té doby, co stará Hrabalka s tvářící vrásčitou a scvrklou jako ropuší kůže přiklopila zející prázdnotu vyrabovaného prostoru zhroucenými vrátky a zápasíc se svým přeraženým křížem s námahou vlekla dolejší část těla za vozem, hrkajícím uprostřed dědiny do nového domova, od té doby pustla chalupa stále žalostněji... (Ibidem, s. 57)

Podobnou situaci vidíme rovněž v povídce *Rozárka Lukášová*.

p. Rozárka Lukášová (Vigilie):

Jeroným Lukáš koupil chalupu v Lípově až po narození svého prvního dítěte a takřka úplně na dluh, ale Rozárka už neměla vzpomínky na jiný domov. Odrolené zdi s houštím kopřiv, po jejichž omítce vídala se sklonkem dne stoupati stín, pěšinka v zahradě mezi dvěma hranicemi dříví a zpěv sedmihlásky na hrušce za stodolou – to jí byly věci známé a přítomné od prvního úsvitu paměti. Byla s nimi spřátelena velmi důvěrně a znala dobře tón, jaký vydává stín větve, když se dotkne měkké sladkosti trávy, a nevýslovnou tesknoutu, když na sebe pohledí bílá silnice, ubíhající do modré dálky, a nížce shrbený štít Lukášovy střechy, uzavírající malý schoulený prostor. Tvář otceva zdála se jí táž od prvního dětství až do dne její smrti: rozkazovala bez odvolání. Tíha starostí a divoká vášnivost povahy spleť se v ní v neobyčejně tvrdý uzel. (Ibidem, s. 23)

Nebo v povídce *Husopas*.

p. Husopas (Vigilie):

Třebas bydlel v pastoušce a jeho hodnost byla nižší než hodnost pacholka a nádeníka, byl to domorodec ryzího jádra a celý utvořen z pralátky své dědiny, jejíž barvy věrně nosil a jejíž tradice byl posledním rytířem. On, dříve hajný a teď pasák hus, a pak ještě starý zvoník Zendulka, jenž vláčoval své dlouhé nohy po pěšinách podél humen a z potíciho se kamení v kapli za oltářem předpovídal déšť – počestná obci kozlovická, plač a naříkej nad ztrátou

nenahraditelnou, neboť nebudeš mít podobných dvou sluhů do konce svých dní! Nevím už ani, který z nich odešel dřív, ale husopas Vrána byl jistě z tvrdšího křemene. Tento zavalitý stařík, jehož tvář byla sice rozryta dlátem času a rozorána blesky nesčíslných pohrom – strnisko bílých vousů trčelo z těch úžasných brázd – přesto však vzdorovala pevností a zdravým ruměncem, ničím nezkrusený stařík běhal dosud křepce za svým vrtošivým stádem a praskání jeho nesmírného biče, ono hrdinské praskání, jež zůstane nezapomenutelným zvukem našich chlapeckých let, rozléhalo se kulatým prostorem kozlovické návsi, slavné praskání, jehož tajemství vzal s sebou husopas do hrobu. Přítomní i budoucí husopasové kozlovičtí, kdo z vás si troufá zapraskat jako nebožtík starý Vrána? (Ibidem, s. 45)

Koncepce časoprostoru se tak ve *Vigiliích* proměňuje, což se děje na pozadí modifikace základní struktury definované kategoriemi subjekt–čas–prostor, ke kterým se výrazněji připojuje kategorie abstrakt. Na rozdíl od významové tendence, kterou nabývá časoprostor ve *Dvojím domově*, kde se stává výrazně subjektivně reflektovanou kategorií, se ve *Vigiliích* výrazněji posunuje do roviny nadosobního časoprostoru, který však není časem neosobním, objektivně realistickým. Subjekt se naopak stává přirozenou, organickou součástí onoho již výše jmenovaného časoprostorového palimpsestu.

p. Rozárka Lukášová (*Vigilie*):

Děvčátko se za chvíli vskutku poněkud utišilo. Nad korunami stromů, pohroužených v sebe v nevyzpytatelném zamyšlení, valil se s teskným šumotem proud záhadného ticha, jež barvila lidská bolest rudými pramenky krve. Proč je to tak? Trp! trp! říkaly hvězdy, zírající dolů s nesmírným soucitem. Rozárka uslyšela jakýsi slib na dně něměho šumění, ale rty se jí zkřivily porozuměním tolik předčasným a krutým. [...] Jeroným Lukáš nejel odpoledne na pole, ale až k nemocné do světnice bylo slyšet, jak kdesi na dvoře zatlouká hřebíky. Ve světnici bylo ticho až bolestné, stíny stromů před okny, prozářené sluncem, pohybovaly se po bílých stěnách s takovou sladkostí a něhou, že se srdce svíralo. Okno bylo otevřené a najednou jím zavanula prudká vůně jasmínu. Nemocná otevřela oči (Rozárka se choulila na stoličce za almarou), vdechla zhluboka nejomamnější vůni této země a zalomivši rukama zaúpěla tak, že se Rozárka v hrůze přitiskla k veřejím. (Ibidem, s. 27, 29)

S uvedenými skutečnostmi souvisí i změna sémantické kategorie pohybu – z motivu cesty (chůze), která je v souboru *Dvojí domov* především reálným pohybem postav ve fikčním světě (srov. malý Jeník v povídce *Domek*, otec, matka a syn v povídce *Bouře*, kráva v povídce *Dvojí domov*, chlapec v povídce *Do města*) a s ním úzce spojeným, ale i omezujícím hlediskem postav, se uvedená kategorie přesouvá k motivu metaforické cesty-pouti (s tím se významně pojí kategorie abstrakt), která se znovu modifikuje v oddíle *Zeměžluč* (srov. kontexty k lemmatu *cesta*). S těmito posuny přirozeně souvisí i změna vyprávěcích strategií. Od souboru *Vigilie* směřuje vyprávěcí strategie k epičnosti, rámuje vlastní příběh širším časoprostorovým záběrem, vytváří širší epický časoprostor.

Uvedené tendence, o kterých hovoříme v souvislosti s knihou *Vigilie*, bychom mohli najít rovněž u některých povídek *Dvojího domova*, např. *Mámení* nebo *Dobyvatel*, které však Čep v dalším vydání, jež se stalo součástí knihy *Zeměžluč*, vynechal. *Vigilie* také disponují ústředním principem synekdochy, avšak na rozdíl od *Dvojího domova* je realizován horizontálně.

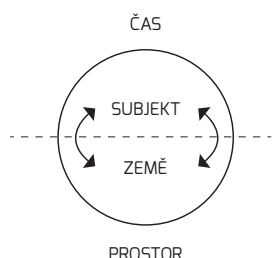


Schéma 20: Horizontalita jako synekdochický princip.

V (sub)korpusu *Zeměžluč – oddíl* (1931) vidíme opět vysoce frekventované lemma *oko*, avšak paradigma tematické oblasti je přece jen oproti oběma předchozím (sub)korpusům modifikované. Z tabulky substantiv a jejich taxonomie je zjevné, že kategorie subjektu (+HUM) se podobně jako v (sub)korpusu *Dvojí domov* (1926) dostává do popředí; je tvořena rovněž hyponymně-hyperonymními vztahy, avšak na rozdíl od uvedeného (sub)korpusu se zde významněji realizují a zapojují i ostatní sémantické kategorie. Zdá se, jako by (sub)korpus *Zeměžluč – oddíl* byl jakousi konfuzí tvořenou oběma předchozími tendencemi, tj. směřováním k výrazné kategorii subjektu a současně směřováním k tematizaci nadosobního časoprostoru. Tento aspekt se promítá rovněž do narativní strategie, která na rozdíl od *Vigilií* (1928) nevytváří tak výrazný epický časoprostorový rámec, avšak není ani zcela bezprostředním uvedením do děje z pozice personického vypravěče, jak je tomu často ve *Dvojím domově* (1926).

p. Starcův smích (Zeměžluč):

Pohyboval se takřka nehmotně mezi stromy řidnoucími v světle měsíčním, v kterém se zdi stávaly průsvitnými a stíny se vyjeveně prodlužovaly. Obešel chalupu, která bděla spíc, a šoural se přes synův dvůr, otevřený jako mrtvá dlaň, nahlédaje do chlévů, z nichž unikalo pokojné oddechování krav. Pak postál u studny a rozhlédl se po nebi rozestřeném v nesmírné teskné dálce, a jeho bezzubá ústa se špičatou bradou se rozpačitě rozvezla a oči se mu naplnily bludným svitem. Bílá hlava se však záhy nachýlila a výměnkář se vlekl zahradou, přičemž rukávy kabátu, který mu visel s ramen, napodobovaly bezvládné pohyby jeho vysílených paží. (Čep 1931, s. 125)

Vyprávěcí er-forma si ve vstupní povídce *Starcův smích* zachovává objektivní odstup od postavy, přestože se v narativním segmentu objeví subjektivně hodnotící části (*dvůr*

jako mrtvá dlaně), které jsou spíše samostatnou vypravěčskou intencí evokovat prostředí nežli výsledkem narativní transformace personického hlediska postavy. Spojnici k narativní strategii *Dvojího domova* vytvářejí tyto incipity zaměřením na centrální událost, kterou v daném případě tvoří starcova reflexe pomíjivosti času a života, vyjádřená vnitřními monology. Podobně tomu je i u dalších povídek.

p. Hoře z lásky (Zeměžluč):

Hleděla chvíli vyjeveně na svou zástěru postříknutou blátem, nic nechápajíc. Učinila jen nepatrný posunek, sotva naznačivši nejistý krok, a pak strnula v zoufalství beze slz a beze slova, sama nad sebou užaslá, že nekřičí, že neběží za tím mužem, který se od ní vzdaloval po rozmoklé silnici, vinoucí se k šedému obzoru. Ještě před měsícem, ještě včera by se nebyla dala takto odbýt. Ale dnes věděla, že je po všem. Pláč v ní vysychal u samého zdroje a srdce i údy jí ztuhly v nehybnost mrtvolky. (Ibidem, s. 131)

p. Veselá pohřební (Zeměžluč):

Pacholek dlouho nevstával. Děvečky, málem ještě v polospánku, zívající a s roztřesenými vlasy pobíhaly po zásplí (zářijové ráno bylo už chladné a bosé nohy zábly) a hospodyně, pozbyvši trpělivosti, pootevřela už podruhé dveře do stáje a zavolala: „No tak, Filipe, hneš-li sebou!“

Stavějíc hrnce na plotnu a jsouc obklopena pološerem nevyvětrané kuchyně, jejíž nepořádek (podzimní práce nenechávaly chvílky oddechu) se kladl na dno jejího vědomí jako nejasná tesknost životní bíd, sotva měla kdy, aby si v duchu zahubovala na čeledína – byl to vlastně pouhý chlapec – že se tahá do noci s holkami a ráno ho pak není možná dostat z pelechu. Kašpar – její muž – přišel z hospody také ne brzy, slyšela jeho boty, jak jedna po druhé udeřily o podlahu, slyšela ho, jak jí cosi povídá, a cosi mu na to i odpovídala, ale její duch, zatěžkaný nesmírným balvanem únavy, spočíval na dně přehlubokých vod jako ukovaný, a když několikrát marně zacloumala závažím, poddala se znova sladkému tonutí a vlny spánku se přes ni opět převálily. (Ibidem, s. 140–141)

Ve druhé ukázce uvedený incipit předchází vlastní centrální události, jíž je sebevražda mladého člověka. Ačkoli v povídkách *Zeměžluče* stojí v centru dění subjekt, který „je středem sítě motivů, souběžnicí hodnotového tázání textu“ (Kubíček 2014, s. 140), neplatí už jako ve *Dvojím domově*, že „je sjednocujícím bodem perspektivy“ (Ibidem, s. 140). Vnímání a reflexe postav obvykle zřetelně rámcuje narativní hlas objektivní nebo komentující (rétorické) er-formy.²¹⁰ Ta spoluurčuje významnou situaci subjektu ve fikčním světě, jež se ve výsledku promítá do roviny nadosobního času a prostoru překračující

²¹⁰ „Je konec,“ opakovala si, octnuvši se najednou tváří v tvář slepé a bezcitné zdi a vidouc za sebou mosty pobožené. Stála uprostřed rozjeté silnice a dívala se tupě na kalužiny, čeřící se loudavými kruhy podzimního deště. [...] Oči opuštěné ženy, odrážející zoufalou šedost oblohy a skleslost únavy nevýslovné, zatěkaly po liduprázdném obzoru a náhle se v nich cosi změnilo, jejich výraz se bolestně oživil a ústa se jí krutě zkřivila. Teprve teď si všimla, že u kraje silnice stojí kamenný kříž, lůžko nenáviděného a nepochopitelného utrpení. Dala se na útěk, motajíc se jako opilá.“ (Čep 1931, s. 132)

úzký rámec subjektem (postavou) bezprostředně vnímaného historického místa a subjektivního času.

p. *Starcův smích* (*Zeměžluč*):

Nad korunami švestek vyhlédají za zídou střechy statků, kreslíci po tolik večerů touž křivku na západním nebi, a stařec si najednou, jako by se před ním byla rozhrnula jakási záclona, v nesmírném ohromení uvědomuje přelud času, který nám představuje věci, jako by šly za sebou, a s rozšířenými očima hledí zpět na řadu dní a let, jež se mu po celý život kladly vedle sebe jako místnosti oddělené neprůhlednými zdmi – a hle, stěny jsou znenadání jako tenká průsvitná plátna a obrysy událostí se najednou prostupují a obličej se kryjí, přecházejíce jeden v druhý nepatrnou změnou rysů kolem očí a úst – „ach, jak je to vlastně pořád totéž,“ žasne stařec a cítí, že takto je všechno rázem mnohem jednodušší, ale zároveň strašnější a nepochopitelnější, toto nekonečné vyměňování kukly, příprava k vzletnutí záhadného motýla... A rázem chápe, že čas je mukou bytostí padlých, strašlivým trestem zatracenců, věcí neznámou v končinách věčné blaženosti. (Ibidem, s. 128–129)

Motiv cesty je v *Zeměžluči* přirozeně spojený s kategorií času a prostoru. Jeho symbolický význam je mnohem výraznější, než je tomu ve *Vigiliích*.²¹¹ Jak vidíme ze vztahů mezi sémantickými kategoriemi substantiv, nejenže je základní tematická struktura osnovy „subjekt–čas–prostor“ výrazně indexována zástupnými znaky konkrétního významu, ale významně se na tuto osnovu napojuje také třída abstrakt. V *Zeměžluči* dochází k mnohem zásadnější a explicitnější tematizaci symbolického významu motivu cesty, který se nespojuje výhradně s postavami příběhu, ale rovněž s vyprávěcí ich-formou. V jejím kontextu je tento motiv spojen s významem uplynulého času, zašlé vzpomínky i poutnictví v existenciálním smyslu.

p. *Starosvětská* (*Zeměžluč*):

Ve světničce s trámovým stropem se už docela setmělo a hlavy dědoušků sotva prosvítají v bledém odrazu oken. Už nerozeznávám jejich tváře a za chvíli mi docela splynou s tmou. Chvilku mi sice ještě utkví v paměti jejich podoba, dovedu si ještě připomenout zvuk jejich hlasu, ale jednoho dne se přistihnu, že mi už jejich rysy vymlel čas. Tento trámový strop, který se klene nad jejich hlavami, tady možná zůstane, ale jiní lidé budou pod ním jíst a spát. Potok za zahrádkou bude bublat hluchým uším a já, naskytne-li se mi někdy jít touto cestou, udělám velkou okliku, abych neviděl přes plot do okénka, za kterým mi zhasly ty dva drahé stíny. (Ibidem 1931, s. 167)

S uvedenými motivy i tematikou je úzce spojen motiv smrti, který je ostatně v Čepových prózách obvyklý. Tento motiv je současně jakýmsi výrazným katalyzátorem, který

211 Srov. kontexty lemmatu *cesta* a významně posílenou sémantickou třídu abstrakt a zejména vazby mezi sémantickými třídami subjekt, prostor, abstrakta (viz Schéma 18).

spojuje oblasti jevového a nadsmyslového světa, spojnicí, na jejíž ose se projektuje vztah mezi generacemi. Pro Čepovy prózy je velice důležitý vztah, jež postavy zaujímají k oblasti transcendentna. Důležitý je moment (náhlého) rozpoznání vlastní limitovanosti vzhledem k nezměrnosti času a prostoru, což se realizuje buď z personické narativní perspektivy, nebo na pozadí příběhové linie, jejímž ústředním motivem je symbolický motiv cesty-životní pouti. Postavy Čepových próz se téměř vždy konfrontují se smrtí, avšak jestliže se ve *Dvojím domově* a *Vigiliích* smrt stává komplementární součástí světa postav, které reflektují a zkoumají její hranice, takže se ve výsledku s tímto motivem propojují v organický celek, v povídkách *Zeměžluče* vystává mezi postavami a uvedeným tématem většinou ostrý předěl. Postavy před smrtí stojí v němém a nechápavém úžasu. V *Zeměžluči* je mezi světem živých a mrtvých mnohem výraznější propast, než je tomu v předchozích dvou knihách.

p. Veselá pohřební (*Zeměžluč*):

Teď tu leží nehybně po nenapravitelném činu, se zkřivenými dětskými ústy, odpovídaje strnulým pošklebkem na výlevy matčiny bolesti a lásky, na všechny dohady a na všechna obviňování, sám už teď tváří v tvář skutečnosti, sám vědoucí mezi nevědomými, ale skrblicí lakotně svým poznáním, jehož nečekal a jež ho ohromilo, a jeho zlomyslný úsměv jako by říkal zevlounům a pomahačům: „Zkuste a uvidíte!“ Ale ti jsou hlší a lhostejní k řeči nebožtíkově a rozmrzení nářkem žebraččiným začínají se rozcházet, hovoříce o posledních oračkách a o blízkosti setby. (Ibidem, s. 145–146)

3.6 Synekdocha jako ústřední organizující princip

Radikální úpravou prvních dvou knih a jejich včleněním do souboru *Zeměžluč* dosáhl Čep toho, že celá kniha je kompozičně vyváženým celkem. Lze se domnívat, že úpravy prováděné u druhých vydání *Dvojího domova* a *Vigilií* probíhaly nejen v dosahu strukturně-intenčního pole oddílu *Zeměžluč*, ale že i poslední oddíl této knihy byl psán po zkušenostech s předchozími dvěma knihami. To se projevuje nejen v dominanci ústředních motivů a témat, ale rovněž v invariantním řešení sémantických tříd a jejich vztahů. Ve *Vigiliích* (1928) je oslaben výrazný motiv pohybu (v korpusu reprezentovaný lemmatem *jít*). Na rozdíl od *Dvojího domova* (1926) je rovnoměrněji rozložen mezi kategorie času a prostoru. To ovšem neznámá, že by tento motiv byl vytěsňen, pouze byl z jednoznačné strukturní dominanty, kterou tvořil v prvním vydání *Dvojího domova* (1926), rovnoměrněji včleněn mezi ostatní sémantické kategorie. Ve druhé verzi *Vigilií* (1931) dokonce sledujeme, že kategorie subjektu nabývá dominantnějšího postavení; lemma *jít* je odsunuto na čtvrtou pozici. V oddíle *Zeměžluč* (1931) se dokonce ocitá pod hranicí *h-bodu*. Zajímavé ovšem je, že v (sub)korpusu *Vigilie* (1928) se nad *h-bodem* vedle

lemmatu *jít* objevuje lemma *cesta*, které ve své významové struktuře obsahuje aspekt pohybu. Lemmata *jít* a *cesta* pak vytvářejí úzkou sémantickou korelaci, která se v jednotlivých knihách významově modifikuje. Je však rozdíl, pokud se Čepův vypravěč soustředí na vlastní akt pohybu postavy, k čemuž ve *Dvojím domově* využívá personického hlediska vypravěče, než když tematizuje motiv cesty, která sice implikuje pohyb, avšak na rozdíl od lemmatu *jít* v sobě zahrnuje širší prostorovou perspektivu ústící do symbolické roviny. *Jít* předpokládá užší perspektivu postavy, zatímco *cesta* sugeruje komplexnější perspektivu časoprostoru.

V analyzovaných Čepových prózách dochází k modifikaci kategorie času i prostoru takovým způsobem, že individuálně prožívané a zakoušené kategorie, jež jsou implikovány narativně zprostředkovaným vnímáním postavy, jejím subjektivním prožitkem, se postupně proměňují v nadosobní významy času a prostoru představující konstanty, které rámuji životní pouť. Jak vidíme na frekvenčních distribucích, princip synekdochčnosti je stále zachovávan, tematika se radikálně nemění. Modifikuje se však pojetí obou základních kategorií (času a prostoru) tím, že se strukturně proměňují vztahy mezi sémantickými kategoriemi. Tyto proměny strukturní povahy jsou empiricky reprezentovány a indikovány frekvenční distribucí těch zástupných kvalit – lemmat, jež náleží pod příslušné sémantické kategorie. Samotná analýza frekvenčního rozložení lemmat je jen dílčím projevem globální textové struktury, ke které lze přistupovat exaktním empirickým zkoumáním. Druhou nezbytnou částí je kontext. Z kontextové analýzy lemmatu *jít* jsme usoudili na proměnu jeho použití v uměleckém narativu, což se opět promítá do vyšší strukturní úrovně celkové strategie textu, tzn. do oblasti vypravěče, který od *Vigilií* (1928) zakládá epičtější vyprávění. Oproti úžeji specifikovanému kontextu lemmatu *jít* ve *Dvojím domově* (1926) se ve *Vigiliích* (1928) kontext rozšiřuje o další kategorie, jež korespondují s tendencí k modifikovanému významu prostoru jako stabilní nadindividuální hodnotové kategorie. Lze konstatovat, že jestliže je časoprostor ve *Dvojím domově* (1926) výrazně reflektován skrze subjekt, jehož aktivní vnímání, prožitek (úzkost) určuje této kategorii ve fikčních světech specifické sémantické *axis mundi* jako výrazně subjektivně prožívané (a znejistěné) kvality, ve *Vigiliích* (1928) se akcentuje sémantika nadindividuálního prostoru, s čímž souvisí jeho explicitnější pojetí jako axiologické konstanty. Vrcholu dosahuje tento význam v *Zeměžluči* (1931), kde je časoprostor explicitně tematizován jako symbolická, nadosobní kvalita, která nejenže rámuje příběhy postav, ale dalece je přesahuje. Uvedené významové mody obou základních kategorií jsou strukturními projekcemi fikčních světů Čepových povídek. Nejsou reálnými kvalitami, ale kvalitami sémantickými a jako takové souvisí s celkovou strukturní strategií uměleckého textu, pochopitelně nejen s jedinou strukturní rovinou.

Podobně jako v případě souboru *Dvojí domov*, jehož výslednou redukcí došlo k posílení subjektivně-modálního hlediska a konceptu vertikality, tak i v případě *Vigilií* dochází ve druhé verzi povídkového celku, jež se spolu s *Dvojím domovem* stala součástí knihy *Zeměžluč*, ke zdůraznění motivu cesty, jenž symbolizuje koncept horizontality a přesouvá se směrem k symbolickému významu. Společným svorníkem obou tendencí, které zásadním způsobem zacházejí s kategoriemi času a prostoru, se stává ústřední organizující princip synekdochy. Na úrovni kvantitativně-korpusové analýzy je tento poznatek potvrzen nejen v oblasti analýzy tematiky, ale též v dílčích sondách, které jsme provedli na vybraných Čepových (sub)korpusech. Základní rozměr Čepovy synekdochy se u uvedených souborů (a ve výsledku nejen u nich) realizuje na ose hyponymně-hyperonymních vztahů jako provázanost subjektu (vědomá či nevědomovaná) s transcendentálním rozměrem lidské zkušenosti rozpínající se na dvou základních osách, které tvoří čas (vertikála) a prostor (horizontála).

Kvantitativně-korpusová analýza ukazuje nejen exaktní manifestaci tohoto obecného strukturního principu v oblasti námi sledovaných příznakových autosémantik a jejich sémantických kategorií, ale rovněž proces utváření výsledné podoby celku knihy *Zeměžluč*. Vrátime-li se zpět k zobrazení průběhů relativních frekvencí v jednotlivých (sub)korpusech a v upravených verzích prvních dvou povídkových knih, konstatovali jsme, že knižní celek *Zeměžluč* tvořený třemi samostatnými oddíly představuje z tohoto hlediska směřování k analogičnosti, kterou demonstrovají frekvenční průběhy ústředních motivických impulzů. Zdá se, že k úpravám prvních dvou povídkových knih dochází v intencním poli strukturní organizace oddílů *Zeměžluč*. Avšak to ještě neznamená, že oddíly *Dvojí domov* a *Vigilie* jsou pouze pasivně připojeny. Ve skutečnosti se jedná o vzájemnou strukturní provázanost všech částí, které se propojují nejen ve smyslu časové následnosti, ale též návratnosti. Oddíl *Zeměžluč* vzniká jako logické vyústění prvních dvou knih, současně je celý soubor *Zeměžluč* dán modifikací prvních dvou verzí *Dvojího domova* a *Vigilií*, která již probíhá pod vlivem komplexní struktury prvních tří povídkových knih a její centrální dominanty, kterou tvoří synekdochický vztah subjektu, času a prostoru. Výsledný soubor *Zeměžluč* je sice projevem autorova záměru, ale autor není vyvázán z „radiace“ uvedeného strukturního pole, jehož je nedílnou a přirozenou součástí. To je současně dalším projevem synekdochičnosti v Čepově tvorbě, jež jako centrální organizující princip v důsledku přerůstá svět umělecké fikce a prostupuje se se světem autora.

Čep v závěru autobiografického eseje *Sestra úzkost* (1975), který psal v emigraci, na jednom místě poznamenává: „[N]enapsal jsem všechny knihy, které jsem mohl napsat, a ty, které jsem napsal, mohl jsem napsat líp. Hlavně jsem nenapsal **knihu**, kterou jsem měl napsat a kterou už nikdy nenapíši.“ (Čep 1975, s. 125) S určitou nadsázkou lze říci,

že Čepovým celoživotním úsilím je psát a cizelovat stále stejnou knihu. Jeho touha po dokonalém vyjádření a uměleckém zpracování ústřední myšlenky je současně jeho osobní reflexí vztahu mezi člověkem, světem a transcendentnem. Čepovým úsilím bylo transformovat své pojetí světa silně prostoupené křesťanskou vírou do umělecké formy, resp. pomocí uměleckého tvaru si danou skutečnost *de facto* uvědomovat. Tvorba se u Čepa prostupuje s náboženskou vírou a *vice versa*, avšak jedno nesupluje druhé. V tomto kontextu tedy nabývá pojem synekdochy nového významu. Projektuje se jako sjednocující princip nejen uvnitř fikčních světů Čepových próz, nejen na rozhraní jednotlivých knih, ale rovněž na rovině vztahu díla, autora a jeho neustálého mučivého hledání vlastní existenciální jistoty.

4. Závěr

V humanitních vědách orientovaných na problematiku jazyka je kvantitativní a korpusová analýza tradičně spojována s lingvistikou. Kvantitativní lingvistika jako samostatný obor se začala výrazně prosazovat zejména od druhé poloviny 20. století v souvislosti s rozvojem počítačové technologie. Lingvisté brzy dokázali využít možností rychle se rozvíjejícího moderního oboru – kybernetiky, a to zejména v oblasti strojového zpracování lexikonu. První pokusy o počítačem zpracovaný slovník vycházely z tzv. děrnoštítkové metody. V českém prostředí je toho příkladem autorský slovník Jitky Štindlové *Konkordanční a frekvenční index k Slezským písním Petra Bezruče*. Podobné práce vznikaly i v zahraničí, jak jsme doložili v prvním oddíle. První výzkumy tohoto typu se pochopitelně prováděly na literárních textech, které byly v předkorpusovém období spolehlivým zdrojem jazykového materiálu. Výzkum však byl orientován lingvisticky a jeho úkolem bylo na základě exaktní strojové kvantifikace demonstrovat obecné charakteristiky a zákonitosti lexikonu, mezi které náleží hodnoty, jako jsou bohatství, rozptyl (disperze), koncentrace slovníku nebo vztahy mezi frekvencí lexikálních jednotek a jejich kvalitativní charakteristikou a distribucí v rámci jednotlivých frekvenčních pásem.

Na materiálu české prózy uvedené otázky systematicky zkoumala především Marie Těšitelová, která spolu s Josefem V. Bečkou a Jaroslavem Jelínkem v roce 1961 vydala *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Základní metodologické postupy spojené s kvantitativní analýzou lexikonu Těšitelová formulovala v publikaci *Otázky lexikální statistiky* (1974), kde svůj výzkum založila na analýze konkrétních literárně-uměleckých textů. Potenciál, jež kvantitativní metody nabízejí literární vědě, rozpoznal i Lubomír Doležel ve studii *Pražská škola a statistická teorie básnického jazyka* (1965). Doleželova argumentace se týká zejména otázek možného využití statistiky pro zkoumání textové normovanosti a aktualizace, pro odlišování ne-básnického a básnického jazyka nebo pro stanovení míry konstantnosti a variabilnosti jazykových prostředků v uměleckých textech (Doležel 1965, s. 105–106). Doleželovy úvahy se přirozeně opírají o strukturalistické základy a mimo jiné dokládají úzký vztah mezi strukturální metodologií a kvantitativní analýzou: „Pražská teorie básnického jazyka byla již charakterizována jako teorie statistické povahy. Ačkoli tato charakteristika je v podstatě správná, přece nelze zapomínat, že v pražských pracích nalézáme rovněž velmi cenné analýzy kvalitativně strukturních vztahů v básnickém textu. [...] Na druhé straně musím zdůraznit, že pražská teorie básnického jazyka má charakter prestatistický, protože zde nebyl aplikován

ani ve vymezení, ani v analýze básnického jazyka skutečný aparát matematické statistiky. Hlavní věty teorie jsou statistické svým obsahem, nikoli svou formulací.“ (Ibidem, s. 103–104) Zejména poslední tvrzení je důležité, neboť odkazuje k povaze kvantifikace raných strukturalistických prací. Doležel jasně upozorňuje na rozdíl mezi kvantifikací ve smyslu vlastní metody, která se řídí nejen specifickými postupy, ale determinuje výslednou povahu analýzy, a kvantifikací ve významu pouhého doplnění metodologicky jinak zaměřeného výzkumu. V druhém případě slouží statistické informace jako příloha nekvantitativního výzkumu.

Ačkoli uvedené otázky, a zejména pak výsledky práce Marie Těšitelové vybízely k jejich testování v oblasti literární vědy, k podobným pokusům nedošlo. Veškeré formulace tak zůstaly pouze na teoretické rovině, jak dokládá i studie Miroslava Klivara *Význam kvantitativní lingvistiky pro literární teorii* (1984), ve které autor *de facto* reprodukuje výsledky M. Těšitelové. Pouze na několika málo úvodních místech studie Klivar konstatuje některé velmi obecné skutečnosti: „Prokazatelné výsledky kvantitativní lingvistiky naznačují řadu podnětů pro literární vědu. Kvantitativní lingvistika byla definována jako vědecký obor, jehož předmětem je zjišťovat statistické zákonitosti jazykových jevů a jejich vztahů a získávat statistické údaje o jednotlivých jevech a na jejich základě charakterizovat analyzované celky. Při aplikaci statistických metod se postupuje od statistiky přes počet pravděpodobnosti k matematické statistice. Vychází se od studia souborů přes studium množin souborů k systému množin souborů, k vytváření modelů. [...] Z hledisek teorie literatury jde pak o další nadsystém, v němž lingvistika je jedním ze subsystémů vědecké soustavy výzkumných oborů a metod podřízených literárněteoretickému ‚supersystému‘. [...] Na druhé straně si uvědomujeme určité relativní hranice, v nichž se kvantitativní lingvistika pohybuje, a které nemůže překročit. Za prvé je to jednak její základní zakotvení v lingvistické oblasti, z čehož plyne její vztah k širším nadsystémům v teorii literatury zcela bezprostředně a jasně. Za druhé je to její těžiště v kvantitativní stránce metod výzkumu, jak jsme viděli sice v jednotě s kvalitativním hlediskem (již ve výběru materiálu pro výzkum), ale v jednotě kvantity a kvality je těžiště v kvantitativní stránce vztahu této ‚strojové‘ lingvistiky, jak bývá také někdy nazývána. Za třetí je z hlediska teorie literatury zjevné, že uvedená specifika jsou jen dílčí v celkové koncepci tří vrstev uměleckého díla – hmotné, významové a obsahové. Je tu řada dalších specifik, například ve významové vrstvě vztahy času vnitřního a vnějšího aj., která se projevují ve významové vrstvě jinak než ve hmotné, jsou tu hodnoty ideologické, filozofické, složitý souhrn hodnot mimoestetických, které nesouvisejí se zákonitostmi zkoumanými kvantitativní lingvistikou.“ (Klivar 1984, s. 24, 34)

V druhé polovině 20. století byly otázky týkající se možností kvantitativní analýzy v literární vědě promýšleny spíše hypoteticky, její hlavní vývoj byl orientován zcela jiným

směrem. Od 60. let se v literární teorii výrazně prosazuje naratologie, která přichází s vlastními otázkami a problémy. Krize strukturálního myšlení, se kterým je kvantifikace úzce spojena, je reflektována například v dekonstrukci. Současná literární teorie disponuje rozsáhlou škálou tematických okruhů, jimž věnuje pozornost – gender studies, nový historismus, nová cenzura, reprezentace paměti, postkoloniální studia nebo výrazný příklon k otázkám prostoru, jež nabývají na aktuálnosti zejména od konce 80. let.²¹² Speciálně orientované bádání dnes představuje jeden z výrazných směrů (nejen) literárněvědného myšlení, jež postupně opouštělo vyhraněnou oblast metodologických diskuzí a přecházelo k otázkám spíše ideovým či ideologickým a interdisciplinárním. Ačkoli se na různých úrovních lingvistika a literární věda stále setkávaly, obě disciplíny se přesto vyvíjely samostatně, což je ostatně jedním z hlavních znaků svěbytných vědních oborů. Kvantitativní a formální analýza založená na strojovém algoritmu se stala znovu aktuální především na počátku 90. let, kdy začínají vznikat moderní jazykové korpusy.²¹³ Teprve nedávné výzkumy potvrzují, že literární věda a lingvistika ve spojení s moderní technologií mohou tvořit užší celek. Umožňují to dnešní moderní elektronické jazykové korpusy. Spojení výsledků, které nabízí korpusová a moderní kvantitativní lingvistika, s literárněvědným bádáním se ovšem ne vždy setkává s porozuměním. Jedním z kritiků spojení kvantitativně-korpusové lingvistiky a literární vědy je i holandský literární teoretik Willie van Peer, který ve studii *Quantitative Studies of Literature. A Critique and an Outlook* (1989) konstatuje: „I take it that quantitative studies of literature do nothing less than that: trying to describe as accurately as possible the various linguistic elements and configurations one encounters in literary works of art.“²¹⁴ (Peer 1989, s. 302) Na druhé straně se právě kvantitativní i korpusová analýza ve světě čím dál častěji využívají v oblasti literárněvědného bádání. Jak poznamenává David L. Hoover, možnosti tohoto přístupu se stále rozšiřují: „Quantitative approaches are most naturally associated with questions of authorship and style, but they can also be used to investigate larger interpretive issues like plot, theme, genre, period, tone, and modality.“²¹⁵ (Hoover 2008, s. 517)

V čem spočívá tato názorová dichotomie, kdy je na jedné straně odmítán kvantitativně-korpusový přístup v literární vědě, zatímco na straně druhé, zejména v zahraniční literatuře, vidíme stále rostoucí a zdokonalující se úsilí o vzájemné propojení? Van Peerův názor významně koresponduje s tím, co uvedl již Klivar, totiž že literární systém

212 Viz Bachmann-Medick 2007, s. 286n.

213 Český národní korpus byl založen v roce 1994. Dostupné z WWW: <<http://ucnk.ff.cuni.cz/>>

214 [„Jsem toho názoru, že kvantitativní studium literatury nedělá nic jiného, než že se pokouší co nejpřesněji popsat různé jazykové prvky a konfigurace, se kterými se setkáváme v uměleckém literárním díle.“]

215 [„Kvantitativní přístupy jsou nejpřirozeněji spojeny s otázkami autorství a stylu, ale mohou být rovněž použity k objasnění širších interpretačních problémů, jakými jsou plot, téma, žánr, období, řečový projev a modalita.“]

je složitým komplexem různých strukturních a významových úrovní, které nelze redukovat na společný základ, jímž je jazyková rovina.

Tuto skutečnost nejenže nelze popřít, ale naopak je třeba ji zdůraznit. Literární umělecké dílo je komplexním jevem, který kromě řady imanentních aspektů zahrnuje i oblast širšího rádiusu. Především se jedná o otázky spojené s dobovým a kulturním kontextem díla, kam náleží aspekty recepční nebo pragmatické. Literární dílo nelze vyčerpat z hlediska jediné metodologie; jeho povaha komplexního kulturního jevu, jenž je navíc nositelem složitých a mnohdy variabilních významových kvalit, dovoluje dokonce řadu protichůdných interpretací či metodologických postupů. Od žádné literárněvědné metody nelze proto právem očekávat konečné a jediné relevantní řešení celé řady složitých otázek, jež s problematikou literárního díla souvisí. Smysl plurality literárněvědných diskurzů však spočívá v tom, že umožňuje nezbytnou kritickou konfrontaci mezi jednotlivými přístupy, která stojí na pozadí jejich vlastního odborného vývoje. Otázky kladené napříč metodologickými přístupy nejenže problematizují samotný předmět zkoumání, ale také rozvíjejí každý z těchto teoretických přístupů. Nejinak je tomu i v tomto případě. Očekávat od kvantitativní či korpusové analýzy uměleckého slovesného narativu, že odpoví na komplexní otázky spojené s povahou literárního díla, je základem fatálního omylu. Stejně tak je mylné se domnívat, že uvedená metoda a její pracovní postupy mohou nahradit vlastní literárněvědnou analýzu nebo interpretaci. O ničem takovém zde nebyla řeč. Ostatně ani v lingvistice, která pracuje s korpusy, neplatí, že by samotné korpusové výsledky postačovaly k relevantnímu vysvětlení příslušných jevů.

Je tedy třeba zdůraznit, že korpus stejně jako metoda a postupy statistické a kvantitativní analýzy slouží především jako nástroje, jejichž potenciál spočívá v tom, že v kontextu humanitních věd umožňují maximálně objektivizovat některá tvrzení. Je rovněž mylné se domnívat, že formální analýza jazykové vrstvy uměleckého díla se týká výhradně a pouze této vrstvy. Ačkoli platí, že se kvantitativně-korpusová analýza explicitně vztahuje k jazykové složce uměleckého díla, nemusí referovat výhradně a pouze k ní. Samotné kvantitativní i korpusové techniky náleží obecně pod statistiku jazyka a záleží mimo jiné rovněž na výzkumné strategii, kterou budou využívány. Zde se ocitáme poblíž Doleželova výroku, kterým rozlišil dvojí zacházení se statistickými metodami. Van Peerova kritika pokládá kvantifikaci za pouhé počítání množství výskytů určitého prvku, nicméně kvantita je rovněž odrazem strukturní organizace. Navíc kvantitu v tomto případě nelze nikdy myslet bez kvalitativních aspektů, jimiž jsou v případě literárního díla jazykové jednotky a jejich sémantika. Van Peer rovněž upozorňuje na to, že různé roviny slovesného díla nelze stejně úspěšně kvantifikovat. Čím výše půjdeme na úrovni strukturní abstrakce, tím méně úspěšná bude kvantifikace jejích komponent (viz pozn. 133). To je věc zcela nepopíratelná. Na druhé straně kvantitativně-korpusový přístup neusiluje

o absolutní kvantifikaci všech složek uměleckého díla, ale pouze těch, které jsou úspěšně kvantifikovatelné. Slovem úspěšně se rozumí, že počítat lze pouze takové prvky, které je možné relativně přesně a objektivně identifikovat, např. slova (lexémy). To však automaticky neznamena, že se výzkum založený na kvantifikaci lexémů uměleckého díla musí omezit výhradně na danou oblast. Struktura uměleckého díla je komplexní organizovanou jednotou, ve které jsou vzájemně propojeny nejen prvky na téže strukturní úrovni, ale i jednotlivé strukturní úrovně. Nelze například myslet sémantiku fikčního světa a jeho entit bez funkční komplementarity jazykových, narativních, kompozičních, tematických, motivických a celé řady dalších aspektů, které sice lze teoreticky izolovat a relativně samostatně zkoumat, ale které ve výsledku vytvářejí homogenní celek, organickou strukturu uměleckého díla.

Řekli jsme, že je otázkou pracovní strategie, za jakým účelem budou kvantitativně-korpusové metody využívány. Formální analýza umožňuje zkoumat nejen prostý počet výskytů izolované jednotky, např. slova, ale současně korpusové nástroje dovolují také formalizovat kontexty (kolokace), jež jsou důležité z hlediska významu. Každá jednotka se nejen vřazuje do svého kontextového okolí (syntagmatický aspekt), ale disponuje i paradigmatickou, což například představuje její organizaci pod společnou sémantickou třídou. Charakteristiku sémantických tříd lze realizovat s ohledem na lingvistickou i literárněvědnou klasifikaci (subjekt, objekt, nástroj, prostor, čas apod.). Ve výsledku tedy dochází k propojení několika formálně analyzovatelných vrstev a oblastí – kvalitativní aspekt je určen typem jednotky (nemusí se jednat pouze o slovo, ale například o kolokaci), jejím kontextem a paradigmatickým, tzn. příslušností do nadřazených sémantických tříd a skupin; kvantitativní aspekt specifikuje příznakovost užití dané jednotky v systému. Vždy ovšem platí, že relevantních výsledků se lze dobrat teprve srovnáváním dílčích i komplexnějších jevů. Zejména vyšší jednotky, v našem případě sémantické třídy a jejich vztahy, odkazují k dalším strukturním činitelům v rámci výstavbové strategie uměleckého textu, především k oblasti narativních technik, kompozice, tematiky apod. Samozřejmě se zde již nelze spoléhat výhradně na výsledky kvantitativně-korpusové analýzy, neboť se ocitáme v oblastech komplikovaně kvantifikovatelných. Přesto jsme schopni využít účelně vybudované strategie kvantitativně-korpusové analýzy uměleckého textu k úvahám, které vedou do jeho „vyšších pater“. Výstupní informace uvedené formální analýzy spočívají především v absolutních hodnotách, tedy nemohou substitovat individuální analyticko-interpretační výkon. Absolutní hodnoty jsou však přesně tím, co od kvantitativně-korpusové analýzy očekáváme. Tyto informace jsou validní právě v důsledku své objektivity a „plošné“ reference k fundamentálním oblastem strukturního systému uměleckého textu.

V současné době je vyvíjeno úsilí taxonomizovat vyšší segmenty uměleckého textu, například různé narativní techniky a postupy: „Words themselves, as the smallest clearly meaningful units, are the most frequently counted items, and syntactic categories (noun, verb, infinitive, superlative) are also often of interest, as are word n-grams (sequences) and collocations (words that occur near each other). Thematic or semantic categories (angry words, words related to time), while more difficult to count, have the advantage of being clearly relevant to interpretation, and automated semantic analysis may reduce the effort involved. Phrases, clauses, syntactic patterns, and sentences have often been counted, as have sequences or subcategories of them (prepositional phrases, subordinate clauses, passive sentences). Many of the items above are also used as measure of the lengths of their items: word length in characters, sentence or clause length in letter or words, text length in words, sentences, paragraphs, and so forth. Nonlinguistic textual division ranging from small units like lines and couplets to large structural units like paragraphs, stanzas, scene, acts, and characters (including subcategories like first-person and third-person narrators, and characters divided by age, ethnicity, nationality, class, and gender), and plot elements (marriages, deaths, journeys, subplots). [...] Counting may be limited to the dialogue or narration of a text, to one or more speakers or narrators, or to specific passages.“²¹⁶ (Hoover 2008, s. 519–520) V tomto ohledu je pozoruhodná například studie Larryho L. Stewarta *Charles Brockden Brown: Quantitative Analysis and Literary Interpretation* (2003), ve které autor kvantifikuje rozdíly mezi jednotlivými narativními promluvami vybrané Brockdenovy novely. Stewartova zjištění učiněná na základě kvantifikací vybraných textových markerů²¹⁷ odhalila vnitřní shody mezi zdánlivě rozdílnými personickými vypravěči v různých částech textu. Kromě toho je Stewartova studie dobrým příkladem toho, jak lze kvantitativních analytických metod smysluplně a funkčně využít v literárněvědné práci. Kvantitativní zkoumání soustředěné na umělecké texty má čím dál větší tendenci se prosazovat v zahraniční odborné literatuře, jak

216 [„Samotná slova jako nejmenší jednoznačně plnovýznamové jednotky jsou nejčastějším komponentem kvantifikace. Předmětem zájmu jsou často také syntaktické kategorie (podstatné jméno, sloveso, infinitiv, superlativ), stejně jako otázka n-gramů (posloupnost) a kolokací (slova vyskytující se ve vzájemné blízkosti). Tematické nebo sémantické kategorie (expresivní slova, slova s příznakem temporality), i když jsou obtížně kvantifikovatelné, mají tu výhodu, že jsou pro interpretaci zcela vhodné. Jejich automatická sémantická analýza může snížit vynaložené úsilí. Fráze, vedlejší věty, syntaktické vzorce a hlavní věty byly často kvantifikovány stejně jako jejich části nebo subkategorie (předložkové fráze, vedlejší věty, pasivní větné konstrukce). Řada komponent je často používána k určení délky svých částí: délka slov určena počtem znaků, délka vět určená počtem písmen nebo slov, délka textu počtem slov, vět, odstavců atd. Nelingvistická segmentace textu se rozvíjí od malých částí, jakými jsou verš a kuplet, po větší strukturní celky jako odstavce, scény, jednání, postavy (včetně narativních subkategorií ich-formy, er-formy a postav specifikovaných dle věku, etnické příslušnosti, národnosti, třídy a pohlaví) a základní části příběhu (svatby, umrtí, cesty a zápletky). [...] Kvantifikace může být omezena na dialog nebo akt vyprávění jednoho či více vypravěčů nebo na konkrétní textové pasáže.“]

217 Viz Stewart 2003, s. 132–133.

o tom svědčí mnohé články v časopise *Literary and Linguistic Computing*²¹⁸ nebo ve sbornících, např. *A Companion to Digital Humanities*,²¹⁹ *Digital Literary Studies – Corpus Approaches to Poetry, Prose, and Drama* (2014), kde je mimo jiné k dispozici studie zabývající se stylovou charakteristikou řečových promluv v dramatu. Postupně se této problematice začaly věnovat i samostatné monografie, např. rozsáhlá práce Jana Christopfera Meis-tera z roku 2003 *Computing Action (Narratologia)*, monografie Roberta Franzosiho *Quantitative Narrative Analysis* (2010)²²⁰ nebo připravovaná monografie *Quantitative Analysis of Poetic Texts* Ioana-Iovitze Popescu, Mihaiele Lupey, Diony Tatar a Gabriela Altmanna.²²¹ Výčet prací na dané téma zdaleka nekončí.²²²

Kvantitativně-korpusová analýza, jež byla v této práci představena a realizována, je vedle svých konkrétních výsledků především testováním možností jejího využití v literární vědě. Proto ani prezentované postupy nepřinášejí definitivní a vyčerpávající způsob, jakým lze těchto metod v literárněvědném kontextu využívat.²²³ Specifickou otázkou je například analýza sloves a jejich kontextů v závislosti na typu narativní či řečové promluvy, která by byla schopna významně obohatit korpusový přístup v literárněvědném kontextu. Tato otázka je ovšem sama o sobě složitější, neboť – jak jsme viděli právě u Čepa – je mnohdy velice obtížné stanovit jasné a přesné hranice konkrétní promluvy. Ačkoli by takový postup byl teoreticky možný (viz úpravu Doleželova schématu), při praktické segmentaci se ukazuje, že zde stále existuje vysoké procento subjektivního rozhodování při určení zvláště sporných narativních typů.²²⁴ Jinou skutečností je, že Doleželova taxonomie promluv ve fikčních narrativech není samozřejmě jedinou možnou klasifikací vypravěče a postav. Například již pro Genettovu klasifikaci dvou základních vyprávěcích typů (homodiegetický a heterodiegetický vypravěč) nelze nalézt adekvátní algoritmus, který by dokázal formálně segmentovat text. Každá další specifická segmentace textu na korpusové úrovni na druhou stranu zvyšuje jeho hodnotu, zejména pak jeho další upotřebitelnost, která nebude směřovat výhradně do oblasti lingvistiky.

218 Teprve nedávno nově přejmenováno na *Digital Scholarship in the Humanities*. Dostupné z WWW: <<http://dsh.oxfordjournals.org/>>

219 Mnohé z nich jsou dostupné na internetu.

220 Viz heslo „Computational Narratology“ v *Living Handbook of Narratology*. Dostupné z WWW: <<http://www.lhn.uni-hamburg.de/article/computational-narratology>>

221 Mimo jiné je aktuálně vynakládáno nemalé úsilí na algoritmizaci těch segmentů textu, které nesou speciální informaci; s tím souvisí i rozvoj speciálního metajazyka (viz Snoussi – Gensel – Davoine 2012; Mani – Doran – Harris – Hitze-man – Quimby – Richer – Wellner – Mardis – Clancy 2010).

222 Často bývá formální analýza uměleckých textů využívána k určování autorství.

223 „Kvantitativní přístup k literární analýze může mít různou formu, od dějin knihy přes počítačnickou stylistiku, tematické databáze až po všestrannou lexikální analýzu a další.“ (Moretti 2014, s. 12)

224 Se segmentací promluv v narativu pracují Elena Semino a Mick Short v publikaci *Corpus Stylistics. Speech, Writing and Thought Presentation in a Corpus of English Writing* (2004).

Kromě snahy využívat moderní korpusy a kvantitativní metody k analýze jemnějších významových segmentů uměleckých textů (styl, vypravěč, postava, prostor, verš, metrum²²⁵ apod.), což se jeví jako zcela zásadní otázka, se korpusové techniky dají výborně využít především pro analýzu velkých dat. Konkrétně máme na mysli potenciální existenci větších korpusů literárních textů, které by byly speciálně anotovány a segmentovány právě pro literárněvědné účely. Výsledky korpusového zpracování řady uměleckých textů by mohly přispět rovněž k některým aktuálně řešeným otázkám literární vědy, mezi které náleží také nové pojetí literárního vývoje představené v kolektivní práci *České literární romantično* (2012), jež v podtitulu nese označení pracovní metody, tzv. synopticko-pulzačního mapování kulturních jevů. Podstatou tohoto přístupu je výzkum široké oblasti (nejen) literárních textů, které zahrnují různé směry, tendence, poetiky a školy. Výsledkem je potom mapování složitých vývojových uměleckých procesů jako v čase a prostoru nesymetricky a nelineárně rozložených kvalit (poetických symptomů), jež různým způsobem a rozdílnou intenzitou manifestují svoji dobově historickou a uměleckou příslušnost. Na základě tohoto přístupu se ukazuje, že zejména přechody mezi uměleckým směry nejsou tvořeny ostrými předěly, ale že se jedná o nezřetelné a velice dynamické hranice, které navíc nelze vzhledem ke složitosti fenoménu literárního díla a jeho fungování v kulturním kontextu jednoznačně a zcela definitivně vyložit. Excerpci, třídění a generování absolutních hodnot nejen jednotlivých textů, ale i množin textů a jejich tříd by korpusová databáze nejenom zjednodušila, ale byla by schopna přinést takové výsledky kvantitativně-korpusové analýzy, jež by svojí objektivitou dokázaly jistě přispět k prováděnému výzkumu.

Kvantitativně-korpusová analýza tím, že uspořádává jevy na základě jejich kvantitativního potenciálu, umožňuje objektivizovat strukturní vlastnosti konkrétního systému, ve kterém jsou tyto jevy uspořádány a organizovány. V případě analýzy literárního díla to jistě neznamena, že se tím vyčerpávají veškeré jeho strukturní vlastnosti. Podstatné je ovšem to, že pomocí formální taxonomizace a algoritmizace lze relativně objektivně a exaktně zkoumat fundamentální rovinu uměleckého díla, která s vyššími rovinami vytváří celistvost strukturně organizované jednoty.

225 Výzkum poezie založený na speciálně vybudovaném korpusu je realizován na Ústavu pro českou literaturu AV ČR. Dostupné z WWW: <<http://www.versologie.cz/index.html>>

Quantitative-Corpus Analysis and Literary Studies

Summary

Traditionally, methods of quantitative and corpus analysis have chiefly been applied in linguistics. Only in recent years have (mainly foreign) studies tested their possibilities in the sphere of literary studies. The main purpose of such an approach is to provide literary studies with an exact tool, one based on formally analysed verbal material and not burdened with subjective introspective aspects. The principle of quantitative-corpus analysis does not only comprise the schematised taxonomy of a verbal corpus – it also presupposes its categorical statistical analysis, which is capable of manifesting (representing) absolute structural values and qualities. This is so because quantification is not just the expression of an amount by means of a number, but a representation of the structural distribution of quantified entities. At the same time, this does not mean that a quantitative or corpus approach eliminates the qualitative aspect. The quantification of language is, generally, the representation of a distribution of qualitative elements (words, contexts etc.). Both aspects (the quantitative and qualitative one), therefore, in the present analysis, naturally and organically complement each other.

The aim of the present study is to test the aforementioned possibilities of the functional connection between quantitative-corpus analysis and literary studies, using, as an example, an author-based corpus of the work of Jan Čep, which was compiled for the purpose of the present study from the author's prose texts, including his short-story collections from 1926–1953 and several short-stories only published in magazines.

The basic strategy of author-based corpus analysis is based on so-called relative chronology, which, unlike the usual chronology of an author's work, allows reflection upon overlaps between the periods in which the individual works were created. This aspect enables us to understand an author's work, on the one hand, as a linear development, and, on the other, as a pulsational process of overlapping periods in which the individual works were written (see Scheme 1). This conception opens up the possibility of a functional comparison between the author's individual creative phases, which can be compared both from the point of view of their relative chronology as instances of overlapping areas and, at the same time, from the perspective of quantitative-corpus analysis, which is capable of significantly complementing the aforementioned strategy with an exact approach that deals with the immanent laws of the texts. Both of the main aspects then represent basic poles which complement – as well as determine – each other.

The central feature of the present study is, above all, an analysis of specific quantitative-corpus results, which are subsequently incorporated into separate literary analyses. The analyses focus on Čep's first three short-story collections, which have a very particular character due to their genesis: the first two, *Dvojí domov* (1926) and *Vigilie* (1928), were subsequently radically revised by the author and used as separate sections for another short-story collection, *Zeměžluč* (1931). The present qualitative-corpus analysis is based on the so-called semantic concentration of the text, which works with the distinction between the text's two basic areas, demarcated by the so-called h-point. The h-point in the lexicon defines two zones: one below the h-point in which autosemantics prevail, and one above the h-point in which synsemantics prevail. Through a more developed analysis of thematic words, which also includes semantic classes and contexts (collocations), a unique structural use of synecdoche becomes apparent, which is present and realised in its own way in each of the aforementioned works by Čep. The principle of synecdoche is, however, realised not only as an organising tendency within the individual short-story collections, but also as a macrocompositional strategy of the entire *Zeměžluč* collection. Synecdoche, therefore, surpasses the sphere of individual short-stories and is manifested as the central structural-semantic trajectory of the author's creative intention.

Ediční poznámka

První dvě části monografie byly v redukované podobě publikovány samostatně v časopise *Naše řeč*. Jedná se o studie *Vybraná problematika z autorské lexikografie* (*Naše řeč*, roč. 97, č. 2, 2014, s. 68–77) a *Model korpusově koncipovaného slovníku Jana Čepa v kontextu české autorské lexikografie a kvantitativní analýzy literárního jazyka* (*Naše řeč*, roč. 97, č. 3, 2014, s. 113–131). V upravené a doplněné podobě stejně jako v plném rozsahu jsou uvedené oddíly publikovány v této monografii poprvé.

Bibliografie

a/ Prameny:

- ČEP, Jan. 1921. „Z vesnice“. *Moravskoslezské besedy* III, č. 41–42, s. 3–4, s. 1.
- ČEP, Jan. 1922. „Kde Haná v hory přechází“. *Moravskoslezské besedy* IV, č. 1, s. 1–3.
- ČEP, Jan. 1922. „Touha po středu, po srdci“. *Moravskoslezské besedy* IV, č. 3, s. 1–4.
- ČEP, Jan. 1922. „O pošetilé lásce“. *Moravskoslezské besedy* IV, č. 10, s. 1–5.
- ČEP, Jan. 1922. „Anče“. *Moravskoslezské besedy* IV, č. 20, s. 2–3.
- ČEP, Jan. 1922. „Dětská idyla“. *Moravskoslezské besedy* IV, č. 22, s. 3–4.
- ČEP, Jan. 1922. „Hřeb v hrušce“. *Moravskoslezské besedy* IV, č. 28, s. 3–6.
- ČEP, Jan. 1922. „Kovář“. *Moravskoslezský deník* XXII, č. 184, s. 1–2.
- ČEP, Jan. 1922. „Oheň“. *Moravskoslezské besedy* IV, č. 34, s. 3–6.
- ČEP, Jan. 1922. „Píseň zahrádky“. *Moravskoslezské besedy* IV, č. 34, s. 6–8.
- ČEP, Jan. 1922. „Karel Benda“. *Moravskoslezský deník* XXII, č. 199, s. 1–2.
- ČEP, Jan. 1922. „Červený muškát“. *Moravskoslezské besedy* IV, č. 38, s. 3–6.
- ČEP, Jan. 1922. „Úzkost“. *Moravskoslezský deník* XXV, č. 265, s. 1–3.
- ČEP, Jan. 1922. „Úpění v samotě“. *Moravskoslezské besedy* IV, č. 47, s. 5–7.
- ČEP, Jan. 1923. „Oddělené světy“. *Moravskoslezské besedy* V, č. 2, s. 3–7.
- ČEP, Jan. 1923. „Jindřich, Slávek a ona“. *Moravskoslezské besedy* V, č. 4, s. 6–8.
- ČEP, Jan. 1923. „Kde jsou lidé?“. *Moravskoslezské besedy* V, č. 13, s. 3–5.
- ČEP, Jan. 1924. „Návrat“. *Moravskoslezské besedy* VI, č. 11, s. 3–4.
- ČEP, Jan. 1924. „Táta Čtyrák“. *Moravskoslezský deník* XXVII, č. 219, příl. *Besedy* VI, č. 29, s. 2–3.
- ČEP, Jan. 1924. „Dědina“. *Cesta* VI, č. 50, s. 717–719.
- ČEP, Jan. 1925. „Hrob Janka Dospivy“. *Moravskoslezský deník* XXVIII, č. 176, příl. *Besedy* VII, č. 26, s. 1–2.
- ČEP, Jan. 1925. „Zmrzlá srdce“. *Cesta* VII, č. 14, s. 220–222.
- ČEP, Jan. 1924/1925. „Opilec“. *Černá země* I, č. 8, s. 11–13.

- ČEP, Jan. 1927. „De profundis“. *Cesta IX*, č. 20, s. 309–311.
- ČEP, Jan. 1927. „Oči“. *Cesta IX*, č. 28, s. 443–444.
- ČEP, Jan. 1927. „Putování úvozem“. *Tvar I*, č. 4, s. 73–78.
- ČEP, Jan. 1927. „Noc filipojakubská“. *Tvar I*, č. 7, s. 129–132.
- ČEP, Jan. 1927. „Holubí princezna“. [Rukopis.]
- ČEP, Jan. 1933. „Cesta ze školy“. *Kolo*, č. 1, s. 4–6.
- ČEP, Jan. 1934. „Jakub Neoral“. *Kolo*, č. 1, s. 5–7.
- ČEP, Jan. 1934. „Páter Damian“. *Kolo*, č. 3–4, s. 34–36.
- ČEP, Jan. 1934. „Zastavení“. *Almanach Kmene 1934–1935*, s. 80–84.
- ČEP, Jan. 1935. „Sedláci v hospodě“. *Kolo*, č. 1, s. 2–4.
- ČEP, Jan. 1938. „Detektivní příběh“. *Hovory o knihách II*, č. 1, s. 6.
- ČEP, Jan. 1926. *Dvoji domov*. Praha: Ladislav Kuncič.
- ČEP, Jan. 1928. *Vigilie*. Praha: Rudolf Škeřík.
- ČEP, Jan. 1931. *Zeměžluč*. Praha: Melantrich.
- ČEP, Jan. 1932. *Letnice*. Praha: Melantrich.
- ČEP, Jan. 1934. *Děravý plášť*. Praha: Melantrich.
- ČEP, Jan. 1935. *Hranice stínu*. Praha: Melantrich.
- ČEP, Jan. 1938. *Modrá a zlatá*. Praha: Melantrich.
- ČEP, Jan. 1941. *Tvář pod pavučinou*. Praha: Vyšehrad.
- ČEP, Jan. 1944. *Lístky z alba*. Brno: Josef Stejskal.
- ČEP, Jan. 1946. *Polní tráva*. Praha: František Borový.
- ČEP, Jan. 1953. *Cikáni*. Mnichov: Sdružení československých politických uprchlíků v Německu.
- ČEP, Jan. 1999. „Cikáni a jiné prózy“. In: Čep, Jan. *Polní tráva*. Praha: Vyšehrad, s. 305–375.
- ČEP, Jan. 1956. *Peníz exulantův*. Mnichov: Sdružení československých politických uprchlíků v Německu.
- ČEP, Jan. 1975. *Sestra úzkost*. Řím: Křesťanská akademie.
- GOETHE, Johann Wolfgang. 1964. *Utrpení mladého Werthera. Spříznění volbou*. Přel. E. A. Saudek. Praha: Státní nakladatelství krásné literatury a umění.
- GOETHE, Johann Wolfgang. 1979. *Faust*. Přel. O. Fischer. Praha: Československý spisovatel.

b/ Odborná literatura:

- ABBOTT, Edwin. 1875. *A Concordance to the Works of Alexander Pope*. New York: D. Appleton & Co.
- BACHMANN-MEDICK, Doris. „Spatial Turn“. In: *Cultural Turns. Neuorientierungen in den Kulturwissenschaften*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag, s. 284–328.
- BAKER, Arthur Ernest. 1914. *A Concordance to the Poetical and Dramatic Works of Alfred, Lord Tennyson*. London: Routledge & Kegan Paul Ltd.
- BAKER, Paul – HARDIE, Andrew – McENERY, Tony. 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- BALDWIN, Dane Lewis – BROUGHTON, Leslie Nathan – EVANS, Laura Cooper – HEBEL, John William – STELTER, Benjamin – THAYER, Mary Rebeca. 1963. *A Concordance to the Poems of John Keats*. Washington: The Carnegie Institution of Washington.
- BARTHES, Roland. 2002. „Úvod do strukturální analýzy vyprávění“. In: Kyloušek, Pavel. (ed.): *Znak, struktura, vyprávění. Výbor z prací francouzského strukturalismu*. Přel. J. Fryčera a P. Kylouška. Brno: Host, s. 9–43.
- BARTLETT, John. 1937. *A New and Complete Concordance or Verbal Index to Words, Phrases and Passages in the Dramatic Works of Shakespeare with Supplementary Concordance to the Poems*. London: Macmillan and Co.
- BĚLIČ, Jaromír. 1961. „Předmluva“. In: Jelinek, Jaroslav – Bečka, Josef Václav – Těšitelová, Marie. *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: Státní pedagogické nakladatelství, s. 5–7.
- BESSINGER, Jess B., Jr. – SMITH, Philip H., Jr. 1969. *A Concordance to Beowulf*. Ithaca – New York: Cornell University Press.
- BEVAN, Dean E. 1971. *A Concordance to the Plays and Prefaces of Bernard Shaw*. Detroit – Michigan: Gale Research Company.
- BLATNÁ, Renata. „Metajazyk v lexikografii“. In: Čermák, František – Blatná, Renata. *Manuál lexikografie*. Jinočany: H&H, s. 72–89.
- BRIGHTWELL, Barron. 1970. *A Concordance to the Entire Works of Alfred Tennyson*. New York: Haskell House Publishers Ltd.
- BOOTH, Wayne C. 2007. „Typy vyprávění“. Přel. M. Knápková. *Aluze* 10, č. 2, s. 42–51.
- BRADSHAW, John. 1894. *A Concordance to the Poetical Works of John Milton*. London: W. H. Allen & Co.
- BROUGHTON, Leslie Nathan. 1963. „Preface.“ In: Baldwin, Dane Lewis – Broughton, Leslie Nathan – Evans, Laura Cooper – Hebel, John William – Stelter, Benjamin – Thayer, Mary Rebeca. *A Concordance to the Poems of John Keats*. Washington: The Carnegie Institution of Washington, s. V–XIX.
- BURROWS, F. John. 2007. „All the Way Through: Testing for Authorship in Different Frequency Strata“. *Literary and Linguistic Computing* 22, č. 1, s. 27–47.

- BURROWS, F. John. 1992. „Computers and the Study of Literature“. In: Butler, Christopher S. (ed.): *Computers and Written Texts*. Oxford, UK – Cambridge, USA: Blackwell, s. 167–204.
- COWDEN CLARKE, Mary. 1845. *The Complete Concordance to Shakespeare being a Verbal Index to all the Passages in the Dramatic Works of the Poet*. London: Charles Knight & Co.
- CLEVELAND, Charles Dexter. 1867. *A Complete Concordance to the Poetical Works of John Milton*. London: Sampson Low, Son & Marston.
- COHNOVÁ, Dorrit. 2009. *Co dělá fikci fikci*. Přel. M. Orálek a V. Klusáková. Praha: Academia.
- COOK, Albert Stanburrough. 1911. *A Concordance to Beowulf*. Halle: Max Niemeyer Verlag.
- COOPER, Lane. 1923. *A Concordance of the Latin, Greek and Italian Poems of John Milton*. Saale: Max Niemeyer Verlag.
- CUNLIFFE, Richard John. 1910. *A New Shakespearean Dictionary*. Glasgow–Bombay: Blackie and Son limited 50 old Bailey London E. C.
- CVRČEK, Václav. 2006. „Metoda zjišťování kolokační platnosti frekventovaných bigramů pomocí ranku“. In: Čermák, František a Šulc, Michal (eds.): *Kolokace*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 36–55.
- CVRČEK, Václav. 2008. „Koncept minimální intervence“. *Slovo a slovesnost* 69, č. 4, s. 284–292.
- CVRČEK, Václav. 2013. *Kvantitativní analýza kontextu*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu.
- CVRČEK, Václav – KOVÁŘÍKOVÁ, Dominika. 2011. „Možnosti a meze korpusové lingvistiky“. *Naše řeč* 94, č. 3, s. 113–133.
- CVRČEK, Václav – VONDŘIČKA, Pavel. 2011. „Výzkum variability v korpusech češtiny“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 184–195.
- ČECH, Radek – POPESCU, Ioan-Iovitz – ALTMANN, Gabriel. 2014. *Metody kvantitativní analýzy (nejen) básnických textů*. Olomouc: Qfwfq.
- ČEJKA, Mirek. 1992. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita.
- ČERMÁK, František. 1995. „Jazykový korpus: prostředek a zdroj poznání“. *Slovo a slovesnost* 56, č. 2, s. 119–140.
- ČERMÁK, František. 2004. „Korpusová lingvistika: stručný historický přehled“. In: *Český národní korpus*. Dostupné z WWW: <http://korpus.cz/doc/2002_cnk.rtf>
- ČERMÁK, František. 2005. „Korpus, informace a lingvistika.“ In: *Přednášky z XLVIII. běhu Letní školy slovenských studií*. Praha: Univerzita Karlova, Filozofická fakulta, s. 15–24. Dostupné z WWW: <http://korpus.cz/doc/korp-info-lingv.rtf>
- ČERMÁK, František. 2006a. „Kolokace v lingvistice“. In: Čermák, František a Šulc, Michal (eds.): *Kolokace*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 9–16.

- ČERMÁK, František. 2006b. „Statistické metody hledání frazémů a idiomů v korpusech“. In: Čermák, František a Šulc, Michal (eds.): *Kolokace*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 94–106.
- ČERMÁK, František. 2007. *Frazeologie a idiomatika česká a obecná*. Praha: Karolinum.
- ČERMÁK, František. 2008. „An Author's Dictionary: The Case of Karel Čapek“. In: *Proceedings of the XIII Euralex International Congress*. Barcelona: Universitat Pompeu Fabra, s. 323–329.
- ČERMÁK, František. 2010. *Lexikon a sémantika*. Praha: Nakladatelství Lidové noviny.
- ČERMÁK, František. 2011. „Korpusy včera, dnes a zítra“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 10–29.
- ČERMÁK, František. 1995. „Frazeologie“. In: Čermák, František – Blatná, Renata. *Manuál lexikografie*. Jinočany: H&H, s. 116–136.
- ČERMÁK, František. 1995. „Komputační lexikografie“. In: Čermák, František – Blatná, Renata. *Manuál lexikografie*. Jinočany: H&H, s. 50–71.
- ČERMÁK, František. 1995. „Paradigmatika a syntagmatika slovníku: možnosti a výhledy“. In: Čermák, František – Blatná, Renata. *Manuál lexikografie*. Jinočany: H&H, s. 90–115.
- ČERMÁK, František – KRÁLÍK, Jan – PALA, Karel. 1992. „Počítačová lexikografie a čeština (počítačový fond češtiny)“. *Slovo a slovesnost* 53, č. 1, s. 41–48.
- ČERMÁK, František – BARTOŇ, Tomáš. 2007. *Slovník Karla Čapka*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu.
- ČERMÁK, František – BLATNÁ, Renata. 1995. *Manuál lexikografie*. Jinočany: H&H.
- ČERMÁK, František – BLATNÁ, Renata. 2005. *Jak využívat Český národní korpus*. Praha: Nakladatelství Lidové noviny.
- ČERMÁK, František – BLATNÁ, Renata. 2006. *Korpusová lingvistika: stav a modelové přístupy*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu.
- ČERMÁK, František – CVRČEK, Václav – SCHMIEDTOVÁ, Věra. 2010. *Slovník komunistické totality*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu.
- ČERMÁK, František – CVRČEK, Václav. 2009. *Slovník Bohumila Hrabala*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu.
- ČERMÁK, František – KŘEN, Michal. 2004. *Frekvenční slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- ČERMÁK, František – KŘEN, Michal. 2010. *Frekvenční slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- ČERMÁKOVÁ, Anna. 2006. „Kolokace a valence některých případů substantiv“. In: Čermák, František a Šulc, Michal (eds.): *Kolokace*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 107–141.

- ČERVENKA, Miroslav. 1965. „Nový projekt statistického rozboru verše“. *Česká literatura* 13, č. 6, s. 541–544.
- ČERVENKA, Miroslav. 2006. „Vývojová imanence a kontakty mezi literaturami: májovský daktylotrochej“. In: Zajac, Peter – Tureček, Dalibor – Urválková, Zuzana (eds.): *Mezi texty a metodami*. Olomouc: Periplum, s. 41–71.
- DANĚLKA, Jiří. 1956. „Biblická konkordance“. *Slovo a slovesnost* 17, č. 4, s. 236–237.
- DAVIDOVÁ-GLOGAROVÁ, Jana – ČECH, Radek. 2013. „Tematická koncentrace textu – některé aspekty autorského stylu Ladislava Jehličky“. *Naše řeč* 96, č. 5, s. 234–245.
- DAVIDOVÁ-GLOGAROVÁ, Jana – DAVID, Jaroslav – ČECH, Radek. 2013. „Analýza tematické koncentrace textu: komparace publicistiky Ladislava Jehličky a Karla Čapka“. *Slovo a slovesnost* 74, č. 1, s. 41–54.
- DAVID, Jaroslav – ČECH, Radek – DAVIDOVÁ-GLOGAROVÁ, Jana – RADKOVÁ, Lucie – ŠÚSTKOVÁ, Hana. 2013. *Slovo a text v historickém kontextu. Perspektivy historickosémantické analýzy jazyka*. Brno – Ostrava: Host – Ostravská univerzita.
- DAVIS, Norman – GRAY, Douglas – INGHAM, Patricia – WALLACE-HADRILL, Anne. 1979. *A Chaucer Glossary*. Oxford: The Clarendon Press.
- DILL, Christa. 1959. „Lexika zu einzelnen Schriftstellern“. *Forschungen und Fortschritte* 33, č. 11, s. 340–346, 369–375.
- DILL, Christa. 1987. *Wörterbuch zu Goethes West-östlichem Divan*. Tübingen: Max Niemeyer Verlag.
- DOLEŽEL, Lubomír. 1993. *Narativní způsoby v české literatuře*. Praha: Český spisovatel.
- DOLEŽEL, Lubomír. 1965a. „Pražská škola a statistická teorie básnického jazyka“. *Česká literatura* 13, č. 2, s. 110–113.
- DOLEŽEL, Lubomír. 1965b. „Úvod do informačně teoretické koncepce uměleckých textů“. *Česká literatura* 13, č. 3, s. 258–260.
- DOROSZEWSKI, Witold. 1965. „Uwagi wstępne“. In: Koneczna, Halina a kol. *Słownik języka Jana Chryzostoma Paska*. Wrocław – Warszawa – Kraków: Wydawnictwo Polskiej Akademii nauk, s. V–VIII.
- DYCE, Alexander – LITTLEDALE, Harold. 1970. *A Glossary to the Works of William Shakespeare*. New York: Johnson Reprint Corporation.
- FILIPEC, Josef. „Teorie a praxe jednojazyčného slovníku výkladového“. In: Čermák, František – Blatná, Renata (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 14–49.
- FOŘT, Bohumil. 2006. „S počítači na narativy“. *Česká literatura* 54, č. 5, s. 146–150.
- FRANZOSI, Roberto. 2010. *Quantitative Narrative Analysis*. London: SAGE.
- FUČÍK, Bedřich. 1969. „Básník dvojího domova“. In: Čep, Jan. *Zeměžluč – Letnice – Děravý plášť*. Praha: Československý spisovatel, s. 319–324.

- GÓRSKI, Konrad – HRABEC, Stefan. 1962. *Słownik języka Adama Mickiewicza*. Wrocław – Warszawa – Kraków: Wydawnictwo Polskiej Akademii nauk.
- GRÄFE, Manfred. 1959. „Goethe und die Probleme der neueren deutschen Lexikographie“. *Forschung und Fortschritte* 33, č. 10, s. 303–306.
- HAASOVÁ, Lenka. 1999. „Pokus o stanovení sémantických tříd v díle Jaroslava Seiferta“. In: *Varia VIII. Zborník materiálov z 8. kolokvia mladých jazykovedcov (Modra – Piesok 25.–27. XI. 1998)*. Bratislava: Slovenská jazykovedná spoločnosť pri SAV, s. 75–81.
- HANCOCK, Leslie. 1967. *Word Index to James Joyce's Portrait of the Artist*. London – Amsterdam: Southern Illinois University Press.
- HAUSMANN, Franz Josef. 1989. „Wörterbuchtypologie“. In: Hausmann, Franz Josef, Reichmann, Oskar, Wiegand, Herbert Ernst, Zgusta, Ladislav (eds.): *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Berlin – New York: De Gruyter, s. 968–981.
- HLAVÁČKOVÁ, Dana – GRÁC, Marek – SEDLÁŘOVÁ, Pavlína. 2011. „Budování anotovaného korpusu bez odborníků“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 132–141.
- HOFFMANNOVÁ, Jana – MIKULOVÁ, Marie. 2011. „Korpusy mluvené češtiny a možnosti jejich využití pro poznání rozdílných ‚světů‘ mluvenosti a psanosti“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 78–92.
- HOLMAN, Petr. 1988. „Frekvenční slovník básnického díla Otokara Březiny“. *Naše řeč* 71, č. 1, s. 41–49.
- HOLMAN, Petr. 1993. *Frequenzwörterbuch zum lyrischen Werk von Otokar Březina*. Köln – Weimar – Wien: Böhlau Verlag.
- HOOVER, David L. 2008. „Quantitative Analysis and Literary Studies: History, Goals, and Theoretical Foundation“. In: Schreibman, Susan, Siemens, Ray (eds.): *A Companion to Digital Literary Studies*. Oxford: Blackwell, s. 517–533.
- HOWALD, Blake S. – KATZ, Graham E. 2011. *The Exploitation of Spatial Information in Narrative Discourse*. Dostupné z WWW: <<http://aclweb.org/anthology/W11-0119>>
- HÜHN, Peter a kol. 2009. *The Living Handbook of Narratology*. Hamburg: Hamburg University. Dostupné z WWW: <<http://www.lhn.uni-hamburg.de/>>
- CHATMAN, Seymour. 2008. *Příběh a diskurs: narativní struktura v literatuře a filmu*. Přel. M. Orálek. Brno: Host.
- CHOJAK, Jolanta – PUZYNNINA, Jadwiga – TELEŻYŃSKA, Ewa – WIŚNIEWSKA, Ewa. 1988. *Słownik języka Cypriana Norwida: zeszyt próbny*. Warszawa: Uniwersytet Warszawski.
- IBRAHIM, Robert – PLECHÁČ, Petr. 2014. *Báseň a počítač*. Praha: Ústav pro českou literaturu Akademie věd České republiky.
- JAKOBSON, Roman. 2005. *Formalistická škola a dnešní literární věda ruská*. Praha: Academia.

- JANOVEC, Ladislav. 1998. „Využití programu Concorde pro tvorbu autorských slovníků“. In: *Varia VII*. Bratislava: Slovenská jazykovedná spoločnosť pri SAV, s. 63–70.
- JANOVEC, Ladislav. 1999. „Z autorského slovníku Jaroslava Seiferta – sémantická charakteristika a použití zájmen já a my v některých sbírkách“. In: *Varia VIII. Zborník materiálov z 8. kolokvia mladých jazykovedcov (Modra – Piesok 25.–27. XI. 1998)*. Bratislava: Slovenská jazykovedná spoločnosť pri SAV, s. 82–91.
- JANOVEC, Ladislav. 2002. „Sémantická charakteristika barev v některých sbírkách Jaroslava Seiferta“. In: *Varia IX*. Bratislava: Slovenská jazykovedná spoločnosť pri SAV, s. 37–42.
- JANOVEC, Ladislav. 2003. „Autorský slovník Jaroslava Seiferta – současný stav, změny, koncepce a budoucnost projektu“. In: *Varia X*. Bratislava: Slovenská jazykovedná spoločnosť pri SAV, s. 197–202.
- JELÍNEK, Jaroslav – BEČKA, Josef Václav – TEŠITELOVÁ, Marie. 1961. *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: Státní pedagogické nakladatelství.
- KAPLAN, David M. – BLEI, David M. 2007. „A Computational Approach to Style in American Poetry“. *IEEE International Conference on Data Mining 7*, s. 545–550.
- KARLÍK, Petr – NEKULA, Marek – PLESKALOVÁ, Jana. 2002. *Encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- KLIMENTOVÁ, Jana. 1965. „Modelování rytmické výstavby verše Markovskými procesy“. *Česká literatura* 13, č. 1, s. 69–72.
- KASTNEROVÁ, Martina. 2011. *Shakespeare a teorie interpretace. Hledání adekvátního interpretačního přístupu*. Plzeň: Západočeská univerzita.
- KLIVAR, Miroslav. 1984. „Význam kvantitativní lingvistiky pro literární teorii“. In: *Literárněvědný sborník Památníku národního písemnictví*. Praha: PNP, s. 24–52.
- KONECZNA, Halina a kol. 1965. *Słownik języka Jana Chryzostoma Paska*. Wrocław – Warszawa – Kraków: Wydawnictwo Polskiej Akademii nauk.
- KOPŘIVOVÁ, Marie – KOCEK, Jan. *Manuál korpusového manažeru Bonito*. Dostupné z WWW: <https://korpus.cz/bonito/stat.php>
- KRÁLÍK, Oldřich. 1961. „Autorské slovníky“. *Slovo a slovesnost* 22, č. 3, s. 239–240.
- KRAUS, Jiří. 1968. „Stroje v jazykovědě“. *Naše řeč* 51, č. 5, s. 295–297.
- KŘEN, Michal. 2013. *Odras jazykových změn v synchronních korpusech*. Praha: Nakladatelství Lidové noviny.
- KŘEN, Michal. 2006. „Kolokační míry a čeština: srovnání na datech Českého národního korpusu“. In: *Kolokace*. Praha: Ústav Českého národního korpusu, s. 223–248.
- KUBÍČEK, Tomáš. 2007. „Kdo vypráví vypravěče?“ *Aluze* 10, č. 1, s. 42–47.
- KUBÍČEK, Tomáš. 2014. *Dvojí domov Jana Čepa*. Brno: Host.
- KUCAŁA, Marian. 1994. *Słownik polszczyzny Jana Kochanowskiego*. Kraków: Instytut języka polskiego Polskiej Akademii nauk.

- LEVÝ, Jiří. 1972. *Bude literární věda exaktní vědou? Výbor studií*. Praha: Československý spisovatel.
- LITTLEDAL, Harold. 1970. „Reviser's Preface“. In: Dyce, Alexander – Littledale, Harold. *A Glossary to the Works of William Shakespeare*. New York: Johnson Reprint Corporation, s. VII–X.
- LOCKWOOD, Laura Emma. 1907. *Lexicon to the English Poetical Works of John Milton*. New York: Macmillan Co.
- MACHOVÁ, Svatava. 1997. „Autorský slovník básnického díla Jaroslava Seiferta“. In: *Sborník Pedagogické fakulty Univerzity Karlovy*. Filologické studie XX, s. 29–34.
- MATTAUSCH, Josef. 1982. „Kommunikativ-pragmatische Sprachbetrachtung und Lexikographie. Zur Spezifik des Autorenwörterbuchs.“ *Zeitschrift für Germanistik* 3, s. 303–313.
- MATTAUSCH, Josef. 1990. „Das Autoren-Bedeutungswörterbuch“. *Historical Lexicography of the German Language (Studies in German Language and Literature)* 6, č. 2, s. 1549–1562.
- MATTAUSCH, Josef. 1991. „Textlexikographische Aspekte im Autorenwörterbuch (am Beispiel des Goethe-Wörterbuchs)“. *Historical Lexicography of the German Language II (Studies in German Language and Literature)* 2, č. 6, s. 713–733.
- MANI, Inderjeet – DORAN, Christy – HARRIS, Dave – HITZEMAN, Janet – QUIMBY, Rob – RICHER, Justin – WELLNER, Ben – MARDIS, Scott – CLANCY, Seamus. 2010. „SpatialML: Annotation, Scheme, Resources, and Evaluation“. *Language Resources and Evaluation* 44, č. 3, s. 263–280.
- MEISTER, Jan Christoph. 2003. *Computing Action (Narratologia)*. Berlin – New York: Walter de Gruyter.
- MERKER, Erna. 1966. *Wörterbuch zu Goethes Werther*. Berlin: Akademie Verlag.
- MORETTI, Franco. 2014. *Grafy, mapy, stromy: abstraktní modely literární historie*. Přel. O. Čaplyginová. Praha: Karolinum.
- MUKAŘOVSKÝ, Jan. 1985. „Polákova vznešenost přírody – pokus o rozbor a vývojová zařazení básnické struktury“. In: Mukařovský, Jan. *Studie z poetiky*. Praha: Odeon, s. 504–505.
- ONIONS, Charles Talbot. 1911. *A Shakespeare Glossary*. Oxford: Clarendon Press.
- PANEVOVÁ, Jarmila. 1962. „Využití strojů při zkoumání češtiny“. *Slovo a slovesnost* 23, č. 3, s. 222–227.
- PEER, Willie van. 1989. „Quantitative Studies of Literature. A Critique and an Outlook“. *Computers and the Humanities* 23, s. 301–307.
- POPESCU, Ioan-Iovitz – MAČUTEK, Ján – ALTMANN, Gabriel. 2009. *Aspects of Word Frequencies*. Lüdenscheid: RAM-Verlag.
- POPESCU, Ioan-Iovitz – MAČUTEK, Ján – ALTMANN, Gabriel. 2014. *Aspekty frekvencí slov*. Olomouc: Qfwfq.
- POŘÍZKA, Petr. – SCHÄFER, František. 2010. „Svět jako vědomí a nic Ladislava Klímy v olomouckém korpusu české esejistiky přelomu 19. a 20. století.“ In: Gilk, Erik – Hrabal, Jiří (eds.): *Věčnost není dřevá kapsa, aby se z ní něco ztratilo. Soubor studií věnovaných Ladislavu Klímovi*. Olomouc: Příloha e-1 revue *Aluze*, s. 60–67.

- POŘÍZKA, Petr. 2011. „K možnostem počítačového zpracování literárního textu“. In: Fedrová, Stanislava (ed.): *Česká literatura v intermedialní perspektivě: IV. Kongres světové literárněvědné bohemistiky. Jiná česká literatura (?)*. Praha: Akropolis – ÚČL AV ČR, s. 485–495.
- PUZYNINA, Jadwiga. 1988. „Wstęp“. In: Chojak, Jolanta – Puzynina, Jadwiga – Teleżyńska, Ewa – Wiśniewska, Ewa. *Słownik języka Cypriana Norwida: zeszyt próbny*. Warszawa: Uniwersytet Warszawski, s. 1–8.
- RAMSAY, Stephen. 2008. „Algorithmic Criticism“. In: Schreibmann, Susan – Ramsay, Stephen (eds.): *A Companion to Digital Literary Studies*. Oxford: Blackwell, s. 476–491.
- ROELCKE, Thorsten. 1994. „Individualsprache und Autorenlexikographie (Einführung)“. *Lexicographica X*, s. 1–20.
- RÖSEL, Hubert. 1983. *Wörterbuch zu den tschechischen Schriften des J. A. Comenius*. Münster: Aschendorf.
- SEMINO, Elena – SHORT, Mick. 2004. *Corpus Stylistics. Speech, Writing and Thought Presentation in a Corpus of English Writing*. London – New York: Routledge.
- SKORUPKA, Stanisław – SZLIFERSZTEJNOWA, Salomea. 1965. „Podstawy pracy redakcyjne nad słownikiem“. In: Koneczna, Halina a kol. *Słownik języka Jana Chryzostoma Paska*. Wrocław – Warszawa – Kraków: Wydawnictwo Polskiej Akademii nauk, s. VIII–XXI.
- SCHADEWALDT, Wolfgang. 1966. „Einführung“. In: Schadewaldt, Wolfgang – Werner, Simon – Wissmann, Wilhelm. *Goethe-Wörterbuch (Erster Band, 1. Lieferung A – abrufen)*. Stuttgart – Berlin – Köln – Mainz: Akademie Verlag Berlin, s. IV–XVI.
- SCHADEWALDT, Wolfgang – WERNER, Simon – WISSMANN, Wilhelm. 1966. *Goethe-Wörterbuch (Erster Band, 1. Lieferung A – abrufen)*. Stuttgart – Berlin – Köln – Mainz: Akademie Verlag Berlin.
- SCHMID, Wolf. 2004. *Narativní transformace*. Přel. P. Málek. Brno: Ústav pro českou literaturu Akademie věd České republiky.
- SCHMIDT, Alexander – SARRAZIN, Gregor. 1902. *Shakespeare-Lexicon. A Complete Dictionary of All the English Words, Phrases and Constructions in the Works of the Poet*. New York: G. E. Stechert.
- SCHMIEDTOVÁ, Věra. 2011. „Nadávky a kletby v dílech Karla Čapka, Bohumila Hrabala a v korpusu ORAL2008“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 287–302.
- SNOUSSI, Mouna – GENSEL, Jérôme – DAVOINE, Paule-Annick. 2012. „Extending TimeML and SpatialML Languages to Handle Imperfect Spatio-Temporal Information in the Context of Natural Hazards Studies“. *AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27*. Dostupné z WWW: <http://w.agile-online.org/Conference_Paper/CDs/agile_2012/proceedings/papers/Paper_Snoussi_Extending_TimeML_and_SpatialML_languages_to_handle_imperfect_spatio-temporal_information_in_the_context_of_natural_hazards_studies_2012.pdf>.
- STEWART, Larry L. 2003. „Charles Brockden Brown: Quantitative Analysis and Literary Interpretation“. *Literary and Linguistic Computing* 18, č. 2, s. 129–138.

- ŠTÍCHA, František. 2011. „O skrytosti langue v korpusovém parole aneb k definici ‚kvantové lingvistiky‘“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 41–59.
- ŠTINDLOVÁ, Jitka. 1969. *Konkordanční a frekvenční index k Slezským písním Petra Bezruče*. Praha – Ostrava: Socialistická akademie v Ostravě.
- ŠULC, Michal. 2001. „Tematická reprezentativnost korpusů“. *Slovo a slovesnost* 62, č. 1, s. 63–61.
- ŠULC, Michal. 2006. „Slovní asociace versus korpusové kolokace“. In: Čermák, František a Šulc, Michal (eds.): *Kolokace*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 374–402.
- ŠULC, Michal. 1999. *Korpusová lingvistika: první vstup*. Praha: Karolinum.
- TEJNOR, Antonín. 1990. „In memoriam Jitky Štindlové“. *Naše řeč* 73, č. 5, s. 262–264.
- TĚŠITELOVÁ, Marie. 1948. „Frekvence slov a tvarů ve spise ‚Život a dílo skladatele Foltýna‘ od Karla Čapka“. *Naše řeč* 32, č. 6–7, s. 126–130.
- TĚŠITELOVÁ, Marie. 1955. „Poznámky k slovní zásobě v románu Karla Čapka ‚Život a dílo skladatele Foltýna‘“. *Naše řeč* 38, č. 9–10, s. 297–307.
- TĚŠITELOVÁ, Marie. 1968. „O básnickém jazyce z hlediska statistického“. *Slovo a slovesnost* 29, č. 4, s. 362–368.
- TĚŠITELOVÁ, Marie. 1974. *Otázky lexikální statistiky*. Praha: Academia.
- TĚŠITELOVÁ, Marie. 1987. *Kvantitativní lingvistika*. Praha: Státní pedagogické nakladatelství.
- TĚŠITELOVÁ, Marie. 1997. „100 let od vydání prvního frekvenčního slovníku“. *Naše řeč* 80, č. 4, s. 190–194.
- TODD, Henry John. 1809. *Verbal Index to the Poetry of Milton*. London: Law and Gilbert.
- TRÁVNÍČEK, Mojmír. 1996. *Pouť a vyhnanství: život a dílo Jana Čepa*. Brno: Proglas.
- TRÁVNÍČEK, Mojmír. 1999. „Vydavatelská poznámka“. In: Čep, Jan. *Polní tráva*. Praha: Vyšehrad, s. 376–379.
- TRÁVNÍČEK, Mojmír. 2011. „Bibliografie“. In: Čep, Jan. *Červený muškát a jiné prózy z let 1921–1938*. Šternberk: Sternberg, s. 298–300.
- TUREČEK, Dalibor a kol. 2012. *České literární romantično. Synopticko-pulzační model kulturního jevu*. Brno: Host.
- TYL, Zdeněk. 1953. „Slovník jazyka Puškinova“. *Sovětská věda (jazykověda)* 3, s. 80–81.
- TYL, Zdeněk. 1985. „Slovník k českým spisům J. A. Komenského“. *Naše řeč* 68, č. 5, s. 240–249.
- UMBACH, Horst. 1986. „Individual Sprache und Gemeinsprache. Bemerkungen zum Goethe-Wörterbuch“. *Zeitschrift für germanistische Linguistik*, č. 14, s. 161–174.

- VALIŠOVÁ, Pavlína. 2011. „Výukové materiály založené na korpusu“. In: Čermák, František (ed.): *Korpusová lingvistika Praha 2011 – 2: Výzkum a výstavba korpusů*. Praha: Nakladatelství Lidové noviny a Ústav Českého národního korpusu, s. 313–323.
- VASÁK, Pavel. 1966. „Statistika a sporné autorství“. *Slovo a slovesnost* 27, č. 4, s. 364–370.
- WALSH, Richard. 2007. „Kdo je vypravěč?“ *Aluze* 10, č. 1, s. 48–60.
- WIEGAND, Herbert Ernst. [1986] 2000. „Bedeutungswörterbücher oder sogenannte Indices in der Autorenlexikographie? Die Eröffnung einer Kontroverse.“ In: Wiegand, Herbert Ernst (ed.): *Kleine Schriften: Eine Auswahl Aus Den Jahren 1970 Bis 1999: 2 Bde.: Eine Auswahl aus den Jahren 1970–1999 in zwei Bänden* (1. díl). Berlin – New York: Walter de Gruyter, s. 635–641.
- WILSON, Andrew. 2009. „Vocabulary Richness and Thematic Concentration in Internet Fetish Fantasies and Literary Short Stories“. *Glottology* 2, č. 2, s. 97–107.
- WYKA, Kazimierz. 1962. „Przedmowa“. In: Górski, Konrad – Hrabec, Stefan: *Słownik języka Adama Mickiewicza*. Wrocław – Warszawa – Kraków: Wydawnictwo Polskiej Akademii nauk, s. V–VI.
- ZAJAC, Peter. 2006. „Literárne dejepisectvo ako synoptická mapa“. In: Tureček, Dalibor – Urválková, Zuzana (eds.): *Mezi texty a metodami*. Olomouc: Periplum, s. 13–23.

c/ Elektronické korpusy:

- Český národní korpus – SYN2010. Ústav Českého národního korpusu FF UK, Praha 2010. Dostupné z WWW: <<http://www.korpus.cz>>
- Český národní korpus – cep. Ústav Českého národního korpusu FF UK, Praha 2015. Dostupné z WWW: <<http://www.korpus.cz>>¹
- Český národní korpus – SYN2010 BEL. Ústav Českého národního korpusu FF UK, Praha 2010. Dostupné z WWW: <<http://www.korpus.cz>>
- Český národní korpus – capek. Ústav Českého národního korpusu FF UK, Praha 2007. Dostupné z WWW: <<http://www.korpus.cz>>²
- Český národní korpus – capek_uplny. Ústav Českého národního korpusu FF UK, Praha 2007. Dostupné z WWW: <<http://www.korpus.cz>>³
- PLECHÁČ, Petr – ZMĚLÍK, Richard. 2015. *KORBEL. Manažer beletristických korpusů. Autorský slovník Jana Čepa*. Dostupné z WWW: <slovníkcep.g6.cz/index.php>

1 Korpus není veřejně přístupný.

2 Korpus není veřejně přístupný.

3 Korpus není veřejně přístupný.

d/ Internetové zdroje a odkazy:

A Companion to Digital Humanities. 2004. Dostupné z WWW: <<http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-8&toc.depth=1&toc.id=ss1-2-8&brand=default>>

Alliance of Digital Humanities Organizations. Dostupné z WWW: <<http://adho.org/>>

Das Goethe-Wörterbuch im Internet. Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften. Dostupné z WWW: <<http://gwb.uni-trier.de/de/>>

European Association for Digital Humanities. Dostupné z WWW: <<http://eadh.org/>>

Korpus českého verše. Dostupné z WWW: <<http://www.versologie.cz/kcv.html>>

Literary Studies in the Digital Age. An Evolving Anthology. Dostupné z WWW: <<https://dlsanthology.commons.mla.org/>>

e/ Software:

ČECH, Radek – KUBÁT, Miroslav – MATLACH, Vladimír. *QUITA – Quantitative Text Analyzer*. (Version 1.1.2.1). Dostupné z WWW: <<https://code.google.com/p/oltk/>>

Jmenný rejstřík

Abbott, Edwin 39, 40, 41
Aischylós 3, 38
Akvinský, Tomáš 66
Alighieri, Dante 14, 38, 39, 45
Altmann, Gabriel 221
Arnold, Matthew 47

Baker, Arthur Ernest 41, 42
Baldwin, Dane Lewis 44
Bartlett, John 42, 43, 50
Bartoň, Tomáš 72
Bauer, Jaroslav 57
Bečka, Josef Václav 25, 215
Bělič, Jaromír 25
Bessinger, Jess Balsor Jr. 46, 47
Bevan, Earl Dean 45, 46, 66
Bezruč, Petr 46, 68, 69, 70, 215
Blake, William 47
Blanc, Ludwig Gottfried 39
Blatná, Renata 72
Booth, Wayne C. 106
Botev, Christo 39, 67
Bradshaw, John 45
Brightwell, Daniel Barron 41
Broughton, Leslie Nathan 44
Brown, Charles Brockden 220
Brezina, Otokar 70, 71, 72, 75
Bukowcowa, Zofia 59
Burdová, Olga 68
Burrows, John F. 116
Busa, Roberto 66
Byron, George Gordon 47

Caesarius, Franciscus 58
Cervantes, Miguel de 106
Cicero, Marcus Tullius 15
Cleveland, Charles Dexter 45
Cohnová, Dorrit 111
Cook, Albert Stanburrough 46
Cooper, Lane 45
Cowden Clarke, Mary 42
Cunliffe, Richard John 50, 51
Cvrček, Václav 72, 109

Čapek, Karel 72, 73, 74, 77, 80, 95, 131, 132, 133,
134, 135, 163
Čep, Jan 9, 10, 24, 28, 34, 76, 77, 79, 80, 81, 82, 83,
84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95,
96, 97, 98, 99, 100, 101, 102, 106, 107, 114,
115, 116, 121, 122, 123, 124, 125, 126, 127,
128, 130, 131, 132, 134, 135, 136, 137, 140,
141, 142, 144, 166, 170, 178, 185, 186, 192,
197, 198, 200, 201, 203, 207, 209, 210, 211,
212, 213, 221, 223, 224
Čermák, František 22, 72, 108

Dalimil 74
Data, Krystyna 59
Davis, Norman 51
Dickinson, Emily 47
Dill, Christa 13, 16, 20, 21, 38, 55, 65, 67, 68
Dobrovský, Josef 13
Doležel, Lubomír 102, 103, 104, 105, 215, 216,
218, 221
Dryden, John 40
Dyce, Alexander 48, 49, 50

Eminescu, Mihai 14, 39, 67
Evans, Laura Cooper 44

Fairfax, Edward 50
Ficek, Viktor 68
Franzosi, Roberto 221

Genette, Gérard 221
Gerstenkorn, Alfred 29
Goethe, Johann Wolfgang 25, 32, 33, 34, 35, 39,
48, 49, 52, 53, 56, 62, 63, 64, 65, 67
Gorkij, Maxim 14, 39
Górski, Konrad 55
Graefe, Johanna 52
Gräfe, Manfred 16, 35, 63
Gray, Douglas 51

Hancock, Leslie 66, 67, 68, 69
Hausmann, Franz Josef 19, 76
Hazlitt, William 42

- Hebel, John William 44
 Hellwig, Peter 25, 26
 Hnátková, Milena 72
 Holman, Petr 70, 71, 72
 Homér 13, 38, 45
 Hoover, David L. 217
 Horatius, Quintus Flaccus 45
 Hrabal, Bohumil 72, 73, 74, 77, 80, 95
 Hrabec, Stefan 57

 Chatman, Seymour 111
 Chaucer, Geoffrey 45, 51, 52

 Ingham, Patricia 51

 Jakobson, Roman 116
 James, Henry 106
 Janeczka, Dorota 59
 Janovec, Ladislav 74
 Jelinek, Jaroslav 25, 215
 Jiráková, Míla 68
 Johnson, Samuel 43
 Joyce, James 66, 67
 Juhász, Gyula 14, 39, 67

 Kajtochowa, Krystyna 59
 Karpiński, Franciszek 57
 Kastnerová, Martina 43
 Keats, John 44
 Kennedy, Arthur Garfield 52
 Klemensiewicz, Zenon 57
 Klíma, Ladislav 75
 Klívar, Miroslav 216, 217
 Knapický, Řehoř 58
 Knight, Charles 42
 Kocek, Jan 72
 Kochanowski, Jan 12, 59, 60, 61
 Komenský, Jan Amos 54, 55
 Koneczna, Halina 57
 Koprivová, Marie 72
 Králík, Oldřich 39, 67
 Kratochvíl, Antonín 91
 Křen, Michal 72, 144
 Kubiček, Tomáš 32, 84, 88, 106, 141, 178, 195, 196
 Kucala, Marian 59, 60, 61
 Kučera, Karel 72, 74

 Lankiewicz, Danuta 57
 Levý, Jiří 74
 Littledale, Harold 48, 49, 50
 Lockwood, Laura Emma 45
 Lommatzsch, Ernst 13
 Lotman, Jurij Michajlovič 17, 34
 Lupea, Mihaela 221
 Luther, Martin 38

 Macháčková, Eva 68
 Mácha, Karel Hynek 68
 Machová, Svatava 74
 Mattausch, Josef 14, 15, 16, 17, 18, 19, 20, 21, 31, 33, 37, 38, 39, 47, 48, 53, 61, 63, 64, 65, 75
 Mauersberger, Arno 13
 Meister, Jan Christoph 221
 Merbach, Fritz 52
 Merguet, Hugo 15
 Merker, Erna 52, 54, 55
 Mickiewicz, Adam 12, 14, 39, 55, 56, 57, 59, 67
 Milton, John 38, 45
 Moretti, Franco 9
 Mukařovský, Jan 78

 Nezval, Vítězslav 74
 Norwid, Cyprian Kamil 12, 13, 61, 62
 Novotná, Renata 72

 Onions, Charles Talbot 51

 Palacký, František 13
 Pasek, Jan Chryzostom 14, 57, 58, 59
 Peer, Willie van 78, 217, 218
 Petöfi, Sándor 14, 39, 67, 75
 Petronius 13
 Plutarchos 13
 Pope, Alexander 39, 40, 41
 Popescu, Ioan-Iovitz 121, 221
 Poprawska, Urszula 59
 Preuss, Siegmund 13
 Puchmajer, Antonín Jaroslav 13
 Puškin, Alexandr Sergejevič 14, 39, 67, 75
 Puzynina, Jadwiga 61

 Racine, Jean 47
 Ramsay, Stephen 66
 Renč, Václav 89

- Roelcke, Thorsten 16, 29, 30, 31, 38, 58, 60, 68
 Rösel, Hubert 54
 Rykiel, Barbara 57

 Saudek, Erik A. 53, 65
 Segebade, Johannes 13
 Seifert, Jaroslav 74
 Semino, Elena 221
 Shakespeare, William 14, 42, 43, 48, 50, 51
 Shaw, George Bernard 45, 46, 66
 Short, Mick 221
 Schadewaldt, Wolfgang 32, 33, 62
 Schmidt, Alexander 49, 50
 Schmid, Wolf 117, 118
 Schmiedtová, Věra 72
 Siekierska, Krystyna 57
 Sieradzka, Bożena 59
 Simon, Werner 62
 Smith, Philip Hagar Smith Jr. 46, 47
 Sofoklés 13
 Sokołowska, Teresa 57
 Spielmann, Jacob Reinbold 64
 Stach, Václav 13
 Stelter, Benjamin F. 44
 Stewart, Larry L. 220
 Stluka, Martin 72

 Šafařík, Pavel Josef 13
 Ševčenko, Taras 14, 39, 67, 75
 Šrajetrová, Dominika 72
 Štindlová, Jitka 68, 69, 215
 Šulc, Michal 36, 72

 Tacitus, Publius Cornelius 13, 38
 Tasso, Torquato 50
 Tatar, Diona 221
 Tatlock, John Strong Perry 52
 Tennyson, Alfred 41, 42, 45
 Tešitelová, Marie 25, 38, 74, 107, 108, 215, 216
 Thám, Václav 13
 Thayer, Mary Rebecca 44
 Todd, Henry John 45
 Trávníček, Mojmír 83, 84, 86, 88, 89, 90, 91, 92
 Tureček, Dalibor 98, 99
 Tyl, Zdeněk 54, 75

 Umbach, Horst 32, 33, 34

 Vančura, Vladislav 58
 Vondříčka, Pavel 72

 Waclawičová, Martina 72
 Wallace-Hadrill, Anne 51
 Walsch, Richard 32
 Wiegand, Herbert Ernst 16, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 45
 Wissmann, Wilhelm 62
 Wyttenbach, Daniel Albert 13

 Yeats, William Butler 47

 Zajac, Peter 98
 Zgusta, Ladislav 37

 Żelazko, Kazimierz 57

KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Změlík, Richard

Kvantitativně-korpusová analýza a literární věda : model a realizace autorského korpusu a slovníku

Jana Čepa v kontextu zahraniční a české autorské lexikografie / Richard Změlík. -- 1. vydání. --

Olomouc : Univerzita Palackého v Olomouci, 2015. -- 243 stran

ISBN 978-80-244-4766-7

- 821.162.3-3 * 81'42 * 801.73 * 81'324 * 81'322.2 * 82.0 * 81'374

- Čep, Jan, 1902-1974

- česká próza -- 20. století

- jazyková analýza

- textová analýza

- kvantitativní lingvistika

- korpusová lingvistika

- literární věda

- lexikografie

- monografie

82.0 - Literatura (teorie) [11]

Kvantitativně-korpusová analýza a literární věda

Model a realizace autorského korpusu a slovníku Jana Čepa
v kontextu zahraniční a české autorské lexikografie

Mgr. Richard Změlík, Ph.D.

Publikace vznikla v rámci projektu

Inovace bohemistických studií v mezioborových kontextech (CZ.1.07/2.2.00/28.0178).

Tento projekt je spolufinancován Evropským sociálním fondem
a státním rozpočtem České republiky.

Výkonná redaktorka Agnes Hausknotzová

Odborný redaktor Martin Lukáš

Odpovědná redaktorka Jana Kreiselová

Jazyková korektura Jana Kolářová, Martin Lukáš, Jana Hajdová

Technická redaktorka Jana Hajdová

Autorka obálky Jana Hajdová

Vydala a vytiskla Univerzita Palackého v Olomouci

Křížkovského 8, 771 47 Olomouc

www.vydavatelstvi.upol.cz

www.e-shop.upol.cz

vup@upol.cz

1. vydání

Olomouc 2015

ISBN 978-80-244-4766-7

NEPRODEJNÁ PUBLIKACE

vup 2015/503