



VEDA
vydavateľstvo
Slovenskej
akadémie
vied

Publikácia sa realizovala v rámci projektu Tvorba a rozvoj Slovenského národného korpusu podporovaného Ministerstvom kultúry SR, Ministerstvom školstva, vedy, výskumu a športu SR a Slovenskou akadémiou vied na základe zmluvy č. 0323/2017.

Vydanie publikácie finančne podporila Slovenská akadémia vied na základe odporúčania Edičnej rady SAV.

Recenzenti:

Prof. Mgr. Marián Zouhar, PhD.

Prof. PhDr. Juraj Dolník, DrSc.

MIROSLAV ZUMRÍK

Filozofické aspekty korpusovej lingvistiky



VEDA, vydavateľstvo Slovenskej akadémie vied
Bratislava 2020

Jazykové redaktorky: Mgr. Lucia Miháliková, PhD., Mgr. Kristína Bobeková
Návrh a spracovanie obálky: Mgr. Marek Petržalka
Grafický návrh a zalomenie: Jana Janíková

Vydala a vytlačila VEDA, vydavateľstvo Slovenskej akadémie vied,
Centrum spoločných činností SAV, Dúbravská cesta 5820/9, 841 04 Bratislava,
v roku 2020 ako svoju 4 508. publikáciu.

Počet strán: 80

www.veda.sav.sk

Všetky práva vyhradené.

© Jazykovedný ústav Slovenskej akadémie vied, 2020
© VEDA, vydavateľstvo Slovenskej akadémie vied, 2020

ISBN 978-80-224-1826-3

Obsah

Úvod	6
1. Komplexná skutočnosť	9
2. Vzťah subjektu a jazyka	15
3. Spôsob jestvovania a poznávania jazyka	20
4. Korpusová lingvistika a filozofia technológie	27
5. Rôznorodosť úloh korpusovej lingvistiky	32
6. Korpus ako elektronická vzorka jazyka	36
7. Odhaľovanie skrytej štruktúry dát	41
8. Korpusová lingvistika v celku prístupov k jazyku	47
9. Korpusová lingvistika vo svetle antropologických konštánt	57
Záver	66
Použitá literatúra	74

Úvod

Korpusová lingvistika predstavuje rozvíjajúcu sa oblasť poznávania jazyka. Opiera sa o čoraz rozsiahlejšie elektronické vzorky jazyka, korpusy, a rozvíjajú sa i technologické aj metodologické možnosti, ako tieto vzorky skúmať a spracúvať prostredníctvom korpusových nástrojov. Odhaliť možno napríklad distribučné pomery a iné vzájomné vzťahy medzi jazykovými formami.¹ Kvantitatívne metódy, ktoré sa v korpusovej lingvistike využívajú, umožňujú kvantifikovať kvality jazykových výrazov, textov či štýlov, ktoré sa inak do istej miery odvodzujú z introspekcie jazykovedca. Tu možno spomenúť sériu *Studie z korpusové lingvistiky*, v ktorých sa takto kvantitatívne napríklad vyjadrujú charakteristiky termínov, inými slovami, miera „termínovosti“ textu (*termhood*) (Kováříková, 2017), alebo sa kvantitatívne vyjadrujú charakteristiky prekladovej češtiny, a teda miera jej „prekladovosti“ (Chlumská, 2017). Možno tiež uviesť slovenskú prácu (Garabík – Karčová – Žáková, 2018), v ktorej sa zasa kvantitatívnymi metódami vyjadruje tá vlastnosť polyfunkčného komponentu „sa“, že sa s rozličnými slovesami spája či vyskytuje v rozličnej miere, a tento parameter je v príspevku označený ako takzvaná „savosť“ sloviess (tamže, s. 32).

Rozvoj korpusovej lingvistiky prebieha zároveň s rozvojom príbuzných oblastí: počítačového spracovania prirodzeného jazyka, umelej inteligencie a iných, čoho výsledkom je neustále rastúci objem elektronických jazykových zdrojov i počet nástrojov na skúmanie týchto zdrojov. Preto možno povedať, že sa zvyšuje objem poznatkov o jazyku a v tomto zmysle o jazyku vieme čoraz viac. Ako však uvádza Aristoteles v úvode svojej *Metafyziky*, poznanie v zmysle pravej múdrosti spočíva v odhaľovaní príčin medzi javmi (2015, s. 9). Jaroslav Peregrin, podobne, konštatuje istý rozdiel medzi vedeckým poznávaním jazyka a filozofickým uvažovaním o ňom. Tento rozdiel obrazne vyjadruje tak, že kým empirické skúmanie konkrétneho jazykového materiálu predstavuje obrazný ponor do krajiny, filozofické uvažovanie o zmysle, smerovaní a ďalších súvislostiach empirického skúmania pripomína rozhliadanie sa po tejto krajine.² Zároveň platí, že oba postupy v obraznej krajine sa navzájom dopĺňa-

¹ „Corpus linguistics is the investigation of linguistic research questions that have been framed in terms of the conditional distribution of linguistic phenomena in a linguistic corpus“ (Stefanowitsch, 2018, s. 50).

² „Imagine two researchers both asked to investigate an unknown landscape. One hires a heli-

jú, sú komplementárne.³ Moje doterajšie vnímanie korpusovej lingvistiky je podobné. Na jednej strane existuje a rozrastá sa nesmierne množstvo výskumných otázok, podnetov a zdrojov, rozširuje sa poznanie o jednotlivých aspektoch jednotlivých jazykov, na druhej strane vnímam, že je potrebné uvažovať aj o teoretických rámcoch, v ktorom sa toto rozširovanie poznania odohráva. V práci by som sa preto chcel pokúsiť o usúvzťažnenie rôznorodej oblasti výskumu korpusovej lingvistiky s jej filozofickými aspektmi, s filozofickými myšlienkami, ktoré sa týkajú skutočnosti ako celku, vzájomného vzťahu skutočnosti, mysle a jazyka.

Postupovať budem tak, že najprv identifikujem niektoré otázky, ktoré sa týkajú naznačeného vzťahu medzi svetom, myslením a jazykom, a charakterizujem tento vzťah všeobecnou tézou. Touto tézou je, že realita, ktorá nás obklopuje, je komplexná (kap. 1). Jej komplexnosť priblížim pomocou pojmov z oblasti filozofického uvažovania, konkrétne fenomenologického, i z oblasti vedeckého spracovania dát a zameriam sa na to, ako sa komplexnosť skutočnosti prejavuje vo vzťahu medzi subjektom a jazykom (kap. 2), v otázke, ako jestvuje a ako môžeme spoznávať jazyk (kap. 3) i vo vzťahu k filozofii technológie (kap. 4). Korpusová lingvistika, jej východisková filozofia – empiricizmus – a korpusové nástroje ponúkajú jazykovede rôznorodé možnosti ako skúmať jazyk ako súčasť komplexnej skutočnosti (kap. 5). V ďalšom výklade sa zameriam na tieto možnosti a nástroje, a teda na teoreticko-praktické postupy uplatňované pri korpusovom výskume. Tieto postupy obrazne predstavujú isté „pohyby myslenia“ (Petříček, 2018, s. 149), ktoré vykonáva korpusový lingvista vo vzťahu k dátam. Predbežne som identifikoval tieto pohyby: 1. pohyb od formy k významu jazykových jednotiek, 2. pohyb od jazykových dát k tvrdeniam o týchto dátach (čo sa dá označiť ako pohyb kritického myslenia), 3. pohyb od dát k dátovým štruktúram a pravidelnostiam, vzorom, javom, 4. pohyb od opisu javov k ich vysvetľovaniu, a napokon, 5. filozofický pohyb, teoretický zostup k základom vedy a jednotlivých oblastí poznania. Na základe naznačených pohybov myslenia sa v ďalšom výklade zameriam na reflexiu faktu, že korpus predstavuje elektronickú vzorku jazyka, ktorú tvoria dáta a väčšinou i metadáta (kap. 6). Štruktúry ukryté v týchto dátach je možné odhaľovať štatistickými metódami, pričom odhalené štruktúry sa definujú ako javy a zdroj informácií o týchto javoch (kap. 7). Tieto informácie predstavujú základ ďalšej lingvistickej (kvalitatívnej) interpretácie, uskutočňovanej napríklad prostred-

copter, acquires a birds-eye view of the whole landscape and draws a rough, but comprehensive map. The other takes a camera, a writing pad and various instruments, and walks around, taking pictures and making notes of the kinds of rocks, plants and animals which he finds“ (Peregrin, 2012, s. 1).

³ Tamže.

níctvom explanačných princípov a teoretických prízíem. Vzorky jazyka predstavujú objektívne existujúci predmet vedeckého skúmania. V diskusii medzi jednotlivými prístupmi v lingvistike sa rieši otázka, do akej miery je skúmanie jazykového systému možné prepojiť s ďalšími systémami: fyzikálnymi, biologickými, sociálnymi, psychickými a inými (kap. 8). Diskusia o tejto otázke je pochopiteľná: na jednej strane je zrejmé, že jazykový systém neexistuje vo vzduchoprázdne, na druhej strane sú vzájomné vzťahy jednotlivých systémov veľmi zložité, a preto je z pozície filozofického empiricizmu i kritického racionalizmu potrebné pýtať sa na pôvod, oprávnenie a fundovanosť tvrdení o povahe vzťahu medzi jazykom a inými systémami. Korpusové nástroje a metódy v každom prípade umožňujú vykonávať na vzorke jazykových záznamov operácie, na ktoré má ľudská pozornosť a pamäť obmedzené kapacity. Korpusovú lingvistiku preto interpretujem prostredníctvom takzvaných antropologických konštánt, čo sú filozofické tézy charakterizujúce človeka a aspekty jeho poznávania, a tiež cez prizmu ďalších filozofických téz a pojmov (kap. 9). V závere práce reflektujem aj povahu samotného filozofického uvažovania, pričom takáto seba-reflexia je pre filozofické myslenie pomerne príznačná. Týmto spôsobom by som chcel uchopiť filozofický rozmer korpusovej lingvistiky a podať istú jej filozofickú interpretáciu.

1. Komplexná skutočnosť

Korpusovú lingvistiku a jej postupy pri skúmaní jazyka možno dať do súvislosti s filozofickými a antropologickými otázkami, ktoré sa týkajú vzťahu človeka k jeho skutočnosti (okolitej i vnútornej), k poznávaniu tejto komplexnej skutočnosti, i vo vzťahu k nástrojom, ktoré v procese poznávania človek používa. Medzi takéto antropologické otázky patria: ako sa orientujeme v javovej skutočnosti a vo svete, kedy sa v realite začíname orientovať a kedy sa už orientovať prestávame, a pomocou čoho sa v javovej skutočnosti orientujeme?

Komplexnej skutočnosti človek čelí nielen ako svojmu okoliu, ale aj ako svojmu vnútru, a v prípade, že prijíma viac informácií, než dokáže spracovať, môže nastať situácia, že z kognitívneho hľadiska obrazne narazí na „múr“, ktorým je práve komplexnosť reality. V procese spracúvania informácií sa pritom človeku jazyk bežne javí ako nástroj, ktorým človek dokáže komunikovať, pomenúvať veci, ktoré ho obklopujú alebo naplňujú. Jazyk však funguje tiež ako prostriedok i predmet poznávania a môže sa, napokon, javiť aj ako problém, a to v naznačenom prípade, keď prijímame priveľa (alebo primálo) informácií. V takej situácii narážame na hranice toho, čo o ňom vieme, funkčný vzťah medzi nami ako používateľmi jazyka a jazykom sa narúša alebo sa nevytvorí, napríklad, keď si nevieme spomenúť na nejaké slovo, neovládame nejaký cudzí jazyk, nedokážeme niečo jazykom uchopiť, pomenovať, inými slovami, sám jazyk sa nám javí ako neprehľadne komplexný, komplikovaný, nedostupný, cudzí, nie náš. V tejto súvislosti možno odkázať na koncept xenolingvistiky, v rámci ktorej sa tematizuje práve kategória cudzieho, či už ide o cudzí prvok vo svete, v mysli alebo jazyku (Dolník et al., 2015, s. 13 – 31).

Okrem komplexnosti zmyslových vnemov a podnetov z okolitej reality čelí človek aj komplexnosti poznatkov o tejto realite, a to zvlášť v dobe informačno-komunikačných technológií, ktoré sa vyznačujú explozívnym rastom objemu informácií (Gleick, 2013, s. 15). Skutočnosť je plná podnetov, signálov, ktoré nás môžu až zahltiť, preto potrebujeme istý odstup, aby sme ich dokázali identifikovať, rozlišovať a interpretovať. O poznávaní sveta človekom sa preto dá hovoriť v istých hraniciach. Vzťah poznávania medzi ja a svetom je dynamický, trvá, ak sa prísun a produkcia podnetov a informácií nachádza v týchto hraniciach. Poznávanie sveta prostredníctvom subjektu predstavuje dynamický stav, pri ktorom subjekt spracúva komplexné dáta, či už majú podobu zmyslových podnetov alebo iných informácií. Komplexnosť zmyslových dát,

iných informácií i prítomného šumu spolu vytvárajú „informačné prostredie“ človeka (Svítek, 2013).

Komplexnosť, respektíve „komplikovanosť“ reality v tejto súvislosti tvorí východisko fenomenologicky ladenej práce *Metodická prednosť spleti* (Nitsche, 2016). Komplikovanosť sa v práci charakterizuje ako takzvaná spleť, ako fenomén komplementárne previazaný s oblasťou „obvyklej fenomenality“, kde je možné rozlíšiť objekt a subjekt. To, že ide o východiskový, a teda základný, trvalý, pôvodný aspekt skutočnosti, naznačuje už názov práce, podľa ktorého má spleť v uvažovaní o týchto základných aspektoch „metodickú prednosť“: „Žijeme v komplikovanej dobe. Situácie, do nichž sa dostávame, jsou nepřehledné [...]. V takto komplikovaných situacích se nelze jednoduše přiklonit na jednu stranu, k jedné myšlence, jednomu konceptu či postoji [...]. Ovšem ani pohled na celek různých perspektiv neumožňuje se komplikovanosti vyhnout, neboť její z jednotlivých postojů nelze složit“ (Nitsche, 2016, s. 7 – 8).⁴

Nitsche vo svojej práci navrhuje takzvaný tranzitívne-topologický model fenomenológie a opiera sa o fenomenologické skúmanie pojmu tranzitivity (spätosti spleti a oblasti obvyklej fenomenality) a topológie (určovania pozície subjektu v celku skutočnosti) (tamže, s. 12). Spleťnosť, takzvaná „spleťnosť fenomenality“ je charakteristická pre oblasti, ktoré azda v bežnom pohľade považujeme za oddelené, respektíve za každých okolností oddeliteľné: vonkajškosť, vnútornosť, subjektivita, objektivita. Fenomenologický model navrhnutý v Nitscheho práci ich však považuje za navzájom sa prenikajúce a pôvodne prepojené. Vzájomná a pôvodná spoločná existencia interiority a exteriority, subjektivity a objektivity, ako aj spoločná existencia týchto pólů so spleťou má dôsledky pre základnú pozíciu subjektu v celku skutočnosti. Bežne totiž vychádzame zo skúsenosti, že človek ako jednotliviec interaguje s akýmsi vyčleneným protipólom, či už je to konkrétna vec, problém, druhý človek, pričom interakcia prebieha na pozadí viac-menej zreteľne vyčleneného horizontu, sveta, času. Subjekt, ktorý sa javí ako „prepletený“ s okolím, však dištančnú pozíciu nemôže zaujať vždy, aby sa zorientoval. Komplikovanosť reality sa prejavuje v ťažkosti pri určovaní pozície subjektu vo vzťahu k jeho miestu a okoliu, a teda v ťažkosti pri vytváraní odstupu od predmetne chápanej skutočnosti, pri vymedzovaní hraníc medzi ja a okolím: „Topologičnost vyjadřuje náš zájem o tematiku místa; zvláště o otázku možnosti zakoušejícího lokalizovat se ve vzťahu ke spleťným fenoménům, vůči nimž není možný běžný, distancující průběh lokalizace“ (tamže).

⁴ Tvrdenie, že sa komplikovanosť nedá „poskladať“ z rôznorodých postojů a perspektív, pripomína tézu, že celok je viac než súčet svojich častí. Túto tézu prezentuje M. Svítek v práci s príznačným názvom *Víc než součet částí* (2013), v ktorej vychádza práve z holistického, systémového prístupu k procesu poznania.

S komplikovanosťou skutočnosti sa okrem fenomenológie vysporadúva aj veda: konkrétne v jazykovede sa komplikovanosť (jazykovej) skutočnosti konceptuálne uchopuje napríklad pomocou termínu „komplexnosť (lingvistických) dát“ (*complexity of data*; Moisl, 2015). V tomto prípade ide o komplikovanosť informácií obsiahnutých v jazykových prejavoch a záznamoch, ktoré nás obkolesujú ako možný zdroj poznania. „Nepreniknuteľná“ (*impenetrable*) komplexnosť jazykových dát sa dá redukovať pomocou matematických a štatistických metód, akou je napríklad klastrová (zhluková) analýza. Tieto metódy umožňujú rozdeliť súbor skúmaných dát na segmenty a oblasti, vykazujúce signifikantné pravidelnosti alebo nepravidelnosti a vytvárajúce zhluky dát. Význam klastrovej analýzy a iných analýz a metód v korpusovej lingvistike spočíva v ich využiteľnosti na riešenie problémov, na znižovanie dimenzionality problémov i na „vytváranie hypotéz“ (*hypothesis generation*; tamže) formulovaných na základe evidencie vytvorených zhlukov v potenciálne významných vzájomných vzťahoch. Fakt, že nás obklopujú (a našu skutočnosť vytvárajú) komplexné dáta, tvorí východisko dátovej vedy (*data science*), ako sa to konštatuje v práci *Measuring the Data Universe* (Stahl – Staab, 2018). Autori práce považujú dáta za „základné tehličky“, z ktorých je stvorený („dátový“) svet, je však potrebné docieľiť, aby bol tento dátový svet zrozumiteľný, čo sa dosahuje transformáciou dát, ich takzvanou štandardizáciou.⁵ S odkazom na (aj) túto formuláciu preto vnímam korpusovú lingvistiku ako istú metodologickú a epistemologickú pozíciu, ktorá vychádza z faktu jazykovej (tvarovej a inej javovej) rôznorodosti a komplexnosti, a predstavuje stratégiu, ako túto komplexnosť (potenciálnu, slovami vyššie uvedenej práce, „kliatbu“ (*curse*)) zvládať, alebo slovami inej práce „krotiť“ (*taming the corpus*; Cvrček – Fidler (eds.), 2018), a teda ju spracúvať a získavať z nej informácie.

V súvislosti so stratégiou subjektu, ktorý čelí komplexnej skutočnosti, možno uviesť aj termín „neprijemné problémy“ (*wicked problems*; Ritchey, 2011), s ktorými sa stretávame ako s komplexnými spoločenskými i psychickými javmi. Ich „neprijemnosť“ sa ukazuje pri porovnaní s „krotkými“ (*tame*) problémami, ktoré sa vyznačujú tým, že 1. sú relatívne dobre definované a stabilné (*stable problem statement*); 2. je možné rozhodnúť, či a kedy sa dosiahlo ich riešenie/riešenia (majú takzvaný *definite stopping point*); 3. tieto riešenia možno objektívne zhodnotiť ako správne alebo nesprávne; 4. takéto problémy patria do skupiny podobných problémov (*class of similar problems*), ktoré sú riešiteľné podobným spôsobom; 5. riešenia dobre definovaných problémov

⁵ „The gospel of data analysts starts with the words ‘In the beginning was the data’. The data are the original building blocks, the atoms in our universe and the starting point of our work. They can, at the same time, be our greatest good and our most terrible curse. Because when they do not fit together, they are worthless“ (Stahl – Staab, 2018, s. 12).

možno testovať a možno sa tiež rozhodnúť, že sa od niektorého z riešení odvrátíme (*can be tried and abandoned*) (tamže, s. 19). „Neprijemné“ problémy sú, naopak, charakterizované ako zle definované (*ill-defined*), nejednoznačné (*ambiguous*), subjektívne, pričom často nie je jasné, v čom vlastne spočívajú.⁶ V tomto zmysle si neprijemné problémy nedokážeme pred-staviť, umiestniť pred seba ako pred-met, dobre definovaný problém, ako úlohu. Naše pokusy o riešenie problému sú v takom prípade súčasťou problému, pridávajú mu komplikovanosť, vytvárajú ďalšie problémy.

Rozdiel medzi „krotkými“ a „neprijemnými“ problémami je kvalitatívny. Kým dobre definovaný problém predstavuje konkrétnu úlohu, nedostatočne definovaný problém predstavuje o úlohu navyše: zistiť, čo je vlastne problémom, či ide o jeden alebo viacero problémov, a či ide vôbec o úlohu, ktorú je potrebné riešiť. Je tiež zrejmé, že oba druhy problémov sa navzájom prelínajú a sú pomerne bežné. Obrazne by sme to mohli prirovnáť k situácii, ak by sme mali za úlohu poskladať puzzle, hoci by nebolo isté, či obsahuje dieliky iba jedného, alebo viacerých obrazov, respektíve či sú to vôbec puzzle. Rozdiel medzi problémami tiež možno prirovnáť k situácii, že dostaneme zadanie matematickej úlohy a očakáva sa od nás jej riešenie (čo by samo osebe predstavovalo dobre definovaný problém), zadanie však dostaneme nekompletné (to je neprijemný problém). Bežnú skúsenosť pritom charakterizuje zväčša predstava mechanického, newtonovského poriadku, v rámci ktorého je možné rozlíšiť základnú situáciu, problém, ako aj jeho príčinu a účinok.⁷

Či už ide o dobre alebo zle definované problémy, model vzťahovania sa človeka k skutočnosti, aj tej jazykovej, sa mení, rozširuje, a to vstupom štatistických a iných metód a technológií do života človeka. Rozšírenie ľudského vnímania a poznávania spočíva v tom, že sú v ňom prítomné technológie na reprezentáciu objektov fyzického i virtuálneho sveta, a tiež technológie umožňujúce skúmať tieto objekty. Čoraz viac vecí, oblastí a aspektov ľudského života, či už ide o uspokojovanie fyzických, kultúrnych, intelektuálnych a iných potrieb, sa takpovediac presúva do virtuálneho prostredia a získava virtuálnu podobu, dokumentuje sa a zachytáva, a tiež sa spracúva algoritmami, ktoré v nesmiernych súboroch dát objavujú hlboké súvislosti a ovplyvňujú i reálne

⁶ „They are subjective and strongly stakeholder dependent: there is often little consensus about what the problem actually is, let alone how to resolve it. Above all, WPs won't keep still: they are sets of complex, interacting issues evolving in a dynamic social context. Often, new forms of wicked problems emerge as a result of trying to understand and solve one of them“ (Ritchey, 2011, s. 20).

⁷ „In a world of Newtonian order, where there is a clear relationship between cause and effect, companies can judge what strategies they want to pursue. In a wicked world of complex and shadowy possibilities, enterprises don't know if their strategies are appropriate or what those strategies' consequences might be“ (Camillus, 2008).

životy ľudí. Algoritmy takto hodnotia poskytnuté údaje o ľuďoch, vytvárajú a dotvárajú profily týchto ľudí, predpovedajú ich budúce správanie, a tak generujú signály, ktoré môžu ďalší ľudia prevziať pri rozhodovaní o tých prvých, napríklad pri schvaľovaní hypotéky. Mapujú sa pohyby, cesty, zážitky, predmety, správanie ľudí a život sa čoraz viac odohráva vo virtuálnom prostredí, hoci zároveň ostáva i fyzický.

Komplex načrtnutých dejov predstavuje isté zvrstvenie skutočnosti: dochádza k presunu v pomere fyzickej a virtuálnej reality, k akémusi premiestňovaniu jednak niektorých zdrojov poznania (napríklad pamäti), a jednak niektorých kognitívnych schopností (napríklad pozornosti) do elektronickej a digitálnej podoby, z čoho zasa vyplývajú zmeny v procese poznávania, v jeho zložení a fázach. Tradičnými zdrojmi poznávania sú, napríklad, introspekcia, individuálna skúsenosť, pamäť, ľudia a predmety z nášho okolia, vrátane technických artefaktov, ktoré človeka sprevádzajú počas celých jeho dejín. Náš mozog sa nevyvinul ako funkčne úplne izolovaný systém, predstavuje „sociálny orgán“ (Koukolík, 2008). Pri poznávaní máme okrem toho k dispozícii jednoduché i zložitejšie nástroje. Medzi jednoduchšie patrí pero a papier, prípadne kartotečné lístky, pričom korpusy sa označujú aj ako ich moderná verzia (Čermák, 2017, s. 13 – 24). K dispozícii tiež máme, napríklad, okuliare na čítanie kníh ako artefaktov, ktoré existujú súčasne s tradíciou orálnej komunikácie (Ong, 2006, s. 13 – 24). V súčasnosti však máme k dispozícii i elektronický text, ktorý má v porovnaní s tradičným textom podobné i odlišné vlastnosti. Existujú zaužívané spôsoby, ako v kombinácii skúsenosti, pamäte, interpretačnej schopnosti čítať a analyzovať písaný text a rozumieť hovorenému prejavu. Vychádzame z individuálnej pozornosti a v prípade písaného textu zo zápiskov a výpiskov. Tieto zložky poznávania sa zachovávajú aj pri zapojení elektronických komponentov a zdrojov poznania. Informácie v takto rozšírenom informačnom prostredí spoznáva stále individuálna myseľ, táto však k nim nemá prístup iba prostredníctvom individuálnej pamäte a skúsenosti, viazaných na materiálne nosiče (napríklad knihy, obrázky, múzeá), ale aj prostredníctvom digitálnych nosičov, ktoré sú zabudované do technických artefaktov. Tie majú takzvaný duálny charakter (*dual nature*; Kroes, 2012): predstavujú jednak materiálne predmety, jednak sú definované prostredníctvom svojej funkcie (knihy nám môže slúžiť i ako podložka a na podložku zasa môžeme zapísať dôležitú informáciu či myšlienku). Elektronické zdroje predstavujú artefakt, ktorý možno skúmať prostredníctvom funkcií, zabudovaných v tomto artefakte: prostredníctvom možností korpusového vyhľadávania, matematického modelovania tohto vyhľadávania a prostredníctvom interpretácie výsledkov takéhoto vyhľadávania. Digitálne prostredie sietí a algoritmov sa okrem toho vyznačuje aj do veľkej miery automatizovaným ďalším spracúvaním, analyzovaním dát.

Hoci sa človek naďalej vníma ako „prevažne vizuálna bytosť“ (Haig, 2018), ako bytosť s vtelenou (*embodied*) myslou, a teda ako jednotliviec, jeho vnímanie sa môže oprieť o niektoré ďalšie prvky a vlastnosti vnímaných a skúmaných elektronických artefaktov. Elektronické zdroje sú prístupné prakticky okamžite, sú ihneď k dispozícii (Štícha, 2008), a možno ich vnímať z perspektív, ktoré ľudská pamäť a pozornosť neumožňuje: ide napríklad o vyhľadávanie a analyzovanie veľkého množstva výrazov istého typu v mnohých dokumentoch súčasne, komparatívne, diachrónne alebo podľa zadaných formálnych požiadaviek na typ výrazov. Takéto vyhľadávanie a analyzovanie prekračuje hranice fyzickej ľudskej pamäte. Možno povedať, že kým predtým sme mali kogníciu viazanú na predmety skôr fyzického charakteru (predmety našej skúsenosti, z nášho okolia, knihy, knižnice, ľudí) a interagovali sme skôr s nimi, teraz môžeme interagovať aj s predmetmi iného typu, elektronickými obdobami pamäte, s jej takpovediac rozšíreniami, či už ide o digitálne knihy, knižnice alebo tiež elektronické korpusy. Tieto obdoby tradičných zdrojov možno ako zdroje informácií vnímať a skúmať inými spôsobmi než v prípade tradičných zdrojov. Fyzickú knižnicu zakaždým nedokážeme prehľadať spôsobom, ako je to možné v prípade digitálnych zbierok. Rozdiel medzi týmito vyhľadávaniami spočíva vo vynaloženom čase, ktorý je alebo by bol potrebný na prehľadanie zdroja, ako aj v objeme, ktoré dokáže ľudská myseľ spracovať pri hľadaní napríklad konkrétneho výrazu v jednotlivých dokumentoch. So vstupom virtuálnych reprezentácií fyzických objektov a virtuálnych objektov získavame tiež možnosť rozšírenej manipulácie s pozíciou a perspektívou týchto objektov: tak získavame inak neprístupné, nedisponibilné aspekty predmetov, a predovšetkým ďalšie zdroje informácií o skúmanom predmete, dovtedy nerealizovateľné simultánne perspektívy na skúmané objekty, napríklad vďaka pridanej obraznej „vrstve“ metadát, ktoré sú zdrojom informácií o prvotných dátach. V tejto súvislosti tiež možno uviesť Deleuzov koncept „záhybu“ ako charakteristickej črty barokového myslenia (Deleuze, 2015, s. 9). V dôsledku vrstvenia skutočnosti do deleuzovských záhybov sa nám ponúka viacero zdrojov, odkiaľ môžeme formulovať korelácie i príčinné súvislosti, interpretovať význam a vytvárať nové koncepty.

2. Vzťah subjektu a jazyka

Vychádzajúc z téz egolinguistiky (Dolník, 2008), človek sa ku svetu vzťahuje prostredníctvom seba-projekcií, vytvára si obraz o sebe a iných subjektoch, respektíve dochádza v medziľudskej komunikácii k ich vzájomnej konfrontácii. Človek sa takto vzťahuje k rôznym kognitívnym obsahom v rôznych formách, napríklad ku kultúrnym hodnotám alebo konceptom uloženým v spoločenskom vedomí a sprítomňovaným v individuálnych vedomiach. Ak však vychádzame z tézy, že funkčný vzťah poznávania je dynamický proces, ktorý môže nastať a trvať, a môže sa narušiť, môžeme si tiež položiť otázku, čo robiť v prípade takého prerušenia? Existujú stratégie na zvládanie komplexnosti, ktorá nás nielen obklopuje, ale až prerastá a ochromuje? Ako sa možno správať voči skutočnosti, v ktorej sa nedokážeme zorientovať? Jednou z možností azda je sústrediť sa na vytvorenie akéhosi „malého poriadku“, usporiadania dostupných vecí v našej bezprostrednej blízkosti, kým ten „veľký“ poriadok prerastá naše schopnosti vnímania. Psychologickým efektom takéhoto malého poriadku (napríklad upratanej izby) je pocit, že existuje vymedzený priestor, o ktorom máme prehľad, ktorý môžeme ovládať (upratali sme ju), a teda dokážeme vnímať príčinnú súvislosť medzi naším konaním a dobrým pocitom z jeho výsledku (upratanej izby). Týmto spôsobom sa vytvoril funkčný vzťah medzi subjektom a okolím, takže môžeme povedať, že poznávame svet, že svet existuje ako kozmos (Ajvaz, 2017, s. 19), respektíve ako čosi usporiadané, v čom je možné byť doma, kde je možné „bývať“ (Heidegger, 1993, s. 79 – 103). Ak sa však skutočnosť vyznačuje aj udalosťami v zmysle kvalitatívnych zmien, „udalostí“ (Ajvaz – Pauknerová, 2019, s. 23), dezintegrácie doterajšieho poriadku, ponúka sa aj otázka, či sa takáto stratégia dá uplatniť vždy.

Bežnou predstavou v povedomí človeka je, že svet, jeho obsahy, ako aj myseľ a jej obsahy sú spolu s jazykom nejakým spôsobom k dispozícii vždy, bez ohľadu na čas. Inými slovami, že svet, myseľ a jazyk sú neproblematické a za každých okolností dané rámce, v ktorých sa vyskytuje a deje všetko ostatné: predmety a procesy na svete, myšlienky a pocity v mysli, slová a vety v jazyku. Svet, ja a jazyk skrátka existujú, ja je neprerušným divákom okolitého sveta, jazykových výrazov i vlastných myšlienkových pochodov, psychických stavov. Predpoklad homogenosti času, a teda trvalej existencie ja, sveta i kognície však možno spochybniť, a to napríklad odkazom na vyššie uvedené „neprijemné problémy“ (*wicked problems*) alebo odkazom na termín *konstrukce světa* v práci o vzťahu mozgu a mysle (Horáček et al., 2016, s. 19 – 82): termín

prinajmenšom naznačuje, že svet je čosi, čo sa konštruuje, teda výsledok konštrukcie, a nie vopred daný rámec konania a myslenia. Usúvzťažnenie kognitívnych procesov a vplyvov okolia sa však uskutočňuje aj v prácach o takzvanej ekológii mysle (Bateson, 2019) či ekológii ľudských kognícií (Kusá (ed.), 2011). Obrazne možno povedať, že ak sú kognície čímsi ako „javmi v prostredí“, môže nastať situácia, že v určitých podmienkach prostredia kognície nevzniknú. Takáto kontextualizácia kognície sa napokon naznačuje aj v podnetnej práci N. Katherine Haylesovej o nevedomých základoch vedomých procesov.⁸

Systémy, či už sú jednoduchšie, alebo zložitejšie (samoorganizujúce sa komplexné adaptívne systémy (*complex adaptive systems*)), existujú po istú hranicu. Sú funkčné, pokiaľ nie je narušená ich základná integrita a trvajú podmienky umožňujúce ich existenciu. To naznačuje ich potenciálny zánik, prerušenie kontinuity, stratu identity. Z objektívneho hľadiska i z hľadiska daného subjektu možno zrejme ťažko povedať, čo sa deje v okamihu rozpadu či prerušenia pocitu osobnej identity, vytvárajúcej vzťahom k rôznym kognitívnym i kultúrnym obsahom: takúto svoju ne-kogníciu subjekt nemôže priamo identifikovať, nedokáže si ju (a ani seba) predstaviť. Keďže však ide práve o narušenie kontinuity ja, takýto proces problematizuje hranice ja a ne-ja. Ak, obrazne povedané, myseľ predstavuje „zrkadlo prírody“ (Rorty, 2001) či skôr nástroj na konštruovanie poznania, tak rozbitie zrkadla sa v zrkadle nezobrazí: pokazený nástroj poznanie neskonštruje.

Na jednej strane teda existujú úvahy o kognícii a podmienkach jej možnosti, na druhej strane prebieha skúmanie toho, čo je dostupné skúmaniu. Korpusovú lingvistiku vnímam ako metodologické a epistemologické gesto na tieto dve otázky poznávania (kedy je možné poznávanie, a čo je možné poznať): korpus ako obrazne hmatateľné „telo“ poskytuje skúmajúcemu subjektu výskumný objekt, intersubjektívne platný orientačný bod, a v tomto zmysle mu umožňuje takpovediac neblúdiť, na rozdiel od introspekcie, ktorá je limitovaná kapacitou a mechanizmami individuálnej pamäte. V prípade zložitejších, ťažšie uchopiteľných introspektívnych problémov potom môže nastať situácia podobná tomu, keď niečo hľadáme vo vrecku a nesprávne veci do vrecka priebežne vraciame, takže sa stáva, že vyberáme nesprávny predmet znova a znova, a opakujeme chyby, keďže pamäť a všeobecné vedomie do istej miery podlieha psychickým mechanizmom, ktoré vedomiu samotnému nemusia byť vždy dostupné.

⁸ „Unthought‘ may also be taken to refer to recent discoveries in neuroscience confirming the existence of nonconscious cognitive processes inaccessible to conscious introspection but nevertheless essential for consciousness to function“ (Hayles, 2017, s. 1).

Čo sa týka dostupnosti a disponibility vecí v našom okolí i vnútri, prostredie, v ktorom človek existuje, bolo vstupom technológií, ktoré transformujú pôvodnú podobu reči do podoby elektronického textu, skúmateľnej vzorky jazyka, obohatené o nový prvok. Korpus ako elektronická vzorka jazyka potom predstavuje disponibilný artefakt, pôvodné prostredie si však prinajmenšom sčasti ponecháva ne-disponibilný, jedinečný charakter. Na toto prostredie mimo objektívnej reality, ktorá je doménou logu, identifikovateľných javov a súvislostí, nie je zrejme vždy možné uplatniť spôsoby poznávania uplatňovaných na svet objektívnej reality, pozostávajúcej z dát, javov, štruktúr. Vzhľadom na to, že je jazyk vo svojej autentickej podobe, v jednotlivých realizáciách, nesmierne variabilný a rôznorodý, a tak často možno vedľa seba nájsť jednotlivý príklad i protipríklad nejakého jazykového javu, zdrojom ďalšieho poznania sa stáva napríklad skúmanie distribučných pomerov týchto jednotlivých príkladov. Všeobecne povedané, ak existuje jazykový jav A (napríklad univerbizácia) a zároveň existuje nie-A, nemôžeme povedať, že pravidlom je A, ani nie-A, a je teda potrebné ukázať, v akom rozsahu a za akých okolností existuje jav A, a za akých (v pomere k A) nastáva jav nie-A. Treba zároveň zdôrazniť, že opis javov na úrovni distribúcie jednotiek je iba jedným z krokov pri ďalšom skúmaní jazyka, ktoré postupuje cez „prizmu teoretických rámcov“ (Jarošová, 2018b) k otázke významu jazykových štruktúr, ich funkcií, k princípom, či k ďalším systémovým súvislostiam javov.

Vytvorením disponibilnej vzorky jazyka sa istá jeho podoba stala čímśi objektívne jestvujúcim, vonkajším, materializovaným, hoci zároveň digitalizovaným predmetom. Je k dispozícii a obsahuje skryté štruktúry, je zdrojom informácií, poznatkov a poznania. Napriek tomu ide stále o „ostrovček“ disponibility, ktorý existuje zároveň s doménou pôvodnej skutočnosti, ktorá je do veľkej miery chaotická, uniká predstavivosti, akoby odopierala subjektu uchopenie. Vytvorením disponibilnej vzorky jazyka sa, inými slovami, nestatila pôvodná skutočnosť, ktorá nie je disponibilná vždy, ako o tom svedčia napríklad výpadky pamäte či prejavy afázie⁹, schizofrénie a iné prerušenia kontinuity ja.

Narušenie elementárnej kontinuity subjektu sa v tejto súvislosti v neurovedách označuje ako narušenie minimálneho ja (*disturbance of minimal self*). Minimálne ja, alebo takzvaná ipseita (*ipseity*), predstavuje otázku, ktorú skúma medicína, psychológia, neurofyziológia a iné vedy z klinického i fenomenologického hľadiska (Sass, 2013; Nelson et al., 2014). Štruktúra označovaná ako minimálne ja sa pritom charakterizuje ako implicitná kvalita vedomia vlastného ja, ktoré vďaka tejto kvalite vníma všetky prebiehajúce skúsenosti

⁹ Problematiku používania jazyka pri diagnostikovanej afázii skúma napríklad Flanderková (2019).

ako „svoje“.¹⁰ Ipseita ako elementárna identita (bytie „tým istým“ vo vzťahu k „tomu istému“) sa do veľkej miery viaže na schopnosť individua vnímať, vytvoriť si a udržať funkčné prepojenie medzi oblasťami vnímania, myslenia a reči. Funkčnosť prepojenia medzi týmito oblasťami je ohrozená narastajúcou komplexnosťou vonkajších či vnútorných podnetov, a teda možno povedať, že funkčnosť prepojenia trvá, kým komplexnosť podnetov, ktoré musí ja spracúvať, ostáva v udržateľnej miere. Narušenie ipseity možno potom obrazne vyjadriť tak, že ide o situáciu, ako keby vyhodil loptu do vzduchu jeden, ale chytil ju druhý, respektíve nie-prvý človek (nie „ten istý“), a nechtyl by loptu, ale nie-loptu (nie „to isté“), respektíve nie-chtyl (neurobil „to isté“) atď.

Takto načrtnutá oblasť javov, ktorých príkladom je subjektívna skúsenosť osobnej identity (a jej rozpadu), sa líši od poznania a ontológie, ako ho modeluje logika a filozofická ontológia, a líši sa tiež od pojmu poznatkov (znalostí) a ontológie, s ktorým sa stretávame, napríklad, v oblasti umelej inteligencie, počítačového spracovania prirodzeného jazyka či v informačnej vede.¹¹ Aj v prípade merania takzvanej sémantickej podobnosti slov, jednej z oblastí počítačového spracovania prirodzeného jazyka a umelej inteligencie, sa však vychádza zo základného mechanizmu kognície, z odhaľovania podobností medzi rôznorodými objektmi a stimulmi. Narušenie ipseity preto môžeme opísať ako druh udalosti, kedy sa odhalenie týchto podobností neuskutoční.¹²

Z filozofického hľadiska, konkrétne z hľadiska genetickej fenomenológie (Ajvaz, 2012) je možné uvažovať aj o procesoch, ktorými sa vôbec konštituujú objekty pre subjekt. Svet a poriadok z hľadiska tejto fenomenologickej reflexie existuje iba vtedy, keď je možné objaviť jednotlivé objekty a podnety, inak panuje chaos pred objavením sa poriadku. Proces, ktorý umožňuje emergenciu sveta, jeho poriadku, a teda umožňujúci skúsenosť a sám poriadok vzniku, zanikania a premien javov, procesov a času, sa vo fenomenologickej literatúre označuje ako prvotné založenie (*Urstiftung*), čo je termín pochádzajúci v rôznych obmenách od Edmunda Husserla.¹³

¹⁰ Minimálne ja sa opisuje napríklad ako „prereflective, tacit level of selfhood. It refers to the implicit first-person quality of consciousness, ie, the implicit awareness that all experience articulates itself in first person perspective as ‘my’ experience“ (Nelson et al., 2014, s. 479).

¹¹ „The sense of ‘knowledge’ used in information science is more relaxed. Terms such as ‘knowledge engineering’ and ‘knowledge management’ do not refer to knowledge in the sense of a body of beliefs that are apodictically true, but of a body of beliefs which the scientific community has good reason to believe are true and thus treats in every respect as if they are true“ (Munn – Smith, 2008, s. 8).

¹² „The capacity of assessing the similarity of objects or stimuli has long been characterised as a central component for establishing numerous cognitive processes.“ (Harispe et al., 2015, s. 115).

¹³ „The *Urstiftung* is in fact a sort of act of creation (as an institution of sense), which opens up new directions for the future by configuring all experience to come. To sum up, it is a concept

Svet javov, ktoré je možné usporadúvať a objavovať v nich štruktúry a zákonitosti, teda z hľadiska fenomenológie neexistuje za každých okolností, o čom zasa z hľadiska psychiatrie či neurovied svedčia podoby afázie, schizofrénie a ďalších porúch. Vytvorenie či „vynález“ objektívnej reality (Markoš, 2015), ktorá je disponibilná, objektívne jestvujúca, naopak, ožíva v procese intersubjektívnej komunikácie. Na jednej strane objektívne existuje disponibilný artefakt skúmaný vo vede a na druhej strane realita, ktorá sa na takýto artefakt redukovať nedá. Akým spôsobom vo vzťahu k týmto dvom pólom jestvuje jazyk?

that can describe both dimensions: that which can be called pre-given and makes up the motivating condition for something to emerge, as well as the active moment, namely that of the active participation that 'institutes' sense in the present and for the future“ (Niel, 2017, s. 140).

3. Spôsob jestvovania a poznávania jazyka

Ján Horecký v článku *Spôsob jestvovania jazyka* konštatuje:

„Pri úvahách o jestvovaní jazyka v reči i v mozgu sa [...] dosť jasne nerozlišuje materiálna a ideálna stránka. Kým o zvukovej stránke možno jednoznačne povedať, že hlásky materiálne jestvujú skutočne len pri ich artikulovaní, jazykové výrazy ako celky jestvujú v jazykovom spoločenstve ako ideálne objekty, produkty spoločnej práce. Z faktu, že pri jazykovom dorozumívaní účastník komunikácie netvorí jazykové jednotky (slová, spojenia slov) pri každom jednotlivom komunikačnom akte (v tom zmysle, ako tvorí, konštruuje podľa istého zámeru vety i celkový prejav), ale berie hotové pomenovacie jednotky, možno vyvodzovať, že jazyk (langue) ako systém výrazov skutočne tvorí akúsi veľkú zásobáreň, do ktorej hovoriaci siaha pri konštruovaní prejavu. Hotové pomenovacie jednotky – jazykové znaky – teda jestvujú vo vedomí daného jazykového spoločenstva“ (1979, s. 100).

Horecký takto identifikuje myšlienkovú činnosť, ktorú označuje „úvahy o jestvovaní jazyka“, čo samo osebe naznačuje istý filozofický rozmer, ontologický rozmer jazyka. V citáte sa tiež konštatuje, že jazyk jestvuje prinajmenšom v dvojakom zmysle, v reči a v mozgu. Identifikuje sa i ďalší dvojaký charakter jazyka, jeho dve stránky: materiálna a ideálna, a tiež sa s odkazom na stav dobovej diskusie o tejto otázke konštatuje, že sa obe stránky nerozlišujú dostatočne jasne. Materiálnu stránku jazyka Horecký prezentuje ako to jestvovanie jazyka, ktoré v podobe hlások trvá počas ich artikulácie; jazykové výrazy zasa jestvujú ako „ideálne objekty“ v takzvanom jazykovom spoločenstve – tieto ideálne objekty tvoria systém, ktorý Horecký pripodobňuje k „zásobárni, do ktorej hovoriaci siaha pri konštruovaní prejavu“. Systém ideálnych objektov zasa jestvuje vo vedomí jazykového spoločenstva, spoločenskom vedomí: „jazykový znak a teda aj jazyk ako systém znakov jestvuje v spoločenskom vedomí daného jazykového spoločenstva“ (tamže).

Horecký tak formuluje filozofické tvrdenia o povahe jazyka, o jeho stránkach (materiálnej a ideálnej), a o vzťahu jazyka (jeho stránok) k používateľovi jazyka, k hovoriacemu. Podáva tiež prinajmenšom čiastočnú odpoveď na otázku o spôsobe jestvovania jazyka, ktorú je podľa Horeckého potrebné „rozložiť na dve otázky: kde jestvuje jazyk a ako jestvuje jazyk“ (tamže, s. 97). Čo sa týka miesta jestvovania jazyka, jazyk ako systém podľa Horeckého jestvuje v spoločenskom vedomí. Termín spoločenské vedomie¹⁴, respektíve existencia jazyka

¹⁴ Tento termín, ako ma upozornil recenzent, prof. Zouhar, sa v slovenskej jazykovede druhej polovice dvadsiateho storočia zaužíval prevzatím z marxistickej filozofie.

v spoločenskom vedomí predstavujú nepochybne zaujímavý námet na úvahy o ontologickom, metafyzickom štatúte tohto vedomia, v ďalšej úvahe sa však zameriam na perspektívu individuálnej mysle hovoriaceho, ktorá sa v uvedenej schéme prezentuje ako východisko ďalších autorových úvah. Čo sa týka spôsobu jestvovania jazyka, hovoriaci podľa Horeckého siaha po jazykových výrazoch, ktoré existujú ideálne v spoločenskom vedomí, a tak konštruuje svoj prejav. Pomenúvacie jednotky, jazykové znaky (výrazy), existujú vo vedomí, zároveň ale majú svoju materiálnu podobu (ako hlásky, slová, slovné spojenia, vety, texty). Používateľ jazyka pri tvorbe prejavu používa jazykové výrazy, ktoré existujú v spoločenskom vedomí, ale sú zároveň uložené v pamäti, vedomí hovoriaceho.

Akúsi vstupnú bránu k spoločenskému vedomiu predstavuje myseľ jednotlivca, ktorý je situovaný do biologicky, spoločensky, kultúrne i historicky podmienenej reality. Človek ako používateľ jazyka a „prírodno-sociokultúrna bytosť“ (Dolník, 2010b, s. 17) je priesečníkom skúmania lingvistiky, neurovied, fyziológie, biológie, kognitívnych vied, sociológie, psychológie i ďalších vied, a je tiež predmetom filozofických úvah, napríklad o povahe a vzájomných vzťahoch pojmov ako myseľ, vedomie, ja. Ako príklad takejto filozofickej úvahy o povahe mysle uvádzam rozsahovo menšiu prácu *Proměna vnímání a zkušenost pravdy* od M. Merleau-Pontyho (2017), kde sa konštatuje prinajmenšom dvojaký pohľad na poznávanie človeka.

Prvým pohľadom je filozofická perspektíva vychádzajúca z autonómnosti ľudského subjektu. Vychádzajúc z tejto perspektívy „*se člověk sám sobě jeví jako zcela svobodný. Za pravdivé uznává pouze to, o čem může mít vědomí*“ (tamže, s. 14 – 15). Druhá je perspektíva vied o človeku, ktoré subjekt vnímajú ako rôzne ovplyvňovanú, až determinovanú bytosť. Táto perspektíva vzniká vďaka vedám a ich poznatkom (tamže, s. 15). Merleau-Ponty konštatuje, že hoci vývoj vied o človeku postupuje a prináša čoraz viac poznatkov o jednotlivých oblastiach objektívnej reality, ľudské poznanie a filozofia sa ocitajú v stave krízy: „*Vyvstává nesoulad mezi tím, co o sobě člověk může sám zjistit pomocí reflexe či vědomí, a tím, co se o sobě dozvídá, když objasňuje souvislost mezi svým chováním a vnějšími podmínkami, na kterých jeho chování zjevně závisí*“ (tamže, s. 14). Obidve perspektívy sú však komplementárne, nemožno sa úplne vzdať ani jednej. Ak by sme, na jednej strane, odmietli filozofickú perspektívu, a teda „*odmítli svědectví vědomí, připravili bychom subjekt o veškeré jistoty, dokonce i ty, které by mohl hledat ve věděni o člověku vypracovaném zevnějšku*“ (tamže, s. 16). Na druhej strane sa nemožno vzdať ani pozitívnych vied, pretože „*je zjevné, že jen čistě pomocí reflexe nemůžeme sami sebe poznat vyčerpávajícím způsobem. Sami sobě rozumíme mnohem více díky vnějšímu věděni např. o minulosti, etnografii, psychopatologii než díky přímému vyjasňování vlastního života*“ (tamže). Filozofii potom ostáva úloha túto komplementárnosť a spo-

ločnú existenciu empirického poznania a filozofického vedomia pochopiť tak, aby sa obe perspektívy (empirické/vedecké a filozofické) mohli dopĺňať a neexistovali iba izolovane, ako *„bohaté, avšak slepé empirické vedomí a filozofické vedomí, jež správně uchopuje podstatný rys člověka, avšak není si vědomo faktu, že se zrodilo“* (tamže, s. 16 – 17).

Podľa Merleau-Pontyho teda ja poznávame a reflektujeme v dvojacom zmysle: na jednej strane v zmysle podmieneného subjektu, subjektu prírodných a spoločenských vied, na druhej strane „slobodné“ ja vo filozofickom zmysle. Vzťah medzi slobodným „filozofickým“ a „neslobodným“, podmieneným, vedecky skúmaným ja je pritom tradičným predmetom filozofickej diskusie. Analytická filozofia, s istou dávkou zjednodušenia, požaduje, aby sa filozofia zaoberala najmä skúmaním významov ako abstraktných a objektívnych entít, napríklad v podobe konceptuálnej analýzy (Zouhar, 2016), čo okrem iného znamená spresňovanie a definovanie (explikáciu) významov používaných výrazov, objasňovanie ich vzájomných vzťahov; čo zároveň značí, že treba odhliadať od psychologických a iných aspektov vedomia, ktoré skúmajú prírodné a spoločenské vedy. Redukciu významov a pojmov na psychické entity a subjektivizmus v logike však odmieta aj husserlovská fenomenológia, ktorá požaduje, aby sa filozofia zaoberala takzvaným eidetickým skúmaním, skúmaním „podstát“, čím by sa odlišovala od empirického skúmania faktov (Husserl, 2014), čo zasa možno považovať za istú podobu platonizmu.

Rôznorodá podmienenosť ľudského subjektu tomuto subjektu stanovuje jeho sociálne, biologické či psychologické a kognitívne limity. Možno tiež povedať, že ľudská kognícia je súčasťou prostredia, je týmto prostredím ovplyvňovaná a spoluutváraná, a teda možno hovoriť o „ekológii ľudských kognícií“. Tak sa ukazuje, že domény ako inteligencia, jazyk, empatia či morálka sú *„spätě s funkčností kognícií v prostrediach, ktoré sú pre ľudí ako živočíšny druh príznačné, a pritom sú ukotvené v evolučných podmienkach, vďaka ktorým sa nadštandardný, druhovo špecifický kognitívny kapitál rozvinul“* (Kusá (ed.), 2011, s. 8). „Ekologická“ podmienenosť subjektu sa týka napríklad danoj pamätevej kapacity, ktorá nie je nekonečná. Rozsah a účinnosť pamäte sa mení v závislosti od momentálnej i celkovej dispozície a kapacity jednotlivca. V dôsledku práve takejto situačnej podmienenosti jednotlivca môže nastať situácia, že používateľ jazyka nemá niektoré jazykové výrazy pri zamýšľanom konštruovaní prejavu k dispozícii, a to v rozličnej miere a rozličnom trvaní. V takom prípade si jednotlivec môže pomôcť výrazmi, ktoré k dispozícii má alebo môže to, čo chce povedať, vyjadriť inými výrazovými formami: zvukmi, gestami, pohybmi.

Ak budeme ľudské poznávanie okolitého sveta prostredníctvom kognitívnych rámcov a modelov, ľudské orientovanie sa v svítekovskom „informačnom prostredí“ považovať za podobu toho, ako prostredníctvom poznania uchopiť

istú časť sveta, možno si položiť aj otázku, „či možno jazyk poznať a „uchopiť“ v jeho úplnosti“ (Findra, 1996, s. 9). Odpoveď na túto otázku je podľa Findru za predpokladu, že „sa na jazyk vo všeobecnosti pozeráme z hľadiska nášho individuálneho poznania“, záporná: „Najmä slovnú zásobu nemôže žiaden jednotlivec poznať v jej komplexnosti, pretože by musel mať uložené v pamäti státisíce slov“ (tamže). Týmto konštatovaním Findra nadväzuje na ontologickú otázku o spôsobe jestvovania jazyka, ktorú sformuloval Horecký, rozširuje ju však aj o epistemologickú otázku poznateľnosti jazyka:

„Kde žije jazyk ako systém a ako ho môžeme poznať? Vo svojej ideálnej podobe, vo svojej ‚komplexnosti‘ je jazyk uložený vo vedomí národného spoločenstva, ktorému slúži ako komunikačný prostriedok, a teda aj vo vedomí každého príslušníka tohto spoločenstva. Vedomie, pamäť jednotlivca ho však v tejto komplexnosti nemôže nikdy obsiahnuť. V skutočnosti v jeho materiálnej podobe jazyk spoznávame v texte. Presnejšie: odráža sa, ‚žije‘ vo všetkých textoch, ktoré boli vytvorené v tomto jazyku. Text je konkrétna, materiálna celosť, v ňom sa nám aj jazyk javí ako fyzická kategória, ktorú možno vnímať zmyslami. V ňom spoznávame lexiku i gramatiku jazyka. Takto, prostredníctvom textu rekonštruujeme aj ideálnu podobu jazyka, jazyk ako abstraktný systém“ (tamže).

Aj podľa Findru teda jazyk ako systém jestvuje vo svojej ideálnej podobe a komplexnosti v spoločenskom vedomí (vedomí národného spoločenstva), a ako taký je jednotlivcovi k dispozícii. Jednotlivec podmienený vlastným vedomím a pamäťou však túto komplexnosť nemôže obsiahnuť. Jazyk ako systém jestvujúci v spoločenskom vedomí má, podobne ako podľa Horeckého, svoju materiálnu stránku (podobu), ktorou je text. Texty vytvorené v danom jazyku predstavujú obrazné zrkadlo jazyka ako systému, keďže sa v nich tento systém „odráža“, „žije“ v nich, je podľa nich „rekonštruovaný“. Otázku, ako je možné poznať jazyk v jeho úplnosti, tak možno upraviť na otázku, ako je možné poznávať jeho materiálnu podobu, text, respektíve texty, v ich rôznorodosti, autentickej podobe a čo najväčšej úplnosti. Metaforická charakteristika hovoriaceho, ktorý podľa Horeckého „berie hotové pomenovacie jednotky“ z „jazyka ako systému výrazov“, potom aj u Findru predstavuje istú základnú schému vzťahu medzi hovoriacim ako jednotlivcom, používateľom jazyka, a jazykom. Jazyk v tejto schéme existuje jednak vo svojej materiálnej podobe ako reťazec zvukov, hlások, znakov zaznamenaných na papieri a prostredníctvom iných médií, a teda ako materiálny artefakt; a jednak existuje vo svojej ideálnej podobe, kde sa k nemu pristupuje ako k systému. U Horeckého existuje na jednej strane jednotlivec, používateľ jazyka, a možno predpokladať, že si jazykové výrazy berie zo zásobárne spoločenského vedomia bezprostredne, vychádzajúc najmä zo svojich individuálnych poznatkov, skúseností a pamäte. Konštrukcia prejavu sa však, na druhej strane, zrejme opiera aj o ďalšie ľudské

i knižné zdroje, v prípade tvorby odborných textov sú to, napríklad, odkazy na citované pramene; umelecké i iné texty vznikajú často v niekoľkých cykloch v spolupráci autorov, prekladateľov, redaktorov, korektorov.

Predstava bezprostredného „siahania“ do zásobárne ideálnych objektov pri konštruovaní jazykových prejavov a pri spoznávaní jazyka sa tiež rozširuje o prítomnosť a pôsobenie rozličných artefaktov a technológií. Tieto procesy a predmety tvoria súčasť prírodno-kultúrneho prostredia človeka a predstavujú i prostriedky, ako s týmto prostredím vstupovať do rôznych interakcií, napríklad pri tvorbe jazykových prejavov. Na tvorbu písomných prejavov sa používajú materiálne nástroje ako pero a papier, písací stroj, ktorých výsledkom je materiálny text, ale aj nástroje, ako sú počítače, ktoré umožňujú tvorbu a spracovanie i skúmanie elektronických textov. V prípade digitálnych nástrojov sa tiež zdôrazňuje čoraz väčšia prepojenosť ich materiálnej a nemateriálnej zložky. Paramita Ghosh (2019) konštatuje, že jedným z trendov v súčasnom rozvoji dátových vied je práve „digitálna zmes“ (*digital mesh*), respektíve „miešanie fyzického a digitálneho sveta“ (*the meshing of the physical and digital world*). Zložený, fyzicko-digitálny charakter informačného prostredia možno označiť za analógiu Nitscheho pojmu „spleť“.

Ďalšie rozvrstvenie perspektív na skutočnosť vyplýva z toho, že technologické artefakty ľudia dokážu vnímať cez prizmu ich funkčnosti alebo aj bez nej: hudobný nástroj sa dá vnímať ako drevený či plechový predmet, zvykne sa však predovšetkým vnímať a používať ako prostriedok na produkciu hudby. Ak sa na hudobný nástroj budeme pozeráť iba ako na fyzický predmet, v modeli vzťahu medzi človekom a jeho okolím zdôrazníme estetický vzťah a proces vnímania. Ak budeme na tento predmet nazerať cez prizmu funkčnosti, v modeli zdôrazníme praktický, pragmatický vzťah, konanie. Tieto vzťahy a prizmy sa rozširujú tak, že sa na daný predmet budeme pozeráť ako na predmet poznania, čím v modeli vzťahu medzi človekom a okolím zdôrazníme teoretický vzťah, proces poznávania. Treba však zároveň povedať, že uvedené tri vzťahy, tri perspektívy, sú prepojené. Technologické artefakty, vrátane digitálnych zdrojov a nástrojov na skúmanie a spracovanie prirodzeného jazyka, korpusov, predstavujú potom prostriedok i ako predmet poznávania. Korpusy ako elektronické vzorky autentických jazykových prejavov rozširujú pôvodný vzťah medzi používateľom jazyka (respektíve jazykovedcom) a jazykom o nový prvok. Predmetom korpusového skúmania sú tieto elektronické záznamy a vzorky jazyka, pričom cieľom korpusového výskumu je opísať a interpretovať ich rôznorodosť, a prostredníctvom nich odvodzovať zákonitosti platné pre systém jazyka, a tiež opísať pravidlá, podľa ktorých sa v komunikácii vyskytujú a ktoré Horecký v článku v úvode označuje ako gramatické pravidlá¹⁵:

¹⁵ Používa sa i termín „gramatikografia“ (*Grammatikographie*; Konopka, 2018).

„V tomto vedomí, podobne ako v prípade hier, jestvujú aj pravidlá pre spájanie jazykových znakov (gramatika)“ (1979, s. 100).

Možno zhrnúť, že Ján Horecký reflektuje spôsob jestvovania jazyka, a teda z filozofického hľadiska nastoľuje ontologickú otázku: akú povahu má jazyk? Odpoveďou je prinajmenšom dvojaké bytie jazyka ako systému, jeho materiálnej a ideálnej stránky. Ontologická otázka sa rozkladá na dve ďalšie otázky, kde a ako jestvuje jazyk, ktoré smerujú k tomu, že jazyk ako systém jestvuje v takzvanom spoločenskom vedomí, ktoré môžeme chápať aj ako metaforickú i metafyzickú doménu ideálnych entít. Otázku, kde a ako jestvuje jazyk, si kladie aj Ján Findra, ktorý pridáva otázku, ako možno jazyk poznať, respektíve ako je ho možné poznať v úplnosti, čo je zasa z filozofického hľadiska epistemologická otázka: ako je možné poznať jazyk. Naznačuje sa tiež, že individuálna myseľ má k jazyku v zásade prístup pri konštruovaní prejavu, nemá ho však k dispozícii v jeho úplnosti, a to vzhľadom na biologické, psychologické i iné limity a možnosti jeho individuálnej pamäte, pozornosti a skúsenosti. Zároveň sa naznačuje, že človek disponuje materiálnoú podobou jazyka, textami a inými jazykovými zdrojmi, ktoré už možné skúmať je. Texty a ďalšie zdroje predstavujú artefakty, ktoré sa ako materiálne objekty stávajú predmetom skúmania a tiež prostriedkom skúmania jazyka, ktorého sú zrkadlom. To znamená, že sa kapacita ľudskej pamäte a kognitívnych schopností sprostredkovane zväčšuje prostredníctvom technických artefaktov. Takýmito artefaktmi sú aj jazykové korpuse ako elektronické zdroje, rozsiahle vzorky jazyka, čím sa okrem Horeckého ontologickej a Findrovej epistemologickej otázky vyslovuje aj tretia, technologická a všeobecne praktická či metodologická otázka vo vzťahu k jazyku: pomocou čoho možno jazyk poznať? To, že artefakty sú zároveň predmetom i prostriedkom skúmania, platí aj v prípade korpusovej lingvistiky: korpuse sú jednak predmetom skúmania a technologického vývoja, ale najmä prostriedkom poznávania jazyka. Korpuse, napokon, predstavujú aj predmet filozofického myslenia, ktoré Emil Višňovský porovnáva s vedeckým poznávaním: „Kým poslaním vedy je poznanie, filozofia je slobodné tvorivé (konštruktívne) myslenie, presnejšie ‚pojmové myslenie‘, t. j. myslenie o svete v špecificky filozofických, relatívne abstraktných pojmoch“ (s. 91 – 92). Filozofické uvažovanie podľa tejto charakteristiky neprináša poznanie v zmysle vedecky dokázateľných tvrdení na empirickej báze, čo však neznamená, že sa takéto myslenie nemôže výsledkov a konceptov vedeckého skúmania pridržať. Úlohou filozofického myslenia je „nachádzať aspekty, prístupy (uhly pohľadu) a súvislosti, ktoré by sme inak nevideli a ktoré sú relevantné“ (tamže, s. 93) a jeho výsledkom je sieť filozofických pojmov (takzvaný „filozofický slovník“), v ktorých myslíme o zvolenom predmete (tamže, s. 91 – 92).

Kedže jazykové korpuse ako elektronicky spracované vzorky autentických jazykových prejavov predstavujú istý typ technických artefaktov, filozofické

myslenie o tejto téme by malo zohľadniť filozofické aspekty technológie. Dôležitou charakteristikou korpusu v porovnaní s limitami individuálnej mysle však ostáva, že takáto elektronická vzorka jazyka je na rozdiel od momentálnej dispozície individuálnej ľudskej mysle k dispozícii kedykoľvek a vo svojej úplnosti (Štícha, 2008, s. 12). Technológia vo svojich rôznych podobách tak predstavuje spôsob, ako môže človek premieňať zložky svojho sveta na disponibilné artefakty.

4. Korpusová lingvistika a filozofia technológie

Filozofi technológie sa podľa *Stanfordovej encyklopédie filozofie* usilujú reflektovať procesy označované ako tvorbu artefaktov i povahu, filozofickú podstatu týchto artefaktov (Franssen et al., 2018). V stredoeurópskom priestore sa otázkam filozofie technológie výrazne venuje napríklad Mark Coeckelbergh, pričom sa zameriava na výklad učení M. Heideggera, J. Searla, D. Ihdeho, B. Latoura, ale do veľkej miery na výklad L. Wittgensteina, ktorého Coeckelbergh výslovne označuje za „filozofa technológie“ (Coeckelbergh – Funk, 2018; Coeckelbergh et al., 2018). Orientuje sa tiež na etické a iné aspekty technologického rozvoja, najmä na aspekty informačno-komunikačných technológií, kybernetiky, bioniky a robotiky. Inšpirácia Wittgensteinovými myšlienkami sa u Coeckelbergha premietla napríklad do názvu práce *Using Words and Things* (2017), čo pripomína wittgensteinovskú koncepciu významu ako spôsobu používania slov v jazykových hrách: názov naznačuje analógiu medzi používaním slov v komunikácii a používaním vecí ako funkčných artefaktov v životnej praxi. Rozšírenie ontologickej a epistemologickej otázky o technologický aspekt používania a skúmania jazyka možno podložiť i tým, že technologické artefakty tvoria takpovediac organickú súčasť ľudskej spoločnosti, a tak je podľa Coeckelbergha legitímne filozoficky reflektovať aj túto oblasť ľudského života. Technologický vývoj, ktorý priniesol aj možnosť skúmania jazyka prostredníctvom elektronických zdrojov (korpusov), sa neuskutočňuje vo vákuu, ale je, naopak, úzko spätý s komplexnými procesmi v ľudskej spoločnosti. O takúto reflexiu sa práve pokúša filozofia technológie.¹⁶

Čo sa týka využitia technológií pri skúmaní jazyka, korpusová lingvistika v tejto súvislosti historicky viac-menej narážala na istú nedôveru, pôvodne vychádzajúcu od zástancov teoretickej lingvistiky, generatívnej gramatiky a formalistického prístupu k jazyku, ktorí neboli presvedčení o praktickej využiteľnosti korpusov na lingvistický výskum.¹⁷ Podľa zástupcov generatívnej

¹⁶ „Cutting-edge technological developments do not arise in a vacuum but are emergent phenomena in highly complex societal conditions. Philosophers of technology scrutinize that interplay, of technology and society, as it unfolds every day – and in everyday life“ (Coeckelbergh et al., 2018, s. 287).

¹⁷ „In much of the formalist linguistic literature, corpus linguistics is routinely attacked, ridiculed, and dismissed not just as having no practical use, but as having no conceivable use at all in linguistic inquiry“ (Stefanowitsch, 2018, s. 5).

gramatiky sa má, naopak, pozornosť lingvistov zameriavať skôr na štruktúry a systém jazykovej kompetencie než na výstupy tohto systému, na jazykovú performanciu. Tento skeptický postoj generatívnych lingvistov zrejme vyplýva z presvedčenia, ktoré je v zásade pochopiteľné, a síce, že nie je možné skúmať listy bez zohľadnenia faktu, že sú súčasťou stromov. Zároveň ho však vzhľadom na empirický charakter jazykovedy ako vedy možno označiť za paradoxný.¹⁸ Podobný postoj voči vedeckej (ne)využitelnosti informačných technológií, prinajmenšom v istom období, panoval aj v oblasti takej tradičnej empirickej vedy, akou je fyzika, ako to konštatuje Irving Wladawsky-Berger (2014) v rozhovore s názvom *Why Do We Need Data Science When We've Had Statistics for Centuries?*¹⁹

Vzájomné súvislosti jednotlivých faktorov ľudského poznávania (individuálnych, psychických, biologických, sociálnych) a rôznych druhov technologických artefaktov spoločne vytvárajú zložitý priestor ľudskej skutočnosti a praxe. Tento priestor možno vo vzťahu k poznávaniu jazyka ilustrovať tak, že i teoretický, takzvaný „kabinetný“ lingvista (*armchair linguist*), skeptický voči korpusovým dátam a nástrojom na skúmanie jazyka, občas pri svojej analýze môže, ba musí používať papier a ceruzku, alebo si pri čítaní pomáhať okuliarmi. Pri zdôrazňovaní nových možností korpusového skúmania jazyka sa pritom často používajú práve metafory optických prístrojov ako mikroskop či teleskop (Jarošová, 2018a, s. 133). Inými slovami, technológia je vo svojich rôznych podobách v ľudskej realite všadeprítomná (*ubiquitous*; Introna, 2017). Vzájomnú prepojenosť jazyka, ľudskej mysle a (jazykových) technológií konštatuje aj M. Coeckelbergh, pričom koexistenciu ľudských bytostí, slov a vecí charakterizuje ako istý premenlivý hybridný útvar. Filozofické reflektovanie a skúmanie tohto hybridného útvaru zasa vytyčuje ako možnú úlohu pre takzvanú empiricky orientovanú filozofiu technológie.²⁰ Coeckelbergh v tejto sú-

¹⁸ „Corpora clearly constitute recorded observations of language behavior, so their place in linguistic research seems so obvious that anyone unfamiliar with the last fifty years of mainstream linguistic theorizing will wonder why their use would have to be justified at all“ (tamže).

¹⁹ Wladawsky-Berger v rozhovore uvádza: „I still remember a meeting in 1965 with a very eminent physicist from whom I was taking a graduate course. He asked me what I planned to do research on for my degree, and I told him that I was already working with Prof. Roothaan on atomic and molecular calculations. He just said that good theoretical physics should require no more than pencil and paper, rather than these elaborate new computers. In his mind, this wasn't real physics.“

²⁰ „In empirically oriented philosophy of technology, we need more thinking about the precise ways language and technology can and do relate. This requires not a rejection of the empirical turn. On the contrary, it does not require less but more “empirical” work. For instance, we may do more research on how in specific practices hybrids of humans, words, and things emerge and transform. Contemporary philosophy of technology has done much to “bring together” humans and machines (cyborg metaphor), humans and things“ (Coeckelbergh, et al., 2018, s. 15).

vislosti hovorí o vzájomnom „premieša(va)ní sa“ slov, vecí a ľudí počas ich vzájomných interakcií, pričom tvrdí, že myslíme jednak „prostredníctvom“ (*through*) technológie, jednak prostredníctvom slov.²¹

Zložitú koexistenciu slov, vecí (artefaktov) a ľudskej mysle tiež naznačuje termín „duálny charakter“ (*dual nature*) technických artefaktov. Tento termín sa uvádza v práci *Technical Artefacts: Creations of Mind and Matter*, kde sa rozlišuje fyzický rozmer technických artefaktov, i to, že slúžia ľudským zámerom, a teda majú intencionálny rozmer (Kroes, 2012, s. 42). V práci sa tiež rozlišuje epistemologická dualita technických artefaktov, ktorá spočíva v tom, že poznávame štruktúru i funkciu technických artefaktov; a potom ontologická dualita, ktorá znamená, že sa artefakty „skladajú“ z fyzického prvku a ľudského zámeru (tamže). Duálny charakter technických artefaktov sa naznačuje už v názve Kroesovej knihy: technické artefakty predstavujú výtvary, v ktorých sa prepája myseľ a hmota. Ďalším rozšírením otázok, ktoré si kladie filozofia technológie, je špecifický ontologický štatút týchto technických artefaktov: počítačových technológií, programov a nástrojov, a to vzhľadom na ich sčasti „abstraktnú povahu“ a „výslednú komplexnosť“, ako sa konštatuje v hesle o filozofii počítačovej vedy (*philosophy of computational science*) v Stanfordovej encyklopédii filozofie (Turner – Angius, 2019).

Zložitá interakcia ľudí, strojov a slov, napríklad pri skúmaní jazyka prostredníctvom korpusov, sa reflektuje aj v koncepcii rozšírenej mysle, ktorú uviedli v článku *The Extended Mind* autori Clark a Chalmers (1998). Myšlienkové experimenty predstavené v článku „*mají překonat naši předpojatost, že myšlení a kognice se dějí pouze v našich hlavách*“ (Ferenc, 2019). Zástancovia teórie rozšírenej kognície preto „*vidí externí svět a technologické artefakty jako konstitutivní části naší kognice, které jsou intimně propojené s interními procesy v mozku a formují interaktivní, rozšířené a distribuované kognitivní systémy vyššího řádu*“ (tamže). Podľa teórie rozšírenej mysle predstavujú technické artefakty a technológie istú extenziu človeka, a teda, obrazne povedané, možno papier alebo pamäť počítača či jazykový korpus vnímať ako „predĺženie“ pamäte, jej obdobu a materializovanú a intersubjektívne dostupnú podobu. Okuliare, mikroskopy a teleskopy zasa možno vnímať ako „predĺžené“ oči: „*naše kognice prosakuje do prostředí a potřebuje média, technologické objekty a další lidi pro správné fungování, což činí lidské bytosti přirozenými kyborgy*“ (tamže). Myšlienku, že technológia je „*projekcí lidských orgánů, která zvyšovala jejich*

²¹ „There is also a lot of other interesting thinking through technology [...]. But we also need to realize that this is also always a thinking through words. We must reflect on these words, explore through conceptual work how we can bring together words and technologies, and study how words, things, and humans mix and are re-mixed in concrete practices – practices which are at the same time human, technological, and linguistic“ (Coeckelbergh, et al., 2018, s. 15).

sílu a užitečnosť“, pritom vyslovil už Ernst Kapp (1877) v zakladateľskej práci *Grundlinien einer Philosophie der Technik* (Ferenc, 2019, tamže). Človek čeliaci komplexnej skutočnosti uplatňuje voči tejto realite stratégiu „zjednodušovania komplexnosti“, na čo sa človeku ponúkajú prostriedky ako formálne jazyky, mapy, iné vizualizácie a počítačové nástroje. Tieto spoločne podľa G. E. Yoosa predstavujú „protetické, mnemotechnické vizuálne pomôcky“ pre skúmajúci vedecký subjekt, jeho individuálne limitovanú pamäť a pozornosť (*attention*).²²

V teórii rozšírenej kognície sa tak modeluje prepojenosť mysle a jej okolia. Predmetom ďalšieho skúmania je preto práve miesto a povaha tohto prepojenia, takzvaného rozhrania (*interface*). Hovorí sa dokonca o „vede o interface“, ktorej hlavným zameraním by mala byť „*obecná analýza povahy interface, tedy otázka, čo interface je a jak se liší interface technologie od ostatních možných typů interfaců*“ (Ferenc, 2019). Rozhranie nepredstavuje statickú, fixnú entitu, ale má dynamický, procesuálny charakter vzťahu medzi človekom a technológiou. Povaha rozhrania sa vymedzuje ako „*styčný bod, kde sa potkávají člověk a technologie*“ a ktorý má „*vždy za úkol vyrovnávat nekompatibility mezi člověkem a technologií, a to tím způsobem, že vždy podle nějaké strategie schovává komplexitu technologie i světa a nabízí člověku jen ty funkce a možnosti, které potřebuje*“ (tamže). O korpuse ako externalizovanej, elektronickej a intersubjektívne (objektívne) dostupnej a skúmateľnej vzorke jazyka preto možno hovoriť ako o istom druhu rozhrania, ktoré lingvistovi poskytuje potrebné funkcie a možnosti na skúmanie jazyka. Aj v prípade korpusu máme k dispozícii vizuálne rozhranie a rôzne nástroje na vizualizáciu či iné zobrazenie korpusových dát, respektíve na štatistické vyhodnotenie vyhľadávania v nich. Rozhranie, ktoré tvorí na jednej strane korpusová vzorka jazyka, na druhej strane subjekt (korpusový lingvista), umožňuje, aby lingvista dokázal efektívne pracovať s rôznorodosťou dát vo vzorke jazyka, aby dokázal odhaľovať dátové štruktúry a javy založené na týchto dátach. Korpusová technológia v tomto zmysle slúži ako teoreticko-praktická stratégia, ako sa vysporiadať s komplexnosťou výseku skúmanej objektívnej jazykovej skutočnosti, a všeobecne predstavuje jednu z podôb interakcie človeka s nepreniknuteľnou komplexitou sveta. Lingvistom zasa poskytuje „kognitívnu mapu“ jazykovej reality. Spracovanie, mapovanie a poznávanie v tomto prípade jazykovej reality predstavuje zároveň istú jej transformáciu, premieňa vzťah človeka k jazyku: „*Interface tak*

²² „*Formal systems and mapping help the sciences reinforce powers of attention. They help those in the disciplines to remedy the psychological limits of memory, thus providing scientists with prosthetic, mnemonic visual aids in the use of graphics, script, or text, and that are importantly now being used in spreadsheets and computer printouts to engage attention to a great deal of complexity. [...]. The sciences and the technologies have visual aids to memory generated by formal languages and graphics, and now they are being enhanced exponentially by computers*“ (Yoos, 2015, s. 4).

člověku představuje umělý a syntetický obraz či kognitivní mapu reality, čímž nevyhnutelně transformuje naši percepci nejen toho, čím technologie je, ale též transformuje náš vztah ke světu“ (Ferenc, tamže). Transformácia vzťahu medzi človekom a svetom, respektíve človekom a jazykom, spočíva napríklad v tom, že sú človeku dostupné štruktúry jazykových dát, ktoré by si inak vzhľadom na kognitívne limity individuálnej mysle nedokázal všimnúť, pretože vznikajú na rozsiahlejšej textúre dát. Túto transformáciu tiež vyjadrujú konceptuálne metaforý ako mikroskop či teleskop, ktorými sa, ako bolo uvedené vyššie, zdôrazňuje kvantitatívny i kvalitatívny posun v skúmaní jazyka spojený s nástupom rôznorodých podôb a úloh korpusovej lingvistiky.

5. Rôznorodosť úloh korpusovej lingvistiky

Komplexnosťou a rôznorodosťou sa nevyznačuje iba celok skutočnosti a dát: rôznorodé podoby a úlohy má aj sama korpusová lingvistika a jej príbuzné oblasti (počítačové spracovanie prirodzeného jazyka (NLP), umelá inteligencia a ďalšie). Výrazná rôznorodosť korpusovej lingvistiky sa naznačuje napríklad v knihe rozhovorov so súčasnými svetovými korpusovými lingvistami pod názvom *Perspectives on Corpus Linguistics* (Viana – Zyngier – Barnbrook (eds.), 2011). Rôznorodosť podľa tejto práce vidieť už v pohľadoch na niektoré základné otázky korpusovej lingvistiky: jej historický vývoj, korene, príčiny jej rýchleho rozvoja, či v diskusii o jej štatúte, teda pri otázke, či je korpusová lingvistika skôr metodológiou alebo samostatnou vedou. Ďalšou otvorenou otázkou je korpusová metodológia, ako aj úloha dosiahnuť takzvanú reprezentatívnu korpusu, i to, či a do akej miery je pri korpusovej analýze využiteľná kognitívna schopnosť individuálneho bádateľa, jeho intuícia a introspekcia. Hoci sa význam korpusovej lingvistiky vidí často práve v prekonaní limit a subjektívnej povahy intuície a introspekcie, nemožno povedať, že ju všetci lingvisti odmietajú, a introspekciu sa skôr prisudzuje miesto v širšie poňatom, korpusovo založenom výskumnom cykle: poukazuje na fakt, že sa introspekcia v niektorých ohľadoch dá nahradiť priamo korpusovou analýzou (Jarošová, 2018a, s. 133), respektíve sa využíva pri kvalitatívnej interpretácii korpusových výsledkov (Čermák, 2017, s. 91). Rôznorodé sú aj korpusové materiály a metódy: či už ide o využitie korpusov pri vyučovaní jazykov, porovnávanie a prekladanie textov, ich anotáciu (obohacovanie zdrojov o metadáta, teda o rôzne formy lingvistických i iných značiek a informácií). Mnohoraké je využitie korpusov v jednotlivých lingvistických oblastiach (historická lingvistika, pragmatika, syntax, sociolingvistika, (kritická) diskurzívna analýza, kontrastívna rétorika, výskum metafory, žánrov, registrov, komplexnosti jazyka) a prirodzene rôznorodé je ich aplikované využitie, napríklad na interpretáciu rodových kategórií, skúmanie sociálnych funkcií v literárnej vede alebo skúmanie sociálnej konštruovanosti jazyka a podobne.

Rôznorodosť oblastí a úloh počítačového spracovania prirodzeného jazyka sa tiež načrtáva v syntetizujúcej práci *Natural Language and Speech Processing* (Jurafsky – Martin, 2018). Napriek množstvu a rôznorodosti oblastí, úloh a pojmov²³ si možno všimnúť niektoré ich spoločné znaky. Ide zväčša o ope-

²³ V uvedenej práci sa uvádzajú napríklad tieto oblasti a úlohy NLP: normalizácia textu, klasifikácia sentimentu, využitie neurónových sietí, vývoj jazykových modelov, vektorová re-

rácie vykonávané počítačovým programom na dátach, postupnostiach, reťazcoch znakov tak, ako je to v prípade regulárnych výrazov (Jurafsky – Martin, 2018, s. 1), z „perspektívy“ počítačového systému ide o operácie na číslach (Čermák, 2017, s. 14). Identifikovať možno niekoľko typov operácií: normalizáciu, štandardizáciu textu, modelovanie, klasifikáciu, metódu regresie, vektrovej reprezentácie, anotáciu, vyhľadávanie informácií, preklad textov či spracovanie hovorenej reči. Príkladom operácie je teda normalizácia textu, ktorá má najprv podobu rozdeľovania slov na jednotlivé reťazce znakov, oddelené medzerami, na takzvané tokeny, čo sa označuje ako tokenizácia. Všeobecnejší termín „normalizácia“ sa charakterizuje ako konverzia elektronického textu do istej štandardnej, počítačovo ďalej spracovateľnej formy (Jurafsky – Martin, 2018, s. 10). Jednotlivé oblasti a úlohy v počítačovom spracovaní prirodzeného jazyku sú rozvinuté v rozličnej miere, pričom významné úlohy sa v budúcnosti črtajú v oblasti sémantickej anotácie a rôznej reprezentácie významu.²⁴ Vo všetkých týchto oblastiach a úlohách sa však vytvára dynamický vzťah medzi človekom a počítačom, vzniká technologické rozhranie, prostredníctvom ktorého sa záznam písomnej i ústnej komunikácie skúma ako obrazná stopa slov, na základe ktorej vedec poznáva systém jazyka, pričom sa využíva fakt, že je tento záznam skúmateľný ako digitálny artefakt. Počítačové spracovanie prirodzeného jazyka preto možno označiť za výrazný prejav „technologizácie slova“ (Ong, 2006).

Počítačovo podporované skúmanie prirodzeného jazyka v podobe korpusovej lingvistiky okrem toho vychádza z rozhodujúcej dôležitosti jazykového materiálu, skúmania jazykovej empirie, a vychádza tak z filozofického empiricizmu, pričom sa hovorí i o „nekompromisnom empirizme“ korpusovej lingvistiky (Jarošová, 2018b, s. 298). Myšlienku takéhoto nekompromisného empirizmu výrazne presadzuje František Štícha, ktorý predkladá charakteristiku korpusu ako „*sbírký textů, která je neměnně a trvale k rychlé a pohotové dispozici*“ (2008, s. 12). Uvedené vlastnosti korpusových zdrojov (ich nemen-

prezentácia významu slov, morfológická anotácia, spracovanie postupností slov pomocou takzvaných rekurentných neurónových sietí, výskum formálnych gramatík, syntaktická, štatistická a závislostná anotácia, úloha počítačovej reprezentácie významu viet, oblasť počítačovej sémantiky a sémantickej anotácie, oblasť získavania informácií (*knowledge retrieval*), označovanie sémantických rolí, vývoj lexikónov pre rôzne druhy sentimentu, afektu a konotácií, prepájanie entít (*entity linking*), otázka koherentnosti diskurzu, vývoj strojového prekladu, dialógových systémov a chatbotov, rozpoznávanie a syntéza reči.

²⁴ V tejto oblasti ide napríklad o takéto úlohy: „*better characterise semantic measures and their semantics, provide theoretical and software tools for the study of semantic measure, standardise ontology handling, improve models for compositionality, study current models for semantic measures w.r.t. language specificities, promote interdisciplinarity, study the algorithmic complexity of measures, support context-specific selection of semantic measures*“ (Harispe et al., 2015, s. 115 – 121).

nosť, trvalosť a rýchla a pohotová disponibilita) sú príznakom transformácie, k akej v dôsledku nástupu počítačového spracovania prirodzeného jazyka došlo vo vzťahu medzi človekom a „jeho jazykom“ (Horecký, 2000). Štícha vidí prínos tejto transformácie v tom, že lingvista má vďaka nej popri nevyhnutne limitovaných bezprostredných zdrojoch vlastnej pamäte, skúsenosti, intuície a introspekcie k dispozícii „úžasný fenomén obrovského množství tvarové variantnosti slov spolu s nejrůznějšími okolnostmi, které tuto variantnost – leckdy po celá staletí – provázejí a podmiňují, který lze nyní díky korpusům studovat nesrovnatelně lépe, hlouběji, objektivněji a efektivněji, než tomu bylo v minulosti“ (2005, s. 10). Štícha tiež naznačuje rôzne jazykové roviny a druhy javov, ktoré je možné skúmať prostredníctvom korpusov: na rozsiahlom jazykovom materiáli sa zisťujú možnosti (prípustné variantné tvary) i medze (hranice prípustnosti tvarov) na rovine formálnej morfológie, funkčnej morfológie, tvorenia slov, syntaxe, slovosledu. Vo vzťahu k budúcemu vývoju korpusovej lingvistiky zasa uvažuje o situácii, že termín korpusová lingvistika sa v budúcom vývoji lingvistiky ukáže ako dočasný, temporálny: „Empiricko-teoretická a deskriptivní lingvistika budoucnosti, bude-li soustavně a cílevědomě vycházet z práce s velkými počítačovými korpusy, nebude mít zapotřebí připojovat před své jméno epiteton „korpusová“ (Štícha, 2008, s. 12). Korpusová lingvistika by v takom prípade získala význam prechodného historického obdobia, počas ktorého sa tento nástroj a materiálna báza v jazykovede postupne adaptovali ako jej bežná, až samozrejmalá súčasť. Využívanie korpusov pri jazykovednom výskume pritom aj podľa W. Teuberta predstavuje v súčasnosti už viac-menej samozrejmosť.²⁵

Z filozofického, fenomenologického hľadiska však introspekcia stále predstavuje relevantný primárny zdroj poznania, a to konkrétne zdroj poznávania „v prvej osobe“, ako ho navrhuje Ivan M. Havel (2018) vo svojich „introspektorských“ analýzach vlastných snov. Fenomenologická filozofia sa pritom odvracia od subjektivismu a zameriava sa – podobne ako analytické prístupy k filozofii a napokon aj empiricky založená a teoreticky fundovaná veda – na fenomenálne dané zistenia o základných štruktúrach poznávania a prežívania (v prípade subjektívne založenej fenomenológie) či skutočnosti vôbec (v prípade asubjektívne založenej fenomenológie). V lingvistike má svoje miesto i intuícia, ktorá vyvažuje „korpusové poznanie“, hoci ich vzťah nie je úplne jasný (Čermák, 2017, s. 121). Podľa Jarošovej zasa môžeme vďaka korpusom vynechať stupeň introspekcie, keďže máme k dispozícii štatisticky vytvorené schémy toho, aké slová, tvary, druhy slov sa reálne vyskytujú či nevyskytujú v kvantitatívne merateľných kombináciách s ďalšími slovami, tvarmi, slov-

²⁵ „No introspection can claim credence without verification through real language data“ (Teubert, 2005, s. 1).

nými druhmi: „Dnes už nielenže môžeme každú generalizáciu vytvorenú len na obmedzených dokladoch dodatočne overiť na veľkom množstve elektronicky uložených dát, ale môžeme vynechať stupeň introspekcie a pozrieť sa na schémy vytvorené štatisticky spracovanou spájateľnosťou skúmaných slov“ (Jarošová, 2018a, s. 133).

6. Korpus ako elektronická vzorka jazyka

Charakteristikou korpusovej lingvistiky, jej úlohami a definíciami sa zaoberajú práce o korpusovej metodológii (napr. Stefanowitsch, 2018) i príručky korpusovej lingvistiky (z veľkého množstva napr. McEnery – Hardie (2012)), príručky používania korpusov (Šimková et al., 2017), celostný pohľad na problematiku korpusovej lingvistiky, počnúc povahou korpusových dát, spôsobmi ich spracovania, až po náčrt možných výskumných otázok korpusovej lingvistiky v budúcnosti podáva Čermák (2017).

Vychádza pritom zo základnej charakteristiky toho, čo je informácia a písaný text. Tlačený text ako nosič informácií porovnáva s povahou elektronického textu, elektronických korpusov, pričom informáciu možno uchopiť rôzne, v jazykovednom výskume však v zásade predstavuje poznatok o jazyku získaný analýzou jazykového materiálu a interpretovaný prostredníctvom sociálnych, semiotických a pragmatických, ale i funkčných a iných rámcov (Čermák, 2017, s. 12). Zdrojom informácií pri skúmaní elektronických zdrojov, korpusov, sú textové jednotky (takzvané *tokens*), čo sú reťazce znakov oddelené medzerou, dostupné na skúmanie a prehľadávanie vo svojich pôvodných kontextoch, čím spolu vytvárajú elektronický text. Kým tlačený text sa prehľadáva individuálnymi „očami“, individuálnou „pamäťou“, elektronický text umožňuje prehľadávanie (a teda „videnie“) pomocou korpusových nástrojov či, napríklad, regulárnych výrazov, a možno ho tiež zobrazíť pomocou grafických nástrojov. Vyhľadávanie v korpusových zdrojoch sa teda, prirodzene, uskutočňuje aj pomocou vizuálneho kontaktu, ale na rozdiel od fyzického čítania, listovania v materiáloch sa uskutočňuje tiež na základe formálnych požiadaviek na vyhľadávané reťazce znakov: tieto nástroje predstavujú pre človeka akési elektronické „oči“, ktorými môže skúmať správanie sa vybraných konkrétnych výrazov v jazyku (teda pôvodných dát), alebo konkrétnych typov výrazov (skúmaných na základe pridaných značiek dát – anotácií, metadát).

Písaný text má jednak lineárnu i „lineárne priestorovú“ povahu (vľavo-vpravo) i lineárne časovú povahu (najprv-potom), pričom paralelné ďalšie informácie, najmä intonácia, sa v písme nezachytia (Čermák, 2017, s. 13). Digitalizácia textu a iných zdrojov zároveň znamená vytváranie elektronických artefaktov, ktoré sú vnímateľné a skúmateľné iným spôsobom, než to umožňujú fyzické artefakty: pri fyzickom artefakte sa, napríklad, môžeme zamerať na istý aspekt, tvar alebo iný príznak nejakého keramického výrobku či architektonického diela, a skúmať ho na základe konkrétnych výrobkov a stavieb, alebo

pomocou ich trojdimenzionálnych modelov. Je zrejmé, že v prípade fyzických výrobkov a stavieb sme limitovaní iným spôsobom, než v prípade skúmania prostredníctvom elektronických artefaktov, modelov, reprezentácií pôvodných artefaktov. Pri skúmaní vo virtuálnom prostredí skúmame v istom zmysle tieto reprezentácie, zároveň však platí, že reprezentáciu (v tomto prípade jazykového správania) môžeme vnímať z iných uhlov a inými spôsobmi, než aké nám poskytuje fyzická skutočnosť, respektíve nám ich táto poskytuje v inej miere: po skúmanom architektonickom objekte sa, napríklad, nemôžeme pohybovať ľubovoľnými smermi a v ľubovoľných uhloch. Iným príkladom je monitorovanie, vizualizácia nedostupných prírodných priestorov, ako sú jaskyne, a to pomocou sond, ktoré vytvárajú presný model vnútra jaskyne, ktorý by sme inak získali len ťažko. Takto opäť môžeme vnímať to, čo by sme vzhľadom na nedostupnosť priestorov jaskyne a limity našej kognície, napríklad rozmerov nášho tela a schopnosti pohybu po rôznych povrchoch a smeroch, vidieť nemohli. Virtuálna reprezentácia, model fyzického sveta, navyše nie je iba reprezentáciou, znakom pôvodného sveta, ale má i nové vlastnosti, napríklad disponibilitu a manipulovateľnosť jednotlivých častí a celku dát, čím sa z fyzického predmetu ako zdroja skúsenosti stáva virtuálny predmet ako zdroj potenciálneho poznania.

Korpusové dáta, či už nespracované alebo obohatené o anotácie predstavujú analógiu pôvodných, tlačených dát, a podľa Čermáka sú *„zhruba řečeno moderním elektronickým ekvivalentem starých a rozsahem omezených manuálních excerpt z různých (většinou psaných) zdrojů (v podobě kartotéčních lístků)“* (Čermák, 2017, s. 13). Texty, pôvodne písané a tlačené, sa zachytávajú, znázorňujú v kódach, ktorými sa tieto dáta prenášajú do inej podoby, konkrétne premieňajú sa na čísla (tamže, s. 14). Vo vytvorenom korpuse sa potom vyhľadáva priamo slovo, tvar alebo istý typ výrazu či kombinácia slov, tvarov, typov výrazov, a to zadaním požiadavky na vyhľadávanie pomocou regulárnych výrazov či pomocou značiek (tamže, s. 42). Vyhľadaním výskytov danej formy, respektíve kombinácie foriem sa získava materiál, ktorý sa potom štatisticky analyzuje, čím sa objavuje istá pôvodne skrytá vrstva informácií (napríklad o vzájomných distribučných pomeroch skúmaných textových jednotiek a javov) a ďalej sa vyhodnocuje kvalitatívne, vzhľadom na kontext formy či kombinácie foriem, ale tiež vzhľadom na použitý teoretický rámec.

Značkovanie korpusových dát prebieha ako do veľkej miery automatické prepojenie dát s anotačnými značkami, ktoré sú „štruktúrne“ i „lingvistické“ (tamže), vyjadrujú typ znaku alebo typ výrazu, slovnodruhovú príslušnosť či ďalšie morfológické, ale i sémantické kategórie. Tým sa dáta obohacujú o svoje metadáta. Metadáta tvoria, obrazne povedané, ďalšie vrstvy skúmaného textu, hoci sú „jasne externé“ (tamže, s. 33), zároveň však štatistické a iné korpusové analýzy rozširujú poznanie o inak „neviditeľných“, intuíciou neuchopiteľných

súvislostiach a javoch aj na základe neanotovaných dát. V súvislosti so značkováním je potrebné zdôrazniť, že sa uskutočňuje do veľkej miery automaticky, keďže manuálne značkovanie celého korpusu by vzhľadom na objem spracúvaných dát nebolo fyzicky možné. Značkovanie sa preto najprv na základe istých pravidiel natrénuje na vzorke a počítač potom, opäť na základe pravidiel, prideluje značky (napríklad morfológické tagy. S automatickým rozpoznávaním tvarov je v dôsledku toho najmä v prípade tvarovo bohatých jazykov a súvisiacej homonymie tvarov spojená aj istá chybovosť, s ktorou je však možné pracovať, zohľadniť ju, odstraňovať a zakomponovať do vyhodnocovania výsledkov i do zlepšovania postupov vyhľadávania. Automatické značkovanie sa uskutočňuje prostredníctvom výberu z takzvaného elektronického slovníka: tvar sa porovnávaním s týmto slovníkom priradí k niektorej položke v slovníku, priradí sa mu značka zodpovedajúca tejto položke a tvar sa určí týmito morfológickými značkami (Čermák, 2017, s. 33). Z kontextu foriem, ich vzájomných kombinácií a distribučných pomerov v jednotlivých dokumentoch, častiach korpusu alebo v celom korpuse sa po zohľadnení limit vyhľadávania získava množstvo informácií: tak, obrazne povedané, čítame v druhej a ďalších vrstvách elektronického textu.

Na jednej strane máme teda pôvodný elektronický text, na druhej jeho druhú vrstvu, vrstvu značiek (metadát), ktorá predstavuje „externý“ zdroj informácií o prvej vrstve. Prostriedkom, ako získavať informácie o prvej vrstve, je skúmanie distribučných pomerov a iných zákonitostí či pravidelností. Čermák uvádza takýto príklad: *„a označené jako konjunkce a následující tak označené jako adverbium [...] však dohromady jsou pouze komponenty vyšší jednotky, která je však především partikulí a tak a má tuto funkci jen díky kontextu, resp. souvýskytu obou slov“* (tamže, s. 93). Korpusový kontext sa definuje ako okolie daného slova či prvku v texte, ktoré je potrebné na určenie konkrétneho významu a/alebo funkcie takého slova (tamže, s. 25). Existujú však i názory, že z korpusu z hľadiska získavania poznatkov principiálne vyťažíme „iba to, čo sme do neho vložili“.²⁶ Toto tvrdenie treba vnímať skôr v tom zmysle, že použitý spôsob anotácie korpusu, zloženie a štruktúra anotačných značiek „určuje“, aké výskumné otázky si budeme môcť položiť.²⁷ Ak sa sada anotačných značiek (takzvaný *tagset*) bude skladať z desiatich značiek, výskumné otázky sa zrejme budú týkať štruktúrnych, funkčných a distribučných vzťahov na základe týchto desiatich značiek (typov či vlastností tokenov). Pridaná „vrstva“ metadát sa teda odvodzuje z použitej koncepcie značkovania²⁸, ktorú zasa ovplyvňuje te-

²⁶ „We can only get out of a corpus what we put in“ (Wallis, 2007).

²⁷ „Annotation decisions determine the types of research questions that we can ask“ (tamže).

²⁸ Ku koncepcii morfológickej anotácie zdrojov Slovenského národného korpusu pozri napríklad Garabík et al. (2004).

oretická pozícia autora koncepcie, a v tomto zmysle je teória „idiosynkratická“ (Čermák, 2017, s. 37).

V korpusovej lingvistike je pritom nevyhnutná štatistická analýza, ktorá však predstavuje iba jeden z krokov vo výskumnom procese, nie je jeho výsledkom (tamže, s. 101). Výsledky tejto analýzy však „*môžu ukazovať na existenciu dosiaľ neregistrovaných a nepopsaných řetězců složených z více slov, tvarů, a tedy případně i nových jevů*“ (tamže, s. 95). Okrem toho sa v korpusovej lingvistike uplatňujú rôzne metódy (testy) i takzvané asociačné miery, ktoré spoločne predstavujú spôsob, ako štatisticky vyhodnocovať výsledky korpusového vyhľadávania (napríklad konkrétneho výrazu alebo kombinácie výrazov).²⁹

Korpusové vyhľadávania prinášajú „výsledky škálovitej povahy“ a táto škálovitosť a gradientnosť je príznačná pre povahu samotného jazyka ako „*otevřeného, proměnlivého a nedeterministického systému*“ (tamže, s. 120). Korpusové dáta sa na rozdiel od „starého škatulkovania“ javia ako „splývavé“, príslušnosť výrazu k nejakej triede výrazov, a teda jeho kategoriálna príslušnosť sa hodnotí stupňom (tamže, s. 97). Na príklade výrazu „Dovolíte?“, ktorého pragmatická funkcia sa nedá odvodiť zo slovníkového významu lexémy „dovoliť“, Čermák ukazuje, že „*token, tvar lexému nemá stejné vlastnosti ani chování jako (celý) lexém*“ (tamže, s. 98). S odkazom na prácu *Víc než součet částí* (Svítek, 2013) teda azda možno povedať, že vo svetle korpusových analýz neplatí len to, že celok je viac než súčet svojich častí, ale aj naopak, časť (token, tvar lexémy) je niekedy viac než celok (celá lexéma). To zároveň pripomína tézu, že význam nemá kompozičný charakter. Napríklad, význam frazémy nezískame zložením významu jej častí (Čermák, 2017, s. 100). V tejto súvislosti aj Jarošová zdôrazňuje, že odhalené štruktúry (jazykových) dát predstavujú iba istý krok korpusovej analýzy, ďalším je interpretácia štruktúr pomocou zvoleného interpretačného rámca, teórie: „*Tieto jednotky treba dôkladne analyzovať a interpretovať z pohľadu určitého teoretického rámca, v ktorom hrá dôležitú úlohu aj funkčná perspektíva, teda otázka, akú funkciu plní daný spoluvýskyt v jazyku*“ (Jarošová, 2018a, s. 133). Aj tu však platí, že skúmaním korpusových zdrojov objavujeme štruktúry, „väčšie celky“, a tak sa obrazným pohybom odhaľuje druhá a ďalšie vrstvy dát a potenciálnych javov: „*Ideme teda od čistých dát smerom hore, k väčším celkom, ktoré nevytvárame my, ale frekvencia a distribúcia jazykových javov*“ (tamže). Korpusový pohľad podobný Hubblovmu teleskopu umožňuje jazykovedcom vidieť „*veci, ktoré by inak nemali šancu vidieť*“ (tamže), umožňuje im, ako to konštatoval aj Čermák, objaviť potenciálne nové javy v jazyku. Skúmanie významu a jazykového systému prostriedkami korpusovej analýzy i spracúvanie korpusových zdrojov pre potreby lexikografického vý-

²⁹ Rôzne (a navzájom kombinovateľné) lexikálne miery, ktoré možno využiť napríklad na identifikáciu potenciálnych kolokácií, rozoberá P. Pecina (2009).

skumu a práce si však vyžadujú „uznanie prvotnosti textu“ a „syntagmatického rozmeru slova“: „*kombinatorika slova je nielen dopĺňajúca metóda výskumu lexiky, ale jej základný a nevyhnutný predpoklad. Význam slova vytvorený ohraňovaním potenciálu obsahov vedomia, bázy znalostí o svete, sa objavuje až vtedy, keď je slovo súčasťou väčšieho, komunikačne relevantného kontextu*“ (tamže, s. 135). Obrazne túto črtu jazyka, jeho spôsob jestvovania, možno vyjadriť tak, že textové komponenty, výrazy, jednotky jazyka netvorí jazyk, jeho textúru, izolovane, ale koexistujú v uzoálne i systematicky vytváraných „trsoch“ jazykových prejavov, od menších jazykových jednotiek až po príznaky veľkých textových celkov.

Vo svetle korpusových analýz sa tak odhaľuje akoby spojitý a do veľkej miery, hoci nie nevyhnutne a úplne kontextovo daný (Zouhar, 2011, s. 28) význam jazykových jednotiek: ten má „*difúzny, vágny a vysoko variabilný charakter*“ (Jarošová, 2018a, s. 135). A. Jarošová, podobne ako Jan Kořenský, vychádza z interdisciplinárne otvoreného procesuálneho modelovania reči, ktoré v porovnaní s predchádzajúcim obdobím „*radikálne revolucionizuje lingvistické myslenie v celú té podobu, jak se utvářelo po staletí včetně období klasických strukturalismů*“ (Kořenský, 2014, s. 5). Vo svojej všeobecnej jazykovednej i špeciálne lexikografickej perspektíve A. Jarošová siaha po pojme komunikačnej udalosti, čo len podčiarkuje situačnú, kontextovú viazanosť jazykových výrazov, v ktorej význam vzniká udalostne, na rozdiel od významu ako statickej, slovníkovo viazanej entity: „*Komunikačná udalosť generuje aktuálny význam*“ (Jarošová, 2018a, s. 139 – 140). Z filozofického hľadiska nie je nezaujímavé, že kontextovú viazanosť významu, prvotnosť textu a syntagmatický rozmer slova si uvedomoval i L. Wittgenstein v súvislosti s koncepciou jazykových hier a používania slov v týchto hrách. Poznať význam slova znamená v duchu tejto koncepcie vedieť slovo použiť v kombinácii s ďalšími slovami, v ich kontexte. Možno dokonca povedať, že Wittgensteinove myšlienky tak nepredstavujú iba príspevok k filozofii technológie, ale tiež, konkrétne v *Logicko-filozofickom traktáte* (1922), myšlienково anticipujú postupy korpusovej technológie.³⁰

³⁰ „*In creating pre-computational samples of the world through language, philosophers and some linguists hope to come to a fuller understanding of how factual truth [...], or the logic inherent in states of affairs, is represented within those worlds*“ (Louw, 2011, s. 172).

7. Odhaľovanie skrytej štruktúry dát

Prostredie, v ktorom sa človek pohybuje, sa na jednej strane vyznačuje často až nepreniknuteľným množstvom zmyslových podnetov, informácií, a to v podobe poznatkov, subjektívnych hodnotových systémov, seba-prezentácií, ale vyznačuje sa tiež komplexnosťou dát, ktoré mu poskytuje napríklad korpus ako vzorka jazyka. Na druhej strane existujú prírodne dané limity a iné danosti, ktoré utvárajú spôsob, akým sa človek môže vzťahovať k svetu a informáciám, ale aj subjektívne zaužívané myšlienkové postupy, ktorými jednotlivec ku svetu pristupuje. Tieto myšlienkové schémy si nie vždy uvedomuje, respektíve nie vždy ich dokáže introspektívne uchopiť, takpovediac vystaviť vnútornému zraku, keďže môžu vychádzať z komplexu podvedomých prepojení, akcií a reakcií. Je ťažké zmeniť čosi, čo človeka spoluvytvára, ak sa s týmto niečím nedokáže konfrontovať, obrazne povedané, zazrieť to v zrkadle vedomia: zistiť obrysy toho, alebo len na to ukázať či uvedomiť si to. Výsledkom takejto situovanosti je, že človek má napriek obrovskému množstvu informácií niekedy nedostatok tých, ktoré práve potrebuje na zorientovanie sa vo vonkajšom a/alebo vnútornom svete. Tieto situácie predstavujú problémy iného rádu, sú to spomenuté „neprijemné problémy“ (*wicked problems*), problematické tým, že takpovediac nemajú úplné zadanie. Čo v prípade, ak si aj neúplne zadaná otázka vyžaduje odpoveď?

Helen Shulmanová v tejto súvislosti uvádza prípad schizofrenickej pacientky, ktorej paralýza vo vzťahu k okolitému svetu spočívala v tom, že pacientka (a teda ani jej terapeutka) nedokázala zrekonštruovať, získať prístup k istému komplexu vnútorných reprezentácií a traumy.³¹ Problém v takom prípade spočíva v tom, že si daný subjekt na vlastný problém nedokáže spomenúť, nedokáže si ho sprítomniť v mysli, a tak sa s ním konfrontovať ako s problémom, preto ide o problémy iného rádu, naplňajúce znaky „neprijemných“ problémov. Tieto vnútorné, psychické udalosti môžu navonok pôsobiť až banálne, subjektívne však predstavujú akýsi tmel ľudskej integrity, myslenia, osobnosti: stačí chýbajúca „tehlička“, a seba-prezentácia a seba-konštrukcia (ega) sa neuskutoční. Apelovať na zdravý rozum v takom prípade pripomína situáciu, ako keby sa človek so strechou nad hlavou čudoval tomu, kto by mal obývať dom

³¹ „Because of Gina's incomplete memory grip on her experience, I couldn't find a way to formulate a question that would help sort this out. When I tried, I encountered fog“ (Shulman, 2012, s. 42).

bez strechy. Možno tu tiež spomenúť román *Remainder* od Toma McCarthyho (2005), ktorého hlavná postava sa usiluje nákladne a doslova zrekonštruovať fyzické okolnosti, ktoré bezprostredne predchádzali jeho traumatickému výpadku pamäte. Príznačný je v tomto prípade už názov románu: z „ja“ sa kvôli vypadnutému kúsku pamäte stal iba akýsi „zvyšok“ osobnosti, sčasti azda funkčný, avšak neintegrujúci vedomé, podvedomé i nevedomé zložky osobnosti. Psychologička Elyn R. Saksová opísala svoje vlastné, podobné skúsenosti s nefunkčnou integráciou zložiek svojej osobnosti v knihe o tom, ako sa „centrum neudrží“ (*the center cannot hold*; Saks, 2008).

Človek teda potenciálne čelí situácii, kedy musí „odhadnúť informáciu z dát vágnej povahy“ (Půlpán, 2012). Vo vzťahu k takýmto dátam si môže vytvoriť istú ich predstavu, schému, model, ktorým redukuje „rozmernosť“ (dimenziálnosť) komplexnej situácie (skutočnosti). Modely podľa Půlpána predstavujú apriórne predstavy, ktoré sú vo vzťahu k skúmanému súboru dát logicky možné, bezrozporné a predstavujú spôsob, ako transformovať komplexné dáta do uchopiteľnej podoby (2012, s. 9). Model, inými slovami, predstavuje „*návod ke špecifické transformaci dat z určité oblasti pozorování na jejich interpretaci, která je pro nás informací*“ (tamže). Človek, pracujúci napríklad s komplexným výsledkom korpusového vyhľadávania, stojí potom pred otázkou, ku ktorej z možných máp, ku ktorému z modelov (neprotirečivých predstáv), sa má prikloniť pri analýze a interpretácii získaných výsledkov. Podľa Půlpána sa treba prikloniť k tej predstave, ktorá je aktuálne najlepšie zdôvodniteľná a vedie ku konštrukcii modelu, ktorého dôsledky sú odpoveďami na otázku o význame týchto dát (tamže). Model, ktorý možno v závislosti od meniacich sa podmienok postupne modifikovať, je totiž použiteľný do tej miery, do akej je schopný predpovedať a vysvetľovať daný jav v skúmanej oblasti.

Základom vzťahu ku svetu, ktorý charakterizujeme ako vedecké poznávanie, je meranie rôznych veličín, to však neznamená, že prísna veda ostáva iba pri meraní a zaznamenávaní okolitého sveta. Poznanie predstavuje skôr kantovskú syntézu empirického a apriórneho zdroja poznania: „*Většina hodnotných poznatků byla nalezena prostřednictvím matematických vztahů, které ovšem nemusí být jen vztahy mezi metrickými veličinami*“ (tamže, s. 10). Vedecké poznávanie však tvoria okrem teoretickej a metodologickej zložky aj metódy a „prakticko-projektívne postupy“ (Černík, 2014), ide teda o celý komplex apriórnych predstáv, ako aj observačného, klasifikačného, experimentálneho a teoretického poznávania (Filkorn, 2009; s. 32 – 34, 46 – 49, 64 – 110, 187 – 188). Vedecké poznávanie vo svojej štruktúrnej a typovej bohatosti transformuje pôvodnú chaotickú zmyslovú skúsenosť, pozorovanie, individuálny kontakt s okolitým svetom na metodickú, usmerňovanú a intersubjektívne kontrolovateľnú interakciu skúmajúceho subjektu a sveta. Pozorovanie, klasifikácia a experimentovanie sa odohrávajú v transformovanom, „vynájdennom“

prostredí objektívnej reality (Markoš, 2015), uskutočňujú sa ako procesy napojené na konceptuálnu sústavu, na pojmy, prostredníctvom ktorých možno vôbec uchopovať okolitý svet ako predmet vedeckého poznania a budovať konceptuálne schémy ako abstraktné, objektívne jestvujúce entity.

Meranie ako zložka a základ vedeckého prístupu ku komplexnej skutočnosti je pri tejto transformovanej skúsenosti ohrozované vágnosťou a nedostatočným praktickým a metodologickým spracovaním opisovanej skutočnosti, respektíve ohrozované nedostatočným objasnením vzťahu medzi skúmajúcim subjektom a skúmaným súborom dát (Pülpán, 2012, s. 10). Aby bolo možné efektívne použiť matematické pojmy a postupy na modelovanie okolitej skutočnosti, je podľa Pülpána nevyhnutná aktívna a korektná medziodborová spolupráca, respektíve „obojsmerná“ poučenosť, informovanosť medzi výskumom v humanitných vedách a matematickými metódami, ale i naopak (tamže). Vo svete sa však, prirodzene, dokážeme orientovať aj nevedecky, provizórne, improvizovane. Ak chceme tento spôsob orientovania istým spôsobom transformovať na vedecké poznávanie, teda tak, že chceme objavovať skryté príčiny, narážame na druhú komplexnosť, komplexnosť dát, v ktorej sa opäť musíme pokúsiť orientovať. Na vyjadrenie tohto aspektu skutočnosti možno použiť aj pojem „viacrozmerných dát“ a zamerať sa na ich analýzu.

Tak ako v prípade modelu ako apriórnej predstavy, ktorá umožňuje transformáciu dát do uchopiteľnejšej podoby, aj v prípade analýzy viacrozmerných dát ide o ich transformáciu do uchopiteľnejšej podoby, o redukciu ich rozmerosti, o objavenie „štruktúry ukrytej v dátach“ (Meloun et al., 2012, s. 24); a podobne ako v prípade odhadu informácie z dát vágnej povahy, aj tu môže medzi verejnosťou panovať istá obava z využitia viacrozmernej štatistickej analýzy dát. Problém však v skutočnosti spočíva v korektnej formulácii danej úlohy a interpretácii, pochopení výsledkov štatistických analýz, hoci platí, že vďaka viacrozmerným štatistickým technikám je získavanie informácií z dát efektívnejšie než pri využití jednorozmerných dát (tamže, s. 12). Odhalené ukryté štruktúry v dátach navyše prinášajú hlbšie pochopenie súvislostí medzi skúmanými javmi a otvárajú priestor pre interdisciplinaritu (tamže, s. 14). Na tomto mieste sa ponúka porovnanie viacrozmerných techník skúmania dát s „jednorozmerným“, či skôr menej rozmerným spôsobom vnímania, ktorý, povedzme, bežne postačuje na orientáciu v každodennej skutočnosti. Pri analýze viacrozmerných vstupných dát je tiež dôležitá ich kvalita. Už táto selekcia a dôraz na kvalitu dát znamená isté prvotné spracovanie, istú transformáciu výseku skutočnosti: „*Data musí obsahovať užitočnou informáciou*“ (tamže, s. 18). Cieľom je, podobne ako pri odhade informácií z dát vágnej povahy, maximalizácia informačného zisku (tamže, s. 13), ktorý možno vyjadriť obrazne: „*Ta spousta zdánlivě nesouvisějících čísel najednou dostává řeč. Místo mlhavých, občasných náznaků začíná být jasno*“ (tamže, s. 14). Objektívna skutočnosť –

príroda – a dáta, ktoré predstavujú predmet skúmania prírodných a iných vied, sú viacrozmerné v tom zmysle, že príčinné súvislosti medzi nimi si vyžadujú zohľadnenie viacerých faktorov a hľadísk: „*Příroda je vícerozměrná. Většina technologických, ale i ostatních dat zpracovávaných statistickými metodami je vícerozměrná. Chceme-li studovat jeden jediný jev, musíme si uvědomit, že je ovlivněn řadou faktorů*“ (tamže, s. 17).

Keďže v skúmanom súbore dát existuje viacero znakov dát, ich skrytá štruktúra sa odhaľuje tak, že sa na súbor dát aplikujú takzvané ordinačné metódy, medzi ktoré patrí napríklad analýza hlavných komponentov (Principal Component Analysis, PCA). Jej hlavná myšlienka spočíva „*v lineární transformaci původního souřadnicového systému do souřadnicového systému tzv. hlavních komponent, které jsou vzájemně ortogonální (nekorelované) a vybrané tak, aby postihovaly maximální množství informací vyjádřené variabilitou mezi objekty*“ (tamže, s. 20 – 21). Cieľom metódy je zníženie dimenzionality dátového súboru, inými slovami „*snížení dimenze úlohy čili redukce počtu znaků bez velké ztráty informace, a to užitím pouze prvních několika hlavních komponent*“ (tamže, s. 68). Existuje, prirodzene, oveľa viac, sčasti súvisiacich, navzájom analogických či inverzných štatistických metód analýzy viacrozmerných dát, napríklad korešpondenčná analýza, viacnásobná korešpondenčná analýza, viacrozmerné škálovanie, zhuková analýza, logistická regresia, viacrozmerná regresia, metóda najmenších štvorcov či (exploračná) faktorová analýza. Výber konkrétnej metódy závisí od konkrétneho typu dát a špecifikách výskumnej otázky či vedeckého odboru (tamže, s. 11) a pri spracovaní viacrozmerných dát je potrebné zohľadniť aj ďalšie aspekty: 1. jedným aspektom môže byť v porovnaní s „rozmernosťou problému“ nedostatočný rozsah spracovaných dát; 2. vzorka komplikovanej skutočnosti je takpovediac stále príliš jednoduchá; 3. problémom môžu byť takzvané nelinearity, neaktivita a iné štruktúrne javy, ktoré je potrebné na skúmanom súbore dát identifikovať; 4. rozdelenie dát len zriedka zodpovedá takzvanému normálnemu rozdeleniu; 5. v dátach sa vyskytujú „podozrivé“ štruktúrne javy a rôzne iné „heterogenity“; 6. štatistické modely sa často tvoria na základe predbežných informácií; 7. pri výbere modelu opisujúceho „správanie sa“ dát panuje „istá neurčitosť“ (tamže).

Spríevodným javom pri odhaľovaní ukrytých štruktúr v dátach je teda, tak ako pri odhade informácií z dát vágnej povahy, istá vágnosť, ktorá sa v tomto prípade týka výberu metód. Objavovanie poriadku akoby bolo zasadené do procesov, ktoré sú takpovediac ukryté v hmle. Štatisticky analyzovaný empirický materiál je potrebné opisovať s obozretnosťou, pamätať treba aj na to, že odhalenie korelácie medzi dátami ešte neznamená odhalenie príčinnej súvislosti (tamže, s. 24) a súčasťou súboru skúmaných dát je aj vplyv použitých výskumných prostriedkov a nástrojov (inštrumentálny šum) či ostatné náhodné chyby. Skúmaný výsek objektívnej reality sa tak skladá z hľadanej ukrytej štruktúry

dát a šumu. Kým šum a náhodné javy budú „vždy prítomné“, v dátach sa môžu ukrývať aj „iné javy“, čo opäť naznačuje možnosť, keď sa z dobre riešiteľného problému stane „neprijemný“ problém (*wicked problem*): „*Instrumentální šum a ostatní náhodní chyby budou vždy přítomny v datech, ale mohou zde být i jiné jevy*“ (tamže, s. 24 – 25).

Jednotlivé metódy štatistickej analýzy viacrozmerých dát sa líšia podľa toho, či hľadajú „*strukturu a vazby ve znacích*“ alebo „*strukturu a vazby v objektech*“ (tamže, s. 55), pričom cieľom týchto štatistických metód je „*zpracovat data tak, aby se zřetelně indikoval model a tak odkryl skrytý jev*“ (tamže, s. 69). Pri odhaľovaní skrytého javu sa analýza zameriava na takzvané sledovanie rozptylu dát a objavovanie smerov ich maximálneho rozptylu. Myšlienka analýzy viacrozmerých dát je totiž založená na predpoklade, že práve pomocou smeru maximálneho rozptylu dát sa dá odhaliť skrytá súvislosť skúmaných dát a javov (tamže). Postupuje sa teda od spracovania a vizualizácie dostupných dát, od pôvodnej informácie v prvej vrstve skutočnosti, k hľadaniu skrytých javov za touto pôvodnou informáciou, k druhej a ďalším vrstvám skutočnosti. Spracovanie dát a odhalenie ich skrytej štruktúry tak otvára priestor pre formuláciu príčinných súvislostí medzi týmito štruktúrami. Tieto príčinné súvislosti nie sú iba výsledkom interpretačnej činnosti subjektu, ale obrazne ich odhaľuje samotná štruktúra dát.

Odhaľovaniu pravidelností v dátach, ktoré sa potom definujú ako javy, sa venuje aj Brian Haig v práci o „filozofii kvantitatívnych metód“ (2018). Rozoberá pritom všeobecné otázky filozofie vedy a bližšie sa venuje povahe a vzájomným vzťahom. Porovnáva tiež využívanosť a využiteľnosť niekoľkých metód a metodologických pozícií v behaviorálnych vedách: okrem vyššie spomenutej exploračnej analýzy dát sú to testy štatistickej signifikancie, bayesiánske chápanie štatistiky, metóda meta-analýzy a – podľa Haiga v behaviorálnych vedách pomerne málo využívaná – metóda exploračnej faktorovej analýzy. Haigova filozofia kvantitatívnych metód sa, podobne ako „*nazeranie na jazyk so zreteľom na jeho vzťah k človeku a ku kultúre*“ (Dolník, 2010b, s. 11), opiera o antropologickú perspektívu, čo je zrejmé: 1. z toho, ako Haig zdôrazňuje potrebu účinného zobrazenia dát (*effective graphical display*; Haig, 2018, s. 18) i dôležitosť odhalenia ich štruktúry (*importance of revelation*; tamže, s. 20); 2. v antropologickej téze o primárne vizuálnom charaktere ľudskej vedomej skúsenosti v zmysle „*raz vidieť je lepšie ako stokrát počuť*“ (*picture is worth a thousand words*; tamže, s. 20); 3. v explicitnom tvrdení, že ľudia sú primárne vizuálne bytosti (*primarily visual creatures*; tamže). Prezentované štatistické metódy človeku napomáhajú k hľadaniu zaujímavých vzorov v spracovaných dátach (*interesting patterns in the data*; tamže, s. 21), pričom Haig tu odkazuje na koncepciu výskumného procesu dátovej vedy podľa J. Tukeyho (1980). Ten vychádzal z predstavy, že pri výskumnom procese zohráva rozhodujúcu úlo-

hu to, aby sa návrh výskumnej otázky takpovediac stvárnil v priebehu práce s materiálom.³²

Haig sa tiež v práci venuje porovnávaniu niektorých základných typov metód: induktívnej, deduktívnej i abduktívnej, pri ktorej uvádza Tukeyho rozlíšenie exploratórnej (*exploratory*) a konfirmatórnej (*confirmatory*) fázy výskumu, čím sa ďalej rozširuje model vedeckého výskumu. Cieľom štatistického výskumu je odhaliť javy definované ako „robustné empirické pravidelnosti“ (*robust empirical regularities*). Javy sú pomerne stabilné, opakujúce sa všeobecné vlastnosti sveta, ktorý sa usilujeme vysvetliť (*relatively stable recurrent general features of the world that we seek to explain*; tamže, s. 24). Ďalším krokom štatisticky podporovaného výskumu je hľadanie príčin, ktoré by vysvetlili výskyt javov.³³ V rámci metodologických úvah Haig ďalej uvádza termín „explanačná teória“: tá predstavuje konštrukciu, ktorú si vytvárame s cieľom pochopiť empirické javy (tamže, s. 27). Proces odhaľovania vzorov (*pattern detection process*; tamže, s. 28) v týchto empirických javoch prebieha v rámci viacúrovňového modelu analýzy dát (*multistage model of data analysis*; tamže, s. 24).

Opäť sa teda ukazuje, že odhaľovanie príčinných súvislostí sa okrem interpretačného výkonu subjektu v nezanedbateľnej miere opiera o odhalené vnútorné štruktúry samotných dát. Tie akoby „napovedali“, ako môžu jednotlivé oblasti skúmaných dát príčinne súvisieť, a odhaľujú tak, do istej miery, okrem javovej, zjavnej stránky skutočnosti aj jej ukrytú, nezjavnú stránku.

³² „We ‚begin‘ not with a properly formulated question, but an idea of a question that cannot be given an answer until it is specified in terms of appropriate constraints“ (Haig, 2018, s. 23).

³³ „Once detected, these phenomena are explained by abductively inferring the existence of underlying causes that are thought to give rise to the phenomena“ (tamže).

8. Korpusová lingvistika v celku prístupov k jazyku

Korpus ako elektronický artefakt predstavuje objektívne existujúcu entitu, a je preto možným predmetom merania a vedeckého skúmania, východiskom uvažovania o jazyku (o jeho systéme i vo vzťahu k iným systémom), ako aj všeobecne poňatého „myslenia o reči“ a jej „procesuálneho modelu“ (Kořen- ský, 2014). Vďaka objektívnej skúmateľnosti elektronickej jazykovej vzorky sa vytvoril priestor na diskusiu medzi prístupmi k tejto vzorke. Spomenutá diskusia prebieha medzi jednotlivými prístupmi k jazyku, jednotlivými lingvistikami (kognitívnou, funkčnou, korpusovou, teoretickou, generatívnou lingvistikou, neuro-, socio-, psycholingvistikou a ďalšími) a týka sa otázky, či je poznávanie jazyka cieľom samým osebe alebo je tiež prostriedkom a východiskom pri poznávaní aj príbuzných systémov (psychických či sociálnych).

V prvom rade sa tu ponúka otázka, ako vôbec k jazyku pristupovať, aká je jeho povaha: je to samostatný systém, ktorý možno a treba študovať samostatne, alebo je súčasťou ďalších systémov, a treba ho študovať prostredníctvom nich? Je zrejmé, že jazyk slúži človeku v širšom rámci ľudskej komunikácie a životnej praxe, že nemožno odhliadať od jeho funkčnosti v celku individuálnej kognície a jeho sociálnej funkcie, zároveň sa však otvára otázka, či stačí jazyk redukovať na túto funkčnosť v rámci väčšieho celku, ostatných systémov. Otázku však možno aj otočiť: je skúmanie jazyka možné fundovane prepojiť so skúmaním ďalších systémov, kognitívnych, biologických, sociálnych, takže by sa štúdium jazyka stalo takpovediac bránou k poznaniu týchto systémov?

Korpusová lingvistika ako „nekompromisne empirická“ pozícia zrejme stelesňuje kritický postoj voči všetkým tvrdeniam týkajúcim sa vzájomného vzťahu „hmatateľnej“ jazykovej empirie, jej záznamov, a „nehmatateľných“ teoretických konštruktov, napríklad kognitívnych rámcov. Z pozície korpusovej lingvistiky sa formuluje otázka, na základe čoho možno tvrdiť, že vzťah jazykového systému a ostatných systémov sa javí práve týmto navrhovaným spôsobom, respektíve na základe čoho majú „nehmatateľné“ konceptuálne rámce práve navrhovanú podobu a štruktúru. Korpusová lingvistika by v tomto zmysle vyjadrovala postoj kritickej pozornosti voči všetkým potenciálnym tvrdeniam o jazyku, o jeho povahe, dynamike, javovej i systémovej stránke. Inými slovami, stelesňovala by funkčný korektív všetkých potenciálnych tvrdení o jazyku. Vzhľadom na svoju empirickú ukotvenosť by však pôsobila aj ako materiállová základňa, a teda zdroj týchto potenciálnych tvrdení.

V takto naznačenom zmysle vníma korpusovú lingvistiku Wolfgang Teu-

bert (2005), vychádzajúc z jazyka ako štruktúry, ktorá pozostáva z rôznorodých, systémových i nesystémových prvkov. K systémovým prvkom patrí podľa Teuberta predpokladaná existencia pravidiel a princíp analógie pri vysvetľovaní jazykových javov (napríklad závislostí medzi jednotlivými tvarmi a ich funkciami). K štúdiu konkrétneho jazyka zároveň možno pristupovať s predpokladom, že predstavuje príklad a nástroj istého univerzálneho systému medziľudskej komunikácie. K nesystémovým črtám jazyka podľa Teuberta patrí zasa do istej miery náhodnosť, idiosynkratickosť (*random idiosyncrasy*) až anomálnosť (*downright anomaly*) jazykových výrazov.³⁴ Túto nesúrodosť konštatuje aj Čermák: „Povaha jazyka je veľmi rôznorodá, je složený z formálnych prvků nestejněho druhu a řádu (tj. nejsou homogenní) a ty se nekombinují ani stejně ani pravidelně“ (2017, s. 101).

Ambíciou korpusovej lingvistiky je podľa Teuberta vyrovnáť sa práve s opisovaným „nepredvídateľným“ aspektom jazyka, uchopiť textúru skúmaného jazyka práve v jej distribučnej konkrétnosti, a to vychádzajúc z predpokladu, že ide o manifestáciu, prostriedok komunikačného systému, nech už ho lokalizujeme výlučne do ľudskej mysle alebo do sociálnej reality a jej diskurzívnych mechanizmov.³⁵ Aj v tomto prípade formuluje podobné zistenie Čermák, keď konštatuje, že zásadou korpusovej lingvistiky je analyzovať text dôsledne, ako artefakt svojho druhu, nie ako zdroj izolovaných výrazov v priebehu ďalšej interpretácie, ale v jeho celosti, systematicky a bez „vynechávok“ (2017, s. 17).

S takto naznačenou povahou jazyka a jeho nehomogénnym charakterom, sa tiež vyrovnáva Juraj Dolník, ktorý zdôrazňuje potrebu skúmania a vysvetľovania nehomogénnosti jazykových foriem pomocou explanačných princípov: „Deskripcia je podriadená explanačným princípom, na základe ktorých sa vysvetľuje stav tejto sústavy – nehomogénnosť. Opis je spätý s otázkou, prečo je tvarový systém taký, aký je, prečo nie je taký, aký by „rozumný“ systém mal byť, totiž pravidelný“ (Dolník, 2010a, s. 9). J. Dolník tiež poukazuje na svojbytnosť jazyka ako predmetu výskumu. Môžeme povedať, že ak by sa jazyk bezo zvyšku dal vysvetliť sociologicky, psychologicky, biologicky či prostredníctvom iného teoretického rámca, lingvistika by prišla o štatút svojbytnej vedy a jazyk by nebol samostatným systémom, ale zložkou a prejavom iných systémov (sociálnych, psychických, biologických): „Jazyk je predmetom vlastnej lingvistiky,

³⁴ „Languages are as much the result of dubious analogy, random idiosyncrasy and downright anomaly as they are a rule-based manifestation of a (hopefully systematic) universal language faculty“ (Teubert, 2005, s. 1 – 2).

³⁵ „This mortifying unruliness goes mostly unnoticed if I use the corpus only to extract the examples which fit with my hypothesis. Corpus linguistics goes further. It also wants to describe what cannot be explained“ (tamže, s. 2).

keď sa chápe ako sebestačný celostný útvar, ako štruktúra osobitného druhu“ (Dolník, 2009, s. 11 – 12).

V diskusii jednotlivých prístupov k jazyku o možnostiach poznávania jazyka a poznávania prostredníctvom jazyka vnímam prinajmenšom dve výrazné pozície. Jedna zdôrazňuje možnú konvergenciu jednotlivých lingvistík, a tak naznačuje možnosť skúmať systém jazyka v súvislosti s ďalšími systémami: túto pozíciu v hojných prácach zastáva skupina kognitívnych lingvistov (Gries – Stefanowitsch (eds.), 2006). Druhá pozícia zdôrazňuje komplementárny charakter jednotlivých „lingvistík“, ktoré však nie je možné zlúčiť. Túto pozíciu zastáva práve Wolfgang Teubert (2005).

Presvedčenie o možnej konvergencii teoretickej, korpusovej, kognitívnej a iných lingvistík vyjadrujú v početných prácach kognitívni lingvisti, sčasti sústredení okolo časopisu *Cognitive Linguistics*³⁶. Ich presvedčenie naznačujú už názvy časopiseckých prác: napríklad o „búrlivom“ (*love-hate*) vzťahu korpusovej a teoretickej lingvistiky (Gries, 2010), o možnej plodnej výmene poznatkov medzi teoretickou, kognitívnou lingvistikou a psycholingvistikou (Gries, 2012), o kľúčových bodoch diskusie v rámci takzvanej korpusovej kognitívnej lingvistiky (Arppe, 2010), či o otázke, čo kognitívna korpusová lingvistika môže a čo nemôže očakávať od neurolingvistiky (Blumenthal-Dramé, 2016).

O tom, že skúmanie takého komplexného javu, ako je jazyk, si vyžaduje kombináciu poznatkov z viacerých oblastí lingvistiky, je presvedčený aj Ray Jackendoff v rozsiahlej práci o „základoch jazyka“³⁷ a možnosť prepojeného skúmania jazykového systému a ďalších zložiek ľudskej kognície sa ďalej konštatuje v edičnej sérii *Lecture Notes in Morphogenesis*, napríklad v kolektívnej práci o „komplexnom rozmere jazyka“ a „emergencii významu“ (*language in complexity, emerging meaning*; La Mantia et al. (eds.), 2017) či v práci o „kvantitatívnej semiotickej analýze“ (*quantitative semiotic analysis*; Compagno (ed.), 2018). Ukazuje sa tu, ako je možné a prínosné prepojenie semiotickej sféry a domény vnímania, respektíve prepojenie kvantitatívnych metód a konceptu komplexnosti so semiosférou a neurofyziológiou.

V kolektívnej práci o kvantitatívnej semiotickej analýze sa uvádza termín „dynamický štrukturalizmus“ (*dynamic structuralism*), čo je pozícia, ktorej korene siahajú do sedemdesiatych rokov minulého storočia a ktorá sa spája s menom matematika Reného Thoma. Ten aplikoval takzvanú teóriu singularít na štrukturalistické uvažovanie v biológii a lingvistike (Petitot, 2018, s. 33).

³⁶ Dostupné na: <https://www.degruyter.com/view/j/cogl>.

³⁷ „Although my reformulation of grammar was motivated largely on grounds internal to linguistics, it turned out also to permit much more fruitful interactions with research in language processing, language acquisition, language use, spatial cognition, social cognition, evolutionary psychology, and neuroscience“ (Jackendoff, 2003, s. XII).

Prepojenie domén vied o jazyku a vied o živote sa naznačuje aj pri termíne neuromorfológia významu, analyzujú sa takzvané „plastické semiotické formy“ (*plastic semiotic forms*) a všeobecne „konštitúcia vizuálnych foriem (Sarti – Bamberi, 2018, s. 55), do súvisu sa kladie semiotika a vedy o zmyslovom vnímaní (*sensory sciences*) a skúma sa kategória významu chápaného ako isté rozhranie medzi textami a číslami.

V kolektívnej práci o komplexnosti a emergentnom charaktere jazyka sa zasa poukazuje na istú tradíciu chápania významu v dejinách lingvistiky. V nadväznosti na „slovníkový model významu“ (Jarošová, 2018a) možno konštatovať zmenu v chápaní povahy kategórií, gramatických pravidiel a významu jazykových jednotiek. Význam sa vo svetle korpusových analýz javí nie ako diskretná, ale skôr ako difúzna, vágna a kontextovo daná entita.³⁸ V ďalšom príspevku v uvedenej druhej kolektívnej práci sa zdôrazňuje, že význam jazykových výrazov závisí zároveň od „imanentných rozmerov“ výrazov (semiotických a jazykových štruktúr), ako aj od rozmerov, ktoré semiotiku prekračujú (zmyslové vnímanie, kognícia). Všetky tieto prekračujúce (*transcendent*) rozmery predstavujú komplexné štruktúry, závisiace od viacerých parametrov. Tvrdí sa, podobne ako v práci o „špirále poznania“ (Jarošová, 2018a), že lingvistický výskum by mal byť vo svojom ďalšom vývoji interdisciplinárne poučený, mal by sa rozvíjať v nadväznosti na zistenia napríklad neurovied (tamže, s. 139).³⁹ Vďaka prepojeniu semiotiky, lingvistiky a kognitívnych i ďalších vied v takzvanej teórii predikatívnych a enunciatívnych operácií (*theory of predicative and enunciative operations*) sa zasa významy výrazov, pojmy (*notions*), modelujú ako zložené entity a v čase premenlivé kognitívne štruktúry, ktoré z biologického i kultúrneho hľadiska existujú pri interakciách jednotlivcov, prostredia a artefaktov.⁴⁰ Na rozdiel od kognitívningvistického dôrazu na konvergenciu jednotlivých prístupov k jazyku zdôrazňuje Wolfgang Teubert (2005) skôr komplementaritu, a teda nezlučiteľnosť týchto prístupov. Svoju „verziu“ korpusovej lingvistiky predkladá v podobe dvadsiatich piatich téz, ktoré sa pokúsim zhrnúť.

³⁸ „Linguistic tradition has produced descriptions that can be assimilated (at least for the most part) to discrete models: discrete models of categories, discrete grammar rules and discrete lists of word senses, among others“ (Col et al., 2017, s. 29).

³⁹ „If semiotics does not wish to confine itself to pure immanent discipline [...], it must integrate variables [...] and think of them as constitutive of its object (sense)“ (Paolucci, 2017, s. 96).

⁴⁰ „[S]tructured sets (or bundles) of properties that include specific aspects of biological, physical and cultural life of human beings. Each of these properties is the cognitive indicator of interactions finely distributed between individuals and individuals, individuals and the environment, individuals and artefacts“ (La Mantia, 2017, s. 119).

Korpusová lingvistika sa podľa autora zameriava na kategóriu významu, ktorý Teubert chápe ako „to, čo sa odovzdáva v komunikácii medzi členmi diskurzného spoločenstva“ (*discourse community*), teda medzi členmi jazykového spoločenstva. Korpusová lingvistika sa nezaoberá psychologickými aspektmi komunikácie, predmetom skúmania v korpusovej lingvistike je zápis textu alebo prepis prejavu, teda toho, čo je reprodukovateľné. To, čo nie je zapísané alebo prepísané, alebo sa nezaznamená, je z pohľadu korpusovej lingvistiky stratené (Teubert, 2005, s. 3).

Teubert, podobne ako Čermák, podotýka, že každý segment textu, slovo, viacslovnú jednotku či frázu možno vnímať z hľadiska formy a z hľadiska významu. Forma reprezentuje význam, pričom význam bez formy, ktorá význam reprezentuje, neexistuje: význam existuje iba v diskurze a z hľadiska korpusovej lingvistiky je nezávislý od úmyslov používateľa jazyka.

Korpusová lingvistika je podľa Teuberta empirická veda, ktorej predmetom sú reálne jazykové dáta, diskurz je zasa celkom textov vytvorených v rámci diskurzného spoločenstva (tamže, s. 4). Túto tézu vnímam ako prejav korpusového empiricizmu, keďže sa skúma diskurz a materiálny význam komunikácie, ktorý je v podobe elektronických textov dostupný, na rozdiel od kognitívnych mechanizmov, ktoré nie sú priamo pozorovateľné. Jazykovedec by podľa Teuberta mal byť schopný najprv definovať a vymedziť svoj predmet výskumu, špecifikovať, ktoré jazykové dáta chce analyzovať a na základe týchto vymedzení, identifikovaných dát a metód ich skúmania a spracovania je potom možné formulovať všeobecné a špeciálne tvrdenia o diskurze. Všeobecné tvrdenia sa týkajú pravidiel alebo probabilistických predpovedí o „správaní“ sa skúmaných predmetov a javov v diskurze (odpovedá sa tým na otázku, ako „sa hovorí“ (*what is spoken*)), špecifické tvrdenia sa týkajú interpretácie typov textov alebo konkrétnych textových segmentov (tým sa odpovedá na otázku, prečo „sa hovorí“ istým spôsobom). Pre vzájomné vymedzenie vzťahu korpusovej, kognitívnej a ďalších lingvistík je dôležité najmä Teubertovo tvrdenie, že skúmanie jazykovej kompetencie je mimo dosahu korpusovej lingvistiky: korpusová lingvistika sa zameriava na javy, ktoré nemožno vysvetliť rekurzívne, prostredníctvom odkazovania na všeobecne platné pravidlá a tvrdenia (tamže, s. 5). Korpusová lingvistika, čím Teubert zrejme myslí hlavne jednu jej idealizovanú podobu označovanú

výrazom „corpus-driven“⁴¹, sa primárne zaoberá náhodnosťami, „idiosynkráziami“ (*idiosyncrasy*) pri používaní jazyka.

Zaujímavé sú i Teubertove tvrdenia o síce bežnej, ale nie samozrejmej schopnosti používať jazyk, pričom za normálnych okolností si podľa neho existenciu jazyka uvedomujeme iba pri poruche v komunikácii (tamže, s. 5), keď sa reč stáva problémom, ako som to naznačil v častiach o afatických poruchách či narušení ipseity. Pri skúmaní jazyka je významným nástrojom matematické, štatistické modelovanie a frekvencia, pomocou ktorých sa, ako podotýka aj Haig (2018), odhaľujú návratné vzory v dátach (*recurrent patterns*; Teubert, 2005, s. 5). Štatistická signifikantnosť spoločného výskytu slov však nepostačuje na to, aby sa na jej základe mohli formulovať všeobecné tvrdenia o diskurze, lexikálne jednotky treba navyše sémanticky interpretovať.

Ďalšie tvrdenia sa týkajú pojmu význam. Ten Teubert charakterizuje ako „parafrázu“ (tamže, s. 7), neexistuje v istom zmysle žiaden „pravý“ a fixný význam nejakej lexikálnej jednotky: každý môže parafrázovať jednotku významu podľa vlastného uváženia, preto je význam akejkoľvek lexikálnej jednotky vždy provizórny.⁴² Diskurz predstavuje „autoreferenčný systém“: prirodzený jazyk je jediný kódovací systém, v ktorom funkcie prvkov nie sú určované zvonka, ale interne, zo strany používateľov jazyka a prostredníctvom používania jednotlivých výrazov, preto sa prirodzené jazyky odlišujú od formálnych kalkulov, akým je matematický kód (tamže). Teubert tu vychádza z danosti a ontologickej prvotnosti diskurzívnej reality, nie tej pred- alebo mimodiskurzívnej. Keď hovoríme, neodkazujeme k realite mimo diskurzu, ale k tomu, čo bolo povedané predtým. Keď takto rokujeme (*negotiate*) o význame textového segmentu,

⁴¹ V prístupe ku korpusovému materiálu možno rozlíšiť dve pozície, ktoré však predstavujú iba rozličné miery, do akých chce lingvista vychádzať z materiálu a jeho textu, respektíve zo svojej intuície a vopred stanovenej koncepcie a hypotézy. Pôvodné rozlíšenie korpusového výskumu typu „corpus-based“ a „corpus-driven“ (Tognini-Bonelli, 2001) sa týka miery, do akej sa korpusový materiál využíva na formulovanie výskumných hypotéz: „*Vlastní lingvistický výzkum korpusu má dvojí povahu podle toho, jak a do jaké míry chce záviset na datech získaných z korpusu. Je-li na nich závislý úplně, tj. pracuje pouze s daty z korpusu (obvykle velkého a reprezentativního), jde o (A) typ přístupu corpus-driven (korpusem řízený, daný). Takový přístup předpokládá, že se zpracují všechna relevantní data korpusu a interpretuje se až výsledek jejich analýzy jako celek. Je-li však závislý na datech jen podpůrným způsobem, kdy si lingvista na základě své intuíce a předem zformulované hypotézy, popř. teorie vhodná data a příklady z korpusu jen selektivně doplňuje, jde o (B) přístup corpus-based (na korpusu založený[...]), ostrá hranice mezi obojím však nebývá a v praxi se oba přístupy často doplňují*“ (Čermák, 2017, s. 100). K rozlíšeniu oboch prístupov pozri tiež Cvrček – Kovářiková (2011, s. 122 – 124).

⁴² „*There is no true and no fixed meaning. Everyone can paraphrase a unit of meaning however they like, therefore the meaning of any lexical item type is always provisional*“ (Teubert, 2005, s. 6 – 7).

robíme to vnútri diskurzu, nie zvonka neho alebo nad ním. Diskurz podľa Teuberta neobsahuje, wittgensteinovsky povedané, súkromný jazyk, žiadne skúsenosti z perspektívy prvej osoby ani subjektívne skúsenosti, realita mimo diskurzu môže do diskurzu vstúpiť iba v podobe svedectva (*testimony*), prepojenie medzi diskurzom a realitou mimo diskurzu nie je súčasťou jazykového systému, je vytvárané každým členom jazykového (diskurzívneho) spoločenstva individuálne. To znamená, že korpusová lingvistika neuvažuje o tom, či jazykový výraz vyjadruje mimojazykovú skutočnosť (obsah mysle alebo časť sveta), ale skúma význam iba ako entitu viazanú na formu. Tému, že význam je parafráza, vysvetľuje Teubert tak, že korpusová lingvistika nerozlišuje medzi lexikálnym významom a encyklopedickým významom: význam slova *citrón* je podľa neho určený všetkými zaznamenanými tvrdeniami o citróne.^{43, 44}

Korpusová lingvistika však napriek tejto empirickej prísnosti podľa Teuberta nie je vedou, ako sú prírodné vedy: lingvistika patrí k humanitným oblastiam, a ako taká je prejavom ľudského úsilia objaviť zmysel vo svojom okolí i vnútre. V rámci korpusovej lingvistiky je legitímnym výsledkom hľadania takého zmyslu jeho interpretácia, nie verifikácia: korpusový lingvista nemá privilégium experta, ktorý by mal rozhodovať o tom, čo je prípustné a čo nie, je súčasťou spoločenstva, a ako taký predkladá svoje zistenia členom spoločenstva a argumentuje v prospech prijatia týchto zistení. V diskurzívnej spoločnosti má každý člen právo do diskurzu prispieť, zákonitosti diskurzu sa preto nevytvárajú riadene a v tomto zmysle azda možno povedať, že diskurz sa „organizuje sám“ (*organizes itself*; tamže, s. 8). V závere článku Teubert napokon explicitne vymedzuje svoj postoj k otázke vzťahu korpusovej a kognitívnej lingvistiky: považuje ich za dve komplementárne, ale absolútne nezlučiteľné paradigmy, kognitívnu lingvistiku zaujímajú „mysle hovoriacich a počúvajúcich“ (*minds of speakers and hearers*; tamže, s. 9) a táto oblasť sa práve korpusovej lingvistiky „netýka“, korpusová lingvistika sa na jazyk pozerá zo sociálneho, nie psychologického hľadiska. Význam jazykovej komunikácie vidí Teubert v zvyšovaní komplexnosti ľudskej spoločnosti, čo možno interpretovať aj ako inú formuláciu myšlienky, že sa ľudská skutočnosť rozrastá o nové prvky, vrátane (jazykových) technológií.

Teubertove tézy považujem za vyjadrenie primátu sociálneho rozmeru jazyka, za prihlásenie sa k myšlienkam filozofického empiricizmu a kritického racionalizmu, i za príspevok k úvahám o možnej konvergencii jednotlivých

⁴³ „Corpus linguistics does not distinguish between lexical meaning and encyclopaedic meaning. The meaning of the unit lemon is everything that has been said about lemons. Lexical items and what they stand for are discourse objects (and not objects of the ‘real world’), constructed through the contributions of the members of the discourse community“ (tamže, s. 7).

⁴⁴ Ďakujem recenzentovi, prof. Zouharovi, za návrh úpravy pôvodnej verzie tohto tvrdenia.

oblastí lingvistiky. Myšlienku, že taká konvergencia má zmysel a dokonca je faktom, zastávajú aj korpusoví lingvisti ako Tony McEnery a Andrew Hardie, pričom konvergenciu konštatujú medzi funkcionalistickou, korpusovou lingvistikou a psycholingvistikou (2012, s. 192 – 224). A. Jarošová, podobne, vidí možnosť vzájomného dopĺňania sa teoretických rámcov, ako sú štrukturalizmus, štruktúrny funkcionalizmus, kognitivismus, korpusová lingvistika (2018a, 2018b) a symbiotický potenciál jednotlivých rámcov označuje výrazom „špirála poznania“ (2018a): *„nie je potrebné vzdať sa užitočných súčastí funkcionálno-systémového hľadiska, treba využívať explanačný potenciál konceptu funkcia a skúmať, ako vplyva na utváranie významu a formy“* (tamže, s. 139). O všeobecnej konvergencii systémových odborov zasa z holistického a systémového hľadiska uvažuje M. Svítek (2013, s. 176) a „optimistom“ v otázke zlučiteľnosti viacerých perspektív na povahu ľudskej kognície je i F. Tvrďý, ktorý však v práci o filozofických aspektoch umelej inteligencie naznačuje aj možnosť, *„že naše mozky jsou natolik složité, že je nebudeme s to úplně pochopit nikdy, a to materiálně ani funkcionálně“* (Tvrďý, 2014, s. 180).

Korpusovú lingvistiku aj na základe načrtnutej diskusie dvoch výrazných pozícií vnímam ako vyjadrenie istého empiristického a metodologického akcentu pri kritickej analýze každého tvrdenia o javovej stránke jazyka, o jej jednotkách (jazykových výrazoch a typoch výrazov či textov) a takto sprostredkované aj o povahe jazyka ako systému. Tento kritický akcent má podobu otázky, na základe čoho, akého materiálu a postupu sú formulované jednotlivé tvrdenia o jazyku. Ide vlastne o požiadavku, aby tvrdenia o jazyku ako empirickom fenoméne spĺňali nároky kladené na korektné tvrdenia v rámci vedeckého diskurzu: aby boli dostatočne empiricky i metodologicky fundované, verifikovateľné či falzifikovateľné. Načrtnutá požiadavka predstavuje potom podobu petříčkovského „pohybu myslenia“: v tomto prípade pohyb smeruje „od introspekcie ku korpusovo podloženým, „informovaným“ tvrdeniam“ (*from introspective to corpus-informed judgments*; Desagulier, 2018, s. 1). Inými slovami, korpusovú lingvistiku možno vnímať ako potenciálne interdisciplinárnu exploračnú „cestu k dobre podloženým a spoľahlivým tvrdeniam o reči“ (*Wege zu wohl dokumentierten und verlässlichen Aussagen über Sprache*; Eichinger, 2007, s. 1).

Oblasť korpusovej lingvistiky je tiež vzhľadom na svoju rôznorodosť a ukotvenosť vo „všadeprítomných“ technológiách prirodzene prepojená aj s oblasťami počítačového spracovania prirodzeného jazyka, umelej inteligencie, digitálnych humanitných vied, dátovej vedy a trendmi v budúcom spracovaní týchto dát. Medzi také trendy patrí dôraz na multimodálnosť korpusov v budúcnosti, vizualizáciu (Čermák, 2017, s. 68), ako aj objavovanie čoraz ďalších vrstiev dát a súvislostí medzi nimi. V rámci digitálnych humanitných vied sa zasa prepája výskum, vývoj i celkový spôsob výskumného a prak-

tického života, životná prax, čoho výsledkom sú i nové spôsoby myslenia, napríklad myslenie „prostredníctvom“ (*through*) médií.⁴⁵ V tejto súvislosti si možno, podobne ako F. Čermák, položiť otázku o ďalšom smerovaní korpusovej lingvistiky, o jej budúcich úlohách a podobách tak, ako ich formulujú predstavitelia korpusovej lingvistiky, počítačového spracovania prirodzeného jazyka i dátovej vedy, a to v úvahách o razantnom rozvoji takzvaných hlbokých neurónových sietí a oblasti Deep Learningu.

V. Dhar (2013, cit. podľa Wladawsky-Berger, 2014) definuje dátovú vedu ako systematické štúdium spôsobov na získavanie poznatkov z dát a porovnáva tiež rozvoj dátovej vedy so štatistickým strojovým učením (*statistical machine learning*). Kým analýza dát (*data analysis*) vysvetľuje nejaký jav tak, že sa získavajú vzory (*patterns*) z jednotlivých súborov dát, dátová veda sa, naopak, usiluje získať poznanie, ktoré človeku umožňuje sa podľa neho rozhodovať a konať. Súbory dát, s ktorými dátová veda pracuje, sú rôznorodé, heterogénne: texty, obrázky, videá. Dátový analytik dostupné dáta nielen jednoducho analyzuje, ale pokúša sa ich „sledovať z mnohých uhlov“, s cieľom objaviť použiteľné, „vykonateľné“ poznanie (*actionable knowledge*), na základe ktorého sa možno rozhodovať pri dosahovaní praktických cieľov.

Charakter nástrojov a predmetov poznávania sa teda mení: od fyzikálnych predmetov a prostredí sa pozornosť presúva na virtuálne reprezentácie predmetov, prostredí a systémov, čo znamená lepšiu aplikovateľnosť perspektív na tieto dáta, ktoré sú sčasti reprezentáciami fyzického sveta a sčasti samostatnými, virtuálnymi predmetmi. Kým fyzický predmet bol viazaný, napríklad svojimi fyzikálnymi vlastnosťami, danosťami fyzického prostredia, ako aj danosťami subjektu, tieto danosti síce ostávajú v platnosti aj pri skúmaní virtuálnych artefaktov, artefakty však už primárne vplyvu fyzického prostredia nepodliehajú. Mení sa aj predstava o zdroji teoretického modelu, o pôvode a povahe teórií, ktorými sa vysvetľujú javy v okolitej skutočnosti. Bežne si myslíme, že teória je výsledkom intelektuálnej invencie vedca, uplatnená na materiálny či abstraktný materiál, predstavujeme si, že pôvodom teórie je ľudská myseľ. Strojové učenie túto predstavu podľa V. Dhara obracia a za zdroj otázky (problému) i za zdroj riešenia tohto problému sa tu, obrazne povedané, považujú priamo dáta, respektíve počítač.⁴⁶ Algoritmy strojového učenia nám podľa Dhara pomáhajú objavovať vzory, ktoré majú „prediktívny“ (*predictive*) charakter. Platí, že lepšie vysvetlenie je také, ktoré vysvetlí skúmaný jav jednak

⁴⁵ „We think through, with and alongside media“ (Hayles, 2012, s. 1).

⁴⁶ „Given a large trove of data, the computer taunts us by saying ‚If only you knew what question to ask me, I would give you some very interesting answers based on the data.‘ Such a capability is powerful since we often do not know what question to ask“ (Dhar, 2013, cit. podľa Wladawsky-Berger, 2014).

jednoduchšie a jednak nielen opíše aktuálnu situáciu, ale dokáže predvídať aj priebeh udalostí v budúcnosti.⁴⁷

V súvislosti s rapidne sa rozvíjajúcou sa oblasťou Deep Learningu sa zasa, podobne ako pri iných perspektívnych zmenách v poznávaní, uvažuje o možnom prínose, či radikálnej premene povahy poznania, ktoré so sebou prinášajú nové technológie na spracovanie jazykových a iných dát. Ozývajú sa navyše hlasy, podľa ktorých explozívny nárast objemu spracúvaných a analyzovaných dát v spojení s Deep Learningom znamenajú koniec doterajších prístupov počítačového spracovania prirodzeného jazyka. Deep Learning podľa Christophera Manninga pripomína akúsi vlnu tsunami, ktorá hrozí, že zmetie počítačové spracovanie prirodzeného jazyka (NLP).⁴⁸ Podľa Manninga sa však netreba obávať, že Deep Learning nahradí NLP, a to z dvoch hlavných dôvodov: 1. podľa predstaviteľov strojového učenia je práve NLP oblasťou, na ktorú sa treba v budúcnosti zamerať, a 2. centrom záujmu NLP sú jazykové technológie, cieľom nie je optimalizácia strojového učenia ako takého, ale poznávanie jazyka (Manning, 2015, s. 702). Manning tak zdôrazňuje samostatnosť výskumných otázok jazykovedy⁴⁹, respektíve samostatnosť jazyka ako systému, pričom práve samostatnosť jazyka ako systému predstavuje podľa Dolníka (2009, s. 11 – 12) podmienku možnosti jazykovedy ako svojbytnej oblasti poznania.

⁴⁷ „To be useful both practically and scientifically, the patterns must be predictive. The emphasis on predictability typically favors Occam's razor, or succinctness, since simpler models are more likely to hold up on future observations than more complex ones, all else being equal“ (tamže).

⁴⁸ Vyjadrené ešte expresívnejším obrazom Neila Lawrenca, „NLP is kind of like a rabbit in the headlights of the Deep Learning machine, waiting to be flattened“ (Manning, 2015, s. 701).

⁴⁹ „It's not about the best method of machine learning – the central issue remains the domain problems. The domain problems will not go away“ (tamže, s. 702).

9. Korpusová lingvistiká vo svetle antropologických konštánt

Korpusovú lingvistiku sa v poslednej kapitole pokúsím interpretovať prostredníctvom niektorých filozofických charakteristík človeka a jeho poznávania, ktoré považujem za „antropologické konštanty“. Moja interpretácia spočíva v tom, že korpusový výskum chápem ako isté teoreticko-praktické vyjadrenie rôznych antropologických konštánt. Výraz antropologická konštanta používa vo svojich výkladoch Juraj Dolník a formuluje ju ako tézu, že „človek je behaviorálno-akčná bytosť, t. j. aktívna bytosť, ktorá funguje na základe koordinácie svojho správania (jej behaviorálna stránka) a konania (jej činnostná stránka; jej aktivita teda zahŕňa správanie a konanie, čiže činnosť)“ (2018b, s. 302).

Táto základná charakteristika sa rozširuje, napríklad, o základné aspekty ľudského poznávania. Antropologickú konštantu, tézu o základnej danosti človeka, takzvanej „prirodzenej zvedavosti“, nachádzame v práci *Perspectives on Corpus Linguistics* (2011), v ktorej sa ozýva úvodná téza Aristotelovej *Metafyziky*.⁵⁰ V rozhovore o filozofických a literárnovedných aspektoch korpusovej lingvistiky v tej istej práci zasa Bill Louw interpretuje rozvoj korpusovej lingvistiky ako príklad, vo všeobecnosti, záujmu o poznávanie a ideu pravdy.⁵¹ Záujem o pravdu možno vystopovať i v korpusovolingvistickom záujme o autentický jazykový materiál, o skutočnosť, a teda je možné interpretovať tento záujem ako potrebu mať prístup k tomu, čo je autentické. Dôležité je, že Louw filozoficky interpretuje i samotný základ, pôvod korpusovej lingvistiky. Ten vníma v období dolníkovskej antropologickej konštanty, v takzvanej „intelektuálnej zvedavosti“. Korpusová lingvistiká je podľa Louwa vyjadrením ľudskej potreby orientovať sa vo svete prostredníctvom modelov, vzoriek sveta.⁵² Pôvod korpusovej lingvistiky teda Louw vníma ako antropologicky danú potrebu zjednodušiť si rozľahlosť, zložitosť a pôvodnosť skutočnosti, upraviť ju prostredníctvom modelov a vzoriek sveta tak, aby bola disponibilná, aby sa dala uchopiť ako predstava, „obraz sveta“ (Heidegger, 2013), ako objektívna realita,

⁵⁰ „Humans are naturally curious beings in search of knowledge“ (Viana – Zyngier – Barnbrook (eds.), 2011, s. XI).

⁵¹ „Interest in CL has grown because everyone is interested in truth“ (Louw, 2011, s. 171).

⁵² „The roots of Corpus Linguistics are to be found within the [...] peculiar trait of human intellectual curiosity: a desire to sample the world and the ‚worlds‘ it purports to contain“ (tamže, s. 172).

objektívne jestvujúci predmet poznávania, ktorý je možné spracúvať a skúmať technologickými postupmi. Vďaka časozberným technológiám sa, napríklad, môžu diachrónne porovnávať jednotlivé podoby danej krajiny a možno ich tak vnímať ako reprezentácie reality súčasne, a tak si azda všimnúť niečo, čo by nám pri postupnom prežívaní týchto období v čase mohlo uniknúť. Z plynutia času sa však, zároveň, úplne neoslobodzujeme, a tak mohutný prísun informácií predstavuje výraznú črtu informačného prostredia, s ktorou sa musíme vyrovnávať.

Ďalšou formuláciou antropologickej konštanty je téza J. Viceníka a V. Černíka, že človek je hľadač vzorov, pravidelností, potrebuje sa vo svete orientovať, hľadať zmysel okolitého sveta, objavovať v ňom skryté pravidelnosti (vzory) a interpretovať ich: „Človek sa niekedy charakterizuje ako hľadač rozmanitých pravidelností a vzorov (*pattern seeker*): trendov, empirických generalizácií, zákonov, pravidiel“ (Černík – Viceník (eds.), 2005, s. 71). Prejavom tejto potreby je zrejme aj hľadanie pravidelností v autentickom jazykovom, korpusovom materiáli. Korpusový záujem o autentickú, rôznorodú a mnohorakú textúru jazyka a jej skryté štruktúry preto vnímam i ako prejav „interpretačnej nastavenosti človeka“. Tento termín používa Juraj Dolník pri opise procesu humanizácie človeka. Ako bytosť, ktorú obklopuje masívna rôznorodosť dát vrátane jazykových foriem, človek narážal na nemožnosť tieto formy explicitne uchopiť a systematizovať (hoci ich intuitívne používa v každodennom živote). V dôsledku istého pnutia medzi človekom a okolím (prírodou) dochádza teda k vzájomnému pôsobeniu medzi objektom a subjektom: „V [...] evolučnej premene humanizujúca sa bytosť navodzovala takú interakciu s prírodou, ktorá nadobúdala kvalitu vzájomného pôsobenia objektu (prírody) a subjektu (humanizujúceho sa živého tvora)“ (2018a, s. 77). Vzájomná interakcia objektu (rôznorodosti jazykových foriem) a subjektu (človeka usilujúceho sa pochopiť túto rôznorodosť a odhaliť jej skrytý systém) sa vyznačuje napätím, ktoré vyplýva z bezprostrednej neuchopiteľnosti jazyka, a preto je potrebné hľadať spôsob, ako tento neuchopiteľný charakter, unikajúcu rôznorodosť a bohatosť jazyka efektívne spracovať: „Vzchádzajúce napätie medzi objektom a rodiacim sa subjektom spočívalo v tom, že navodzovaná potreba a záujem subjektu rozumieť prírode narážala na jej inakosť [...]. Evolučná požiadavka ‚rozumieť tomuto inakému‘ si vynútila objavenie jazyka na premostenie priepasti medzi objektom a subjektom“ (tamže).

Subjekt musel objaviť „jazyk prírody“ (tamže), za ktorý sa tradične považuje matematika. V súvislosti s lingvistikou tak s odkazom na Leonarda Bloomfielda dochádza k jej matematizácii (Dolník, 2009, s. 12). Za obrazný jazyk, teda spôsob, ako sa najmä prakticky zmocniť nejakého výseku skutočnosti, možno označiť aj rôzne technologické riešenia, napríklad, v prípade lingvistiky, experimenty na pomedzí lingvistiky, neurovedy a kognitívnej vedy, ale tiež

rozvoj korpusových technológií a nástrojov. Tieto nástroje umožňujú efektívne spracovať obrovské množstvo dát, a teda umožňujú vytvoriť potenciálne reprezentatívnu vzorku dátovej a javovej rôznorodosti jazyka, v ktorej sa ukrývajú informácie, skryté súvislosti javov, a je potrebné doceliť, aby takto materiál „prehovoril“, získal svoj „jazyk“, predovšetkým jazyk matematiky. Dolník konštatuje, že cestou k objaveniu „prvku jazyka prírody“ je pozorovanie a registrácia (2018a, s. 78), čo sú postupy empirickej vedy. Jazyk prírody, v podobe intelektuálnych i technických nástrojov na zvládanie komplexnej skutočnosti, predstavuje „bohatstvo“ (tamže) a takisto pre subjekt otvára „bohatstvo“ sveta: v prípade jazykovej reality ide o bohatstvo jazykových štruktúr, javov a súvislostí, ktoré možno skúmať na základe korpusu. Dolník v tejto súvislosti hovorí i o evolučnom „údele“ človeka ako interpreta sveta (tamže). Motiváciou termínu „interpretačná nastavenosť človeka“ (tamže, s. 77) je zrejme antropologická téza, že človek je „nastavený na interpretáciu“, čoho iným vyjadrením je, že sa človek potrebuje orientovať vo svete, ktorý ho obklopuje, potrebuje vnímať jeho kontinuitu a uchopovať jeho súvislosti (tamže, s. 79). Iným vyjadrením interpretačnej nastavenosti človeka je, že jednotliviec hľadá vo svete orientačné body, ktoré vníma ako znaky: *„Uzavreté vnútorné svety komunikantov sa dajú prepojiť len znakmi, a tak sú komunikanti odkázaní na interpretáciu [...] človek je na interpretáciu nastavený, čo zahŕňa, že je formovaný na znakové (semi-otické) vnímanie sveta“* (tamže, s. 80). Korpusová lingvistika je potom jednou z podôb, ako sa človek vzťahuje ku svetu prostredníctvom znakov. Takýto semiotický a antropologický pohľad na korpusovú lingvistiku sa reflektuje aj vo fenomenologicky orientovanej filozofii technológie, pričom špecifikom fenomenologického myslenia, a teda akýmsi prienikom jednotlivých fenomenologických modelov a pozícií je, že fenomenológia skúma podmienky, vďaka ktorým sa nám veci javia také, aké (pre nás) sú.⁵³

V nadväznosti na antropologický výklad pôvodu korpusovej lingvistiky u Louwa a Dolníka chcem uviesť i filozofickú reflexiu samotného antropologického myslenia ako prizmy, cez ktorú nazeráme na človeka a jeho jazyk, a tiež filozofickú reflexiu takzvanej „bytnosti“ techniky a výskumu vo všeobecnosti. Takúto filozofickú reflexiu podáva Martin Heidegger v spise *Věk obrazu světa* (2013). Vychádza zo „zabudnutosti bytia“ (*Seinsvergessenheit*), čo je podľa Heideggera jav príznačný pre metafyzické nastavenie západného myslenia. Toto myslenie vníma ako rigidné, umŕtvujúce, keď tvrdí, že antropológia predstavuje *„takový výklad človeka, ktorý v základu už ví, čo človek je, a tudíž se nemůže ptát, kdo člověk je“* (Heidegger, 2013, s. 59). Metafyzika zasa slúži ako východisko filozofického myslenia: umožňuje zamyslenie sa nad „bytnosťou“

⁵³ „Phenomenology investigates the conditions of what makes things appear as such [as that which we take them to be]” (Ihde, 2003, s. 133, cit. podľa Introna, 2017).

súcna a rozhoduje sa v nej o tom, čo budeme považovať za (pravú) skutočnosť, za pravdu: „*V metafyzike dochází k zamýšlení nad bytností jsoucího a rozhoduje se v ní o bytnosti pravdy*“ (tamže, s. 7). Bytostné zamyslenie v zmysle uvažovania o základoch nejakej oblasti poznania a konania, napríklad v špeciálnych vedách, sa potom definuje ako: „*odvaha učinit tím nejproblematictějším pravdivost našich předpokladů a oblast, v níž leží cíle našeho usilování*“ (tamže). Heideggerov výklad sa ďalej zameriava na „bytostné“ javy, medzi ktoré patrí aj novoveká veda a strojová technika, a teda, eventuálne, aj moderné elektronické technológie a spracovanie jazyka: „*K bytostným jevům novověku náleží věda. Stejně významným jevem je i strojová technika*“ (tamže). Strojová technika podľa Heideggera predstavuje „výhonok bytnosti novovekej techniky“ a zároveň akési zhmotnenie novovekej metafyziky (tamže, s. 8).

Dôležité pri výklade korpusovej lingvistiky je to, čo Heidegger označuje ako „bytnosť vedy“. „Bytnosťou“ vedy je podľa neho výskum, ktorý má tiež svoju bytnosť: „*V čem spočívá bytnost výzkumu? V tom, že poznání je situováno jako postup do oblasti jsoucna, tedy do oblasti přírody nebo dějin. Postup zde neznamená jen metodu, neboť každý postup již potřebuje nějaký otevřený okresk jsoucna, v němž se pohybuje. Ale právě otevírání takového okrsku (zvýraz. M. Z.) je tím základním, co musí výzkum při svém postupu učinit. To se děje tak, že pro určitou oblast jsoucna, napr. pro přírodu, se vytvoří základní rozvrh toho, co jsou přírodní procesy*“ (tamže, s. 10). Vytvorenie základného rozvrhu toho, čo sú prírodné procesy, možno interpretovať ako vytvorenie objektívne skúmateľnej prírody, a teda objektívnej reality. Príroda a prírodné procesy predstavujú objekty, súčasti objektívnej reality, na rozdiel od pred-vedeckého sveta vecí: „*V tomto svete neexistují nijaké objekty, iba věci*“ (Markoš, 2015, s. 68). Objektívna realita je podľa Markoša „vynálezom“, čo však neznamená, že je subjektívnym konštruktom a neexistuje objektívne. Predstavuje platformu, na základe ktorej je možné intersubjektívne overovať tvrdenia, odkazovať na takpovediac spoločnú, externú skutočnosť:

„*Homo sapiens sa oproti zvyšku biosféry vyznačuje dvomi vynálezmi. Jedným z nich je objektívna realita. Vedci často hovoria, že všetko okolo nás je objektívna realita. Nie je. Objektívna realita je niečo, čo sme si sformovali my. Môžeme v nej s ľubovoľnou presnosťou určovať akékoľvek veci. Existujú abecedy, nuly, jednotky, čísla, kalkuly, formálna logika a v rámci nich existujú modely. Tu vo svete hmotný bod na nehmotnej niti nemožno zostrojiť, to sa dá iba vo virtuálnom svete objektívnej reality*“ (tamže, s. 69).

Poznávanie si tak vyžaduje istú disponibilitu predmetov poznávania, čo je nemožné v podmienkach jednosmerného plynutia času v herakleitovskej rieke, v podmienkach, keď sa v bezprostrednej skúsenosti objavujú (emergujú) jedinečné, neopakovateľné udalosti, ak ich je vôbec možné identifikovať: „*Vedecké poznanie sa zameriava, prirodzene, na to, čo je opakovateľné*“ (Miko, 1970, s. 40).

Korpusový výskum ako zhromažďovanie a spracovanie písaných, hovorených či inak zaznamenaných prejavov ľudskej činnosti potom v nadväznosti na vyššie uvedené tézy interpretujem ako jednu z podob výskumu, tak ako ho charakterizuje Heidegger. Jazykový úzus, teda autentické prípady použitia jazyka a uzuálne zaužívané spôsoby toho, ako organizovať text a kombinovať jazykové výrazy, predstavujú heideggerovský „rozvrhnutý okrskok predmetov“, ktoré korpusová lingvistika „necháva vystupovať v celej ich vrstevnatosti a spletnosti“, keďže vychádza z „nekompromisného empirizmu“ a hľadanie otázok teda podmieňuje existenciou jazykového materiálu. Práve ten necháva „vystať“ v jeho pôvodnej rôznorodosti: *„Věda se stává výzkumem, když při rozvrhování a zajišťování svého předmětného okrsku postupuje s [...] přísností [...]. Má-li být rozvržený okrsek okrskem předmětu, pak je třeba nechat jej vystoupit v celé jeho vrstevnatosti a spletnosti“* (Heidegger, 2013, s. 13). Vedecké skúmanie „núti súcno“, v tomto prípade jazyk, aby sa „podrobilo“ predstave, v tomto prípade v podobe spracovanej vzorky jazyka, skúmateľného artefaktu: *„Poznání jakožto výzkum nutí jsoucno, aby vydalo počet o tom, jakým způsobem a jak dalece je lze dát k dispozici představě. Výzkum disponuje jsoucnam tehdy, když je buď s to dopředu vypočíst, jak se bude jsoucno chovat, anebo dokáže zpětně dopočítat, jak se v minulosti chovalo“* (tamže, s. 22 – 23). „Vypočítavanie“ v prípade korpusovej lingvistiky znamená štatistickú analýzu jazykových dát, na základe ktorej sa získavajú informácie o nich a umožňuje sa opis, vysvetľovanie jazykových prejavov v ich pôvodnej tvarovej rôznorodosti a distribučných pomeroch.

Novoveký človek si podľa Heideggera podmanil svet tak, že premenil svet na obraz (tamže, s. 33). Súcno (a teda i jazykové zdroje) je pre novovekého človeka vďaka technológiám k dispozícii ako obraz, prípadne ako spracovanie dát, alebo vizualizácia. Disponibilitu súcna sám Heidegger označuje najprv termínom *Machenschaft* a neskôr termínom *Gestell*, ktoré označujú „nadvládu bytnosti techniky“ a funkčnosti. Čo nemá funkciu, účel, čo sa nedá použiť ako materiál či zdroj energie, v tomto zmysle neexistuje, netvorí súčasť sveta, „je ničím“.⁵⁴ V súvislosti s disponibilitou súcna (konkrétne disponibilitou formálnych reprezentácií jazyka) potom možno uviesť myšlienku Julie Kristevy, ktorá sa z pozície psychoanalýzy a semiotiky vymedzuje voči formalizácii semiotickej sféry. Formalizované skúmanie reči zamerané na jej empirické stopy, záznamy reči (v našom prípade korpusové zdroje)

⁵⁴ „To, čo panuje v dnešnej spoločnosti, charakterizuje Heidegger pôvodne slovom ‚Machenschaft‘, mohli bychom snad říci: manipulace vším jsoucnam, k němuž náleží také člověk. Později používá slovo ‚das Gestell‘ a hovoří o nadvládě bytnosti techniky. Jsme v situaci, kdy být znamená být prostředkem pro účely, být materiálem či zdrojem energie v řetězci prostředků a účelů, co takové není, není vlastně ničím“ (Michálek, 2017, s. 9).

predstavuje podľa Kristevy akúsi „umŕtvujúcu“ podobu myslenia, vzťahovania sa k reči a ku svetu:

„Myšlenky archivárů, archeologů či nekrofilů – takové jsou filosofie řeči v naší době, ztělesnění Ideje, fascinované tváří v tvář pozůstatkům diskursivního (mimo jiné) fungování a nahrazující tímto fetišem to, co jej vyprodukovalo. Egypt, Babylón, Mykény: jejich pyramidy, cihly s vyrytými výjevy, jejich fragmentární kódy vidíme v diskurzech našich současníků a domnívame se, že si je přivlastníme, když je budeme systematizovat [...]. Myšlení, setrvale hledající pravdu řeči tak, že formalizuje výpovědi visící v prázdnotě [...]. Tyto archivářské, archeologické a nekrofilní procedury, jež mohly vytvořit základ pro jistý vědecký požadavek – budovat úsudky na základě empirie, systematizovatelné danosti, pozorovatelného předmětu: řeči“ (2004, s. 9).

Systematické spracovanie záznamov jazykovej činnosti tak Kristeva považuje za istý odvrat od živého myslenia o reči a príklon k tomu druhu skutočnosti, ktorý je daný, pozorovateľný, systematicky opísateľný. To, že je empirický predmet trvale a kedykoľvek k dispozícii skúmajúcemu subjektu, sa prinajmenšom na jednej strane považuje práve za hodnotu tohto predmetu, za istú podobu a zdroj pravdy, za „pravdivosť“ (autenticitu) tohto predmetu: myslenie sa takto odvracia od toho, čo je pominuteľné, neopakovateľné, a zameriava sa na to trvalé, opakovateľné, na to, čo je podľa F. Mika základom vedeckého poznávania. Dalo by sa povedať, že ide o vyjadrenie tézy, že pravdivé, skutočné a hodnotné je to, čo je trvalé, nepominuteľné, za každých okolností jestvujúce. Na druhej strane sa však, aspoň podľa J. Kristevy, pohľad skúmajúceho subjektu nezameriava na to, čo tieto disponibilné stopy napríklad reči vytvorilo, vyhyba sa pohľadu na živý zdroj týchto stôp. To zasa možno chápať ako vyjadrenie tézy, že nanajvýš pravdivé, skutočné a hodnotné môže byť práve to pominuteľné, krehké, vzácné, jedinečné. Tento rozdiel v chápaní pravdy, skutočnosti a hodnoty pripomína rozdiel medzi trvalosťou a neživosťou anorganického materiálu a pominuteľnosťou života. Poznanie zamerané iba na to, čo je za každých okolností k dispozícii, sa podobá zbieraniu, triedeniu a skúmaniu suchých listov: je možné nimi manipulovať, vytvárať ich skupiny napríklad podľa farby, a tak opisovať ich formálne vlastnosti, ide však podľa Kristevy o predmety konzervované, nehybné, neživé, zavesené „v prázdnote“, o konzervujúce a umŕtvujúce „procedúry“, o odklon od toho, čo je živé.

Voči takémuto výkladu systematického a formalizovaného poznávania však možno uviesť opačný výklad, podľa ktorého empirický výskum, s ním spojená systematizácia i prípadná formalizácia poznania predstavuje spôsob, ako odkrývať ďalšie vrstvy skutočnosti, a teda spôsob, ako sa k skutočnosti, naopak, prikloniť. Korpusovú lingvistiku s jej dôrazom na empirickú danosť jazykového materiálu vnímam práve ako príklon k skutočnosti, k zisťovaniu toho, čo by na základe dát mohlo byť systémom jazyka, a teda ako príklon

k viditeľnému svetu. Viditeľný svet a to, čo označujeme ako život, chápem ako istý prechodný stav medzi chaotickými stavmi, ktoré svojou komplexnosťou neumožňujú poznanie, štruktúru, ba v hraničných (napríklad schizofrenických) stavoch ani vnímanie štruktúrovaného a rozumom uchopiteľného „sveta“ prostredníctvom nejakého „ja“. Svet v tomto zmysle existuje, kým existujú hranice medzi jeho oblasťami, kým subjekt dokáže stanoviť rozhranie medzi prostrediami a vecami, identifikovať takpovediac ich obrysy či membrány, vymedziť funkčné klauzúry jednotlivých systémov, a to vrátane systému jazyka.

Takúto perspektívu na komplexnosť reality ponúka okrem iných oblastí biológia.⁵⁵ Viditeľný svet, ktorý možno vnímať, skúmať, opisovať a analyzovať, tvoria podľa iných zdrojov z formálneho hľadiska korelácie medzi interagujúcimi prvkami, javmi. Medzi jednotlivými javmi sa zasa vytvárajú systémy vzťahov, vzniká „kaskádová“ dynamika prechodov medzi jednotlivými stavmi systémov. Akýmsi „zákonom“ týchto procesov je pritom podľa J. Pearca (2017) čo najefektívnejšie uvoľňovanie energie.⁵⁶ Korpusové uvažovanie o jazyku v jeho pôvodnej vrstevnatosti a rôznorodosti je preto podľa mňa prepojitelné s procesuálnym modelom komunikácie (Kořenský, 2014; Jarošová, 2018a), ale tiež s filozofickou reflexiou pojmov ako čas, proces, udalosť, prítomnosť.

Proces predstavuje podľa M. Ajvaza dianie, ktoré je „vždy určitým spôsobom nasmerované či orientované“, pričom toto dianie sa prejavuje „určitým pravidelným sledom fází“, a je tiež spojené s ambivalentnosťou, keďže „směrování k rovnováze může být směrováním k řádu i směrováním k rozpadu“ (Ajvaz – Pauknerová, 2019, s. 15). Procesy nie sú navzájom izolované, navzájom sa prestupujú a zároveň platí, že „aby vyvstal proces, musí být vydělen prostřednictvím určitého zaměření pohledu“ (tamže). Proces, či už prírodný alebo kultúrny, tak vzniká až zameraním pohľadu, prostredníctvom vnímania subjektu. Viditeľný, a v tomto zmysle usporiadaný svet z tohto hľadiska predstavuje istý stav medzi chaotickými a neuchopiteľnými komplexnými stavmi: „zaměření

⁵⁵ „One noteworthy way biologists have been looking at living things has been in terms of the forms (morphology) and functions (physiology) that living things exhibit. Especially it is in terms of their anatomy or morphology that we observe living organisms. We find and see membranes that separate an outside from what exists inside of them. We see the forms of all sorts of anatomical structures with boundaries. And it is in morphological structures that we find that what is living (what functions physiologically) and that hides behind these fences. What is living is carrying on functions that separate what is living on the inside from its environmental ecology on the outside“ (Yoos, 2015, s. 98).

⁵⁶ „One can see that the visible, tangible Universe is the set of continuing, progressive correlations between interacting elements, forming systems of relationships. This occurs in a continuing cascade of phase transitions universe-wide. This cascade of phase transitions toward greater correlation of interacting elements, overall, in the universe as a whole, creates both an “arrow of time” and what we call visible “order” in the Universe. And [...] it involves energy minimization in forming the relationships – the correlations“ (Pearce, 2017, s. 1).

pohledu je spjato s různými měřítky, na nichž mimo jiné záleží, vyvstane-li proces, nebo událost“ (Ajvaz – Pauknerová, 2019, s. 15). M. Ajvaz tu nadväzuje na M. Nitscheho, keď takisto hovorí o spleti (procesov), ktorých je jednotlivý proces súčasťou, a tiež na J. Pearca, keď hovorí o „kaskádových rysoch“ procesov: „*jednotlivý proces je tedy součástí spleti procesů, která má kaskádové rysy*“ (tamže). Procesy sa ďalej členia vertikálne do vyšších a nižších, aj horizontálne, do fáz. Súčasťou procesov sú udalosti, ktoré Ajvaz definuje ako významné zlomy spojené s horizontálnym alebo vertikálnym členením (tamže). Aj v procese komunikácie sa pritom stretávame s udalosťami, jednak s „komunikačnými udalosťami“ (Jarošová, 2018a), ale tiež s ruptúrami v procese komunikácie, ako bolo naznačené na začiatku práce: tieto ruptúry majú podobu afatických, schizofrenických a iných narušení vzťahu medzi používateľom jazyka a jazykom.

V procesuálnom modeli, a takisto v súvislosti s pojmom udalosť, vystupuje i pojem fragmentárnosť, ktorý vyjadruje ten príznak procesu, že „*často nejsme schopni sledovat v celém jeho trvání. Setkáváme se buďto s částí (nebo částmi) aktuálně probíhajícího procesu, anebo nějakou částí (částmi) procesu už uplynulého. Tato fragmentárnost zjevování procesu (zvýraz. M. Z.) nemusí znamenat, že nám proces jako celek uniká. Je určitým způsobem spolupřítomný (zvýraz. M. Z.), neboť je s těmi částmi, s nimiž se setkáváme, označován (zvýraz. M. Z.)*“ (Ajvaz – Pauknerová, 2019, s. 19). Ajvaz tak dáva do súvislosti trvanie i prerušenie, fragmentárnosť procesu a jeho vystupovanie v podobe javu. Prostredníctvom termínu „spolupřítomnosť“ „zjavovania“ sa procesu s vnímateľom javu tiež naznačuje, že prítomnosť prekračuje daný okamih smerom do minulosti i do budúcnosti. Proces u Ajvaza „*vystupuje jako význam*“, pričom „*části, jež k němu poukazují, mají povahu znaků [...]. O znaku uplynulého procesu mluvíme jako o stopě; a obecně se dá říci, že minulost nezůstává jen v naší paměti, ale především také v „paměti“ samotného systému, který prochází daným procesem*“ (tamže). Proces odohrávajúci sa v čase má teda charakter predmetu, na ktorý odkazujú, „ukazujú“ jeho časti, takže majú povahu znakov. Proces predstavuje význam týchto znakov a znaky zasa predstavujú „stopy procesu“: na základe týchto „stôp“ možno rekonštruovať systém, stopy sú jeho „pamäťou“. Ajvaz tak prepája procesuálny model skutočnosti s chápaním neustáleho procesu semiózy v rôznych oblastiach ľudského poznávania a ľudskej praxe, pričom vedecké skúmanie sveta možno v tomto zmysle vnímať ako novalisovské „čítanie šifier“, ktorými sú človeku dané procesy a predmety.⁵⁷

⁵⁷ „*Odhalování procesů a událostí ze znaků a stop je činnost, kterou vlastně děláme neustále, vše, s čím se setkáváme, vystupuje jako znak minulých, současných a budoucích procesů a událostí [...]. V této souvislosti se objevuje metafora jazyka, písma, zapisování a čtení: historie se čte či luští z geologického zápisu a z fosilií [...]; z písma střepů [...]; povaha určitého světa se vписuje*

Svet podľa Ajvaza existuje ako prechodné miesto medzi svojím vznikom a zánikom, počas prechodu do iného stavu, a teda ako „kozmos“ v procese „seba-štrukturácie“ a „seba-utvárania“: *„Proces sebestrukturace je procesem utváření jsoucná a bytí v jejich celku – a vzhledem k bytostné vrstevnatosti a kaskádovosti je tento proces už vždy procesem vytváření mnoha různých jsoucn. Sebestrukturace je proces utváření kosmu a tento proces sebeutváření kosmu je vlastním bytím kosmu“* (2017, s. 88). Korpusovú lingvistiku ako oblasť skúmajúcu dostupné „stopy (jazykového) procesu“ preto vnímam ako analogickú „seba-štrukturáciu“ jazyka, tak ako sa táto štruktúracia zaznamenáva na dostupných nosičoch a spracúva dostupnými technológiami, čím sa subjektu, ktorý sa rozhodne pristupovať ku svojmu okoliu prostredníctvom vedeckej metodiky, takpovediac ponúka jazykový svet ako výsek objektívnej reality. V tomto svete možno objavovať *„zcela netušené skryté rady bytí“* (Wilczek, 2011, s. 16) a *„ohromujúci složitost života“* (Stewart, 2014, s. 9).

do podoby domácích nástrojů a dá se z nich vyčíst [...]; do fyzikálního procesu je vepsána jeho prehistorie [...]; připomíná se Novalisovo universální písmo šifer, jimiž se procesy vписují do vzhledu přírodních jsoucn“ (tamže, s. 19 – 20).

Záver

V práci som sa pokúsil uvažovať o filozofických aspektoch korpusovej lingvistiky, čo znamená, že som sa pokúsil formulovať, aké myšlienky z oblasti fenomenológie, filozofickej antropológie, filozofie technológie, epistemológie a príbuzných oblastí poznania možno vnímať na pozadí rozvoja korpusovej lingvistiky. Podobne koncipované uvažovanie o korpusovej lingvistike nachádzam okrem v práci spomenutej knihe rozhovorov s korpusovými lingvistami (Viana et al., 2011) v publikácii o „korpusovolingvistickom diskurze“, respektíve o základných presvedčeniach, na ktorých sa tento diskurz zakladá.⁵⁸ V uvedenej práci sa naznačuje, že sa korpusová lingvistika ako disciplína rozvíja a prináša nové nástroje, zdroje a aplikácie, inými slovami, že sa rozširuje poznanie o rôznych aspektoch a oblastiach jazyka, no napriek tomu je potrebné takpovediac meta-skúmanie, skúmanie tohto skúmania.

Takto formulovaný zámer potom pripomína uvažovanie o základných otázkach terminológie, ako ich charakterizoval Ján Horecký: „*Hoci sa terminologickým otázkam venuje primeraná pozornosť už niekoľko desaťročí, jednako sa ešte stále vynárajú základné otázky, o ktorých treba uvažovať znova a znova*“ (Horecký, 1974, s. 129). Aj tu sa naznačuje, že sa daná oblasť, v tomto prípade terminológia, nepochybne rozvíja, jej otázkam sa venuje primeraná pozornosť, a predsa sa objavujú základné otázky, nad ktorými sa treba zamýšľať „znova a znova“. Ich návratnosť, stálosť, pretrvávajúca prítomnosť signalizujú, že tieto otázky v istom zmysle prekračujú zmnožovanie faktického poznania, spracúvanie ďalších jazykových dát i vývoj nových technológií: týkajú sa skôr celkového zmyslu a smerovania terminológie ako vednej oblasti. Rozvoj faktického poznania tak ako keby takéto základné otázky neodstraňoval, ostávajú v platnosti nezasiahnuté týmto rozvojom, majú teda nadčasovú, a v tomto zmysle filozofickú platnosť. Vnímanie základných otázok špeciálnej oblasti poznania ako filozofických otázok nachádzam i u Jaroslava Peregrina, ktorý konštatuje, že základné otázky, v jeho prípade základné otázky logiky, predstavujú akési ďalšie „poschodie“ otázok, ktoré priamo logika nerieši, hoci sa jej podstatne týkajú. Týmto ďalším poschodím je práve filozofické uvažovanie o základoch jednotlivých vied.⁵⁹

⁵⁸ V anotácii práce sa uvádza: „*With an ever-growing body of corpus linguistic tools, resources and applications, it becomes increasingly important to reflect critically on the underlying assumptions that corpus linguistics is based on*“ (Čermáková – Mahlberg, 2018).

⁵⁹ „*Na určité úrovni znalostí se [...] studenti mohou začít ptát (a to právě tak v logice jako třeba*

Filozofické myslenie má, navyše, tú vlastnosť, že sa okrem predmetu svojho záujmu, ktorým môže byť jav, fakt, pojem či téza, zameriava aj samo na seba, na svoju vlastnú povahu, a tak sa potenciálne problematizuje. Problematizácia toho, že sa v úvahách zaoberáme základom nejakej oblasti, môže spočívať v tom, že sa ukáže, ako sa také hľadanie základov stráca v nejasnosti. Filozofický pohyb myslenia, ktoré smeruje k základom určitej oblasti, či už myslenia, poznania alebo aj vedomia, ale ústi práve do nejasnosti, obrazne vyjadruje Michal Ajvaz v práci o genetickej fenomenológii Edmunda Husserla s názvom *Cesta k pramenům smyslu*. Túto myšlienkovú cestu opisuje ako „*putování proti proudu řeky k jejím pramenům [...], přičemž není předem jisté, jestli nějaký pramen nalezneme – může se nám stát něco podobného tomu, co zažil Claudio Magris při pátrání po prameni Dunaje, když zjistil, že Dunaj vlastně vytéká z vodovodního kohoutku u jakéhosi stavení*“ (Ajvaz, 2012, s. 19). Samotné Husserlovo hľadanie základov našej skutočnosti a intencionálne zameranej skúsenosti zasa neústi ani tak do nejasnosti, ako skôr prekračuje bezprostredne danú ľudskú skúsenosť. Výkony transcendentálneho vedomia totiž prekračujú fakticitu, danosť nášho sveta: „*intencionalita v širším smyslu je nejen silou vládnoucí ve zkušenosti, působící tu v různých podobách a ustavující zkušenost jako takovou; má zároveň povahu explikace universálnější síly, která je starší než (lidská) zkušenost*“ (tamže, s. 29).

Fenomenologické hľadanie základov, skúmanie „konštitúcie“ predmetov a predmetnosti možno pritom vnímať i ako jednu z koncepcií toho, čo znamená filozofovať vôbec. Túto fenomenologickú koncepciu filozofie zaujímavo komentuje Claude Romano v rozhovore o vzťahu fenomenologickej a analytickej filozofie, kde problematizuje práve Husserlov zámer objaviť v transcendentálnom vedomí základy všetkých poznatkov jednotlivých vied. Romano tento zámer vníma ako „karteziánsky fundacionalizmus“⁶⁰, pričom komentuje i pojmy udalosť a zlom, ktoré súvisia s procesmi (akým je aj proces komunikácie), a tvrdí, že myslenie udalosti a zlomu má tradíciu prinajmenšom vo francúzskom myslení. Ajvaz v tejto súvislosti vníma udalosť do istej miery ako

v matematice či v biologii), proč byly za výchozí zvoleny právě ty pojmy, a nikoli jiné, jak přesně máme základním pojům rozumět, proč jsou za správné a adekvátní považovány právě ty a ty postupy, a nikoli jiné. [...] Jelikož takové zkoumání se týká samotných základů příslušné vědy, nemohou být už prováděna pomocí této vědy samotné, ale dostávají se na základnější úroveň, totiž na úroveň filosofickou. Proto mluvíme o filosofii logiky, filosofii matematiky, filosofii té či oné vědy či o filosofii vědy obecně“ (Peregrin – Vlasáková, 2018, s. 17 – 18).

⁶⁰ „Ovšem otázka, kterou si musí položit Husserlův čtenář na začátku 21. století, zní: Je třeba na kartesiánský fundacionalismus přistupovat? Vždyť který dnešní filosof vědy si myslí, že je na místě chtít založit objevy astrofyziky nebo teorie strun, evoluční biologie nebo neeukleidovských geometrií na pravdách, jež by náležely transcendentálnímu vědomí? Tento projekt naprosto zastaral. Má už jen dokumentární význam“ (Romano, 2018, s. 127).

súčasť procesov, ktoré majú prechodne dezintegračný, ale v zásade integračný charakter, prebiehajú ako rozvoj, rozpad a vznikanie nových „syntaxí“ jestvujúcich vecí (2017, s. 36 – 40). Romano zasa o naznačenej tradícii francúzskeho myslenia tvrdí, že podľa tohto myslenia nie všetko, čo sa deje, tvorí udalosť, a tak poukazuje na jeden z významov „udalosti“, „udalosť v druhom zmysle“, ktorá má skôr znaky ruptúry, a teda procesu bez opätovnej integrácie a vlastne bez vlastnej identity, podstaty. Ako príklad by sa azda dalo uviesť subjektívne vnímanie osobnej integrity v okamihu prechodu medzi mániou a depresiou v prípade porúch z maniodepresívneho spektra. Udalosť v tomto zmysle má znaky „neprijemného problému“ (*wicked problem*) a jej podobou môže byť i narušenie ipseity (*disturbance of minimal self*), teda strata a neobnovenie kontinuity vo vzťahu medzi subjektom a jazykom:

„[E]xistuje také jiná tradice, jež je ve francouzském myšlení stále živá a jež chápe událost jako něco, co je charakterizováno svými rysy singularity, zlomu a novosti. Takto proti sobě můžeme postavit fakt (co se obecně děje) a událost: to, co přináší zlom, co se definuje svou radikální novostí. Jde o úzus historiků, chcete-li, ale také celé řady filosofů, hlavně francouzských, od Bergsona k Deleuzovi nebo Badiouovi, abych uvedl některá jména. Ne všechno, co se děje, tvoří událost. Právě událost v tomto druhém smyslu mě zajímá“ (Romano, 2018, s. 130).

Odkaz na Bergsona v uvedenom citáte nie je náhodný. Naznačený „druhý zmysel“ udalosti, skúsenosť „nepredvídateľnej novosti“, predstavuje akýsi myšlienkový kontrapunkt k filozofickému mysleniu smerujúcemu k počiatku, princípu, nadhľadu a prehľadu. Filozofické myslenie, ktoré sa podľa E. Višňovského vyznačuje objavovaním súvislostí v širokom zmysle slova, je v tomto zmysle protikladom skutočnosti, neredukovateľnej na vedecky skúmateľnú objektívnu realitu a jej výseky, na heideggerovské „okrsky“ súcna. S neredukovateľnou komplexnou skutočnosťou súvisí pojem novosť, ktorý sa opäť spája práve s H. Bergsonom:

„Filosofie nepřipouští vznik nepředvídatelné novosti. Vznikání vykládá jako uskutečňování již hotových idejí, jejich rozvinutí v čase. Teze o předchůdnosti idejí však vznikání čehokoli nového de facto popírá. Takový účet vystavuje filosofům v roce 1920 Henri Bergson [...]. Kořen problému spatřuje v nekritickém používání intelektu, jehož přirozeností je uchopovat to, co je ve světě pravidelné a stálé, jemuž ale uniká, rašení novosti“ (Puc, 2017, s. 259).

Romano zároveň vníma filozofické, konkrétne fenomenologické myslenie ako naďalej prínosné a živé, nevníma ho ako prekonanú, spiatočnícku etapu myslenia, ale ako teoretický rámec, ktorý podľa A. Jarošovej „netreba opúšťať“ (2018a, s. 139):

„Pro mě ale fenomenologie nikdy nebyla něčím takovým: není to šik způsob, jak být ve filosofii konzervativní, nebo dokonce zpátečnický. Fenomenologie je pravý opak. Především je to metoda, již charakterizuje opravdová skromnost, a skromnost není v intelektuální oblasti žádná maličkost. Nemá tu místo konstrukce velkých metafyzických systémů; to nejlepší, co lze dělat, je usilovat o správný popis, což v prvé řadě znamená vytvářet v nějaké oblasti dobré pojmy a provádět jemná rozlišení“ (Romano, 2018, s. 134).

Korpusové skúmanie jazyka ako empirické skúmanie stôp jazykovej činnosti vnímam v nadväznosti na uvedenú Romanovu myšlienku ako obdobu fenomenologického úsilia o správny opis, vytváranie dobrých pojmov a jemné rozlišovanie. Korpusová lingvistika podľa mňa predstavuje v prvom rade úsilie o „správny opis“ jazykovej skutočnosti, štatistické metódy umožňujú korpusovým lingvistom navrhovanie „dobrých pojmov“ (termínov pre jazykové javy) a podrobná kvantitatívna i kvalitatívna analýza korpusových dát im zasa umožňuje formulovať jemné rozlíšenia v tvrdeniach o povahe jazyka či jazykov.

Jav, keď sa komplexnosť reality a jej dimenzionálnosť, redukovaná rôznymi nástrojmi a metódami, vracia v podobe komplexnosti poznatkov o tejto realite, sa prinajmenšom z holistického pohľadu systémovej teórie javí ako vlastnosť samotnej reality.⁶¹ Korpusovú lingvistiku som sa preto pokúsil vnímať ako istý epistemologický postoj voči tejto komplexnej skutočnosti, a teda ako istú podobu a spôsob správania sa človeka vo vzťahu k svetu, ako formu ľudského konania. Napríklad, podobne interpretuje J. Paitlová kriticko racionalistickú pozíciu K. R. Poppera a H. Alberta ako „*způsob myšlení a jednání*“ (2018, s. 15). Oblasť poznávania a spracúvania istej časti objektívnej reality vrátane jazykovej reality sa zasa podnetným spôsobom včleňuje do teoreticko-filozoficko-metodologického rámca i pragmatických, prakticko-projektívnych postupov pri vedeckom výskume v sérii slovenských prác⁶² a výskumnú činnosť interpretuje z celkovej perspektívy ľudského konania aj S. Juszczak: „*Vyjadrené vo vedeckých kategóriách sú výskumy jednou z foriem ľudskej aktivity, ktorá je zameraná na hľadanie a poznávanie odkrývaných javov, štúdium ich povahy a zdokonaľovanie infraštruktúry a funkcií skúmaných prostredí a systémov. Výskumná činnosť je súbor poznávacích a konceptuálnych procesov*“ (2003, s. 13).

⁶¹ „Lze popsat řadu stránek o tom, jaké vlastnosti bychom mohli systému přiřadit, přičemž vždy skončíme na úrovni dokonalosti či nedokonalosti popisu formou našeho přirozeného jazyka. Není divu, že také texty prof. Jaroslava Vlčka končily na hranici popsitelnosti větami a jeho odborné práce jsou neshadno čitelné, protože sám text je již formou sémantické reprezentace a popisu určitých systémových vlastností“ (Svítek, 2013, s. 32).

⁶² Černík – Viceník (eds.) 2004; Černík – Viceník (eds.), 2005; Černík – Viceník 2011, 2012; Černík, 2014.

V práci som vychádzal z presvedčenia, že hoci sa veda a jej rozvoj zvykne pomeriavať nárastom faktických poznatkov a schopností čo najefektívnejšie vysvetliť čo najväčší počet javov, pochopiteľný je aj fakt, že explozívny nárast dát a informácií sa stáva predmetom kritickej reflexie i podnetom špekulatívneho uvažovania. Usiloval som sa preto poukázať na koncepty a teórie z oblastí filozofie, najmä fenomenológie, ale aj z oblasti prírodných, spoločenských a humanitných vied, aby som načrtnol súvislosti medzi týmito podobami ľudského myslenia, poznávania a konania. Ústrednou myšlienkou v práci bolo, že ľudské myslenie, poznávanie a konanie sú možné v prípade, ak sa komplexnosť vonkajšej i vnútornej reality človeka udržuje v istých medziach.

Metódy korpusovej lingvistiky potom predstavujú spôsob, ako účinne znížovať dimenzionalitu a komplexnosť dostupných dát tak, aby v nich bolo možné objaviť javy, a tak postupovať od opisu a analýzy k interpretácii súvislostí týchto (jazykových) dát. V kritickej reflexii informačnej spoločnosti sa poukazuje na to, že množstvo dát a informácií, ktoré ľudská spoločnosť neustále produkuje, má okrem pozitívnych stránok (rozširovania poznania) aj výrazne negatívny dopad, pretože potenciálne paralyzuje ľudskú pozornosť. Získavanie nových poznatkov je preto potrebné vyvažovať prakticky, s ohľadom na udržateľné množstvo informácií v našom prostredí. V tejto súvislosti navrhuje M. Svítek termín „múdrosť systému“, ktorú definuje tak, že systém je múdry, ak dokáže 1. *vnímat a vyhodnocovat racionální i emotivní podněty (činný i jalový vstupní informační výkon) ze svého okolí*; 2. *přeměnit je ve své vlastní etické poznání a vědomí okolního světa i role sebe sama v tomto světě*; 3. *neustále uzpůsobovat své výslední etické chování (činný i jalový výstupní informační výkon) ve vztahu ke svému okolí* (Svítek, 2013, s. 196). Špekulatívne napájané uvažovanie o vývoji technológií zasa otvára priestor pre vízie budúcnosti, či už majú podobu „fyziky budúcnosti“ (Kaku, 2013), globálnej konvergence vedných odborov a špeciálne konvergence systémových vied (Svítek, 2013, s. 172 – 197) alebo ďalšieho vývoja oblastí spojených s rozvojom umelej inteligencie (Boström, 2017). N. Maxwell, naopak, poukazuje na potrebu skôr svítekovej „múdrosti systému“, respektíve múdrosti konania, jeho slovami „racionalizácie sociálneho konania“, než kvantitatívneho nárastu poznania o sociálnych systémoch.⁶³

Kvalitatívna premena sociálneho konania jednotlivca na racionálne sociálne konanie v tomto zmysle nenastáva v prvom rade mechanickým prísunom, zhromažďovaním nových poznatkov, ale skôr uplatnením výsledkov vedecké-

⁶³ „Social inquiry needs to be developed, not primarily as social science, but rather as social methodology or social philosophy. The primary task is not to improve knowledge of social phenomena; it is rather to help the social world implement progress-achieving methods generalized from those of science. Knowledge is needed to facilitate what really matters – rational social action“ (Maxwell, 2018).

ho poznávania na metódy sociálneho konania, pričom ide o metódy vo vyššie naznačenom zmysle, teda o „spôsoby myslenia a konania“. Takáto premena v procesoch, ktoré človeka obklopujú, nenastáva s imanentnou nevyhnutnosťou, pretože jednotlivec hľadajúci možnosti, ako zmeniť svoje konanie, sa, na jednej strane, nemôže oprieť o predpoklad vopred daného, metafyzického zákona premeny, čelí bergsonovskej „rašiacej“ novosti, a tak má prípadná kvalitatívna premena, istý posun v bytí, skôr podobu ruptúrneho „evolučného tápania“ (Markoš, 2017). Myšlienka teleológie ako vopred danej účelnosti vývoja zasa, na druhej strane, predstavuje istý metafyzický optimizmus, zbavuje vývoj potenciálne slepých uličiek a udalostí, odstraňuje kategóriu udalosti a novosti vo vyššie naznačenom bergsonovskom zmysle, a pripomína tak „*dávnou představu Boží mysli (sensorum dei), kterou obsáhnout nemáme šanci – a proto nám novinky připadají jako novinky*“ (tamže, s. 186).

V práci som chcel reflektovať fakt nesmierne rôznorodých aspektov jazyka a rovnako rozličných prístupov k jazyku: na jednej strane sa skúmajú jeho biologické i psychologické a sociálne rozmery, na druhej strane sa výskum zameriava na jednotlivé roviny jazyka, hoci zároveň možno povedať, že vedecký výskum sa opiera o istý všeobecný vedecký obraz sveta, napĺňaný poznatkami a postupmi mnohých vedných oblastí a tie sa navzájom viac alebo menej prestupujú. Lingvistický výskum v pravom zmysle slova sa tak zameriava na jazyk ako na samostatný predmet vedeckého skúmania. Aby však bolo možné jazyk skúmať vedecky, je potrebné pri analýze využiť aj poznatky a aparát prinajmenšom matematiky a logiky, pri interpretácii zasa zohľadniť teoretické rámce i poznatky oblastí zameraných na biologické, psychologické či sociálne aspekty jazyka. Lingvistický výskum navyše zakaždým vychádza aj z istých filozofických základov, prinajmenšom z filozofie jazyka a jazykovedy, a tak udržiava živý dialóg s filozofickými otázkami o povahe významu, znaku a ďalších lingvistických, semiotických a filozofických pojmov.

K povahe jazyka, ale nielen jeho, ďalej patrí, že má svoju objektívnu i subjektívnu stránku. Obsahom tej objektívnej je napríklad význam daného slova, ktorý sa môže vymedzovať či spresňovať a má intersubjektívnu platnosť. Obsahom subjektívnej stránky jazyka sú zasa napríklad emocionálne, estetické či iné asociácie, ktoré v nás dané slovo, veta, text, dielo vyvoláva. Takéto asociácie odkazujú na oblasť hodnôt a subjektívnej skúsenosti, ktorá sa nie vždy odohráva v slovách, a preto sa nimi ťažšie vyjadruje. So subjektívnou ťažkosťou vyjadrovania, a teda takouto subjektívnou stránkou jazyka, súvisia aj v práci naznačené neurovedné a psycholingvistické výskumy porúch vzťahu medzi, povedané spolu s Jánom Horeckým, človekom a jeho jazykom. Domnievam sa, že výskum v týchto oblastiach možno označiť za závažný, pretože jazyk predstavuje mocný nástroj pri utváraní a udržiavaní vzťahu so svetom a porucha tohto spojenia so svetom je závažná nemenej.

S tým súvisí i ďalší, inštrumentálny aspekt jazyka. Tento aspekt sa zrkadlí napríklad už v metaforickom spojení „ovládať jazyk“, ale vlastne je prítomný v priebehu celej histórie technológie (ako o tom svedčí súbežná existencia hovorenej reči, písaného slova, pomôcok a prostriedkov na tvorbu a spracovanie jazyka), a najmä v období rozvoja jazykových technológií a počítačového spracovania prirodzeného jazyka. A práve tu sa do obrazu dostáva korpusová lingvistika.

Okrem rôznorodého spôsobu, akým je človeku jazyk dostupný, som tiež v práci reflektoval aj rôznorodosť prístupov k jazyku: okrem toho biologického, psychologického či sociálneho treba spomenúť prístup logiky a analytickej filozofie, zameranej na otázky významu jazykových výrazov, a napokon značnú variabilitu prístupov v rámci samotnej lingvistiky. Túto variabilitu vnímam nie ako obrazné bojové pole, ale skôr ako paletu rozličných akcentov, ktoré má zmysel zohľadniť pri skúmaní takého vrstevnatého fenoménu ako jazyk. Jednotlivé akcenty, teda prístupy, majú podľa mňa schopnosť pôsobiť inšpiratívne na ostatné prístupy, a tak otvárajú priestor pre vzájomné synergické pôsobenia. V práci som sa preto zamerlal na fenomén dynamickeho rozvoja korpusovej lingvistiky jeho včlenením do takto naznačeného prieniku rôznych prístupov k jazyku, a to vedeckých i filozofických, najmä fenomenologických.

Príklon k fenomenológii bol motivovaný najmä reflektovanou komplexnosťou, ktorá sa s podobami a zložkami jazyka (ale nielen jeho) spája, zároveň si ale uvedomujem oprávnený apel analytickej filozofie, ktorá filozofické problémy navrhuje riešiť analýzou jazyka, ktorým o týchto problémoch hovoríme. V závere práce som preto úvahu obrátil aj samotnú fenomenologickú perspektívu. Nekritické používanie jazyka pri fenomenologickej reflexii môže, ako uvádza Michal Ajvaz s odkazom na Claudia Magrisa, vyústiť do paradoxných tvrdení (v príklade sa uvádzalo, že Dunaj vyviera z jedného vodovodného kohútika) a všeobecne do vágneho používania jazyka, pri ktorom je sťažené kantovské verejné používanie rozumu a zužuje sa priestor pre kritický racionalistický prístup k tvrdeniam o povahe sveta.

Zároveň si myslím, že filozofické uvažovanie, ktoré podľa Aristotela vyviera z prvotného údivu ľudí nad svetom, má svoje miesto aj vo vzťahu k rôznorodému fenoménu jazyka. Hodný zamyslenia je už fakt, že jazyk existuje v toľkých podobách, rovinách a je predmetom toľkých prístupov súčasne. Práve tu má azda opäť oprávnenie Nitscheho termín spleť: jazyk je zároveň predmetom vedeckého skúmania i subjektívneho prežívania, funkčný vzťah medzi jazykom a používateľom je zasa zdrojom aj subjektívneho zmyslu napríklad pri tvorbe jazykového diela alebo pri vedeckom skúmaní (nielen) jazyka. V poslednej kapitole som na tento subjektívny zmysel tvorby a skúmania odkázal Wilczekovou metaforou „skrytých rádov bytia“. Dúfam, že sa mi podarilo ukázať alebo

aspoň naznačiť, že súčasťou markošovského „evolučného tápania“ môže byť aj filozofické reflektovanie prostriedkov a cieľov jazykovedného skúmania v podmienkach jeho teoretického i praktického rozvoja.

Použitá literatúra

- Ajvaz, Michal 2012. *Cesta k pramenům smyslu. Genetická fenomenologie Edmunda Husserla*. Praha: Filosofía. 374 s.
- Ajvaz, Michal 2017. *Kosmos jako sebeutváření*. Praha: Pavel Mervart. 536 s.
- Ajvaz, Michal – Pauknerová, Karolína 2019. *Spor o proces a událost*. Praha: Pavel Mervart. 258 s.
- Aristotelés 2015. *Metafyzika A*. Preložil Filip Karfík. Praha: Oikoymenth. 72 s.
- Arppe, Antti et al. 2010. *Cognitive Corpus Linguistics: five points of debate on current theory and methodology*. Corpora, 5, 1, s. 1 – 27.
- Bateson, Gregory 2019. *Ekologie mysli*. Praha: Pavel Mervart. 512 s.
- Blumenthal-Dramé, Alice 2016. *What corpus-based Cognitive Linguistics can and cannot expect from neurolinguistics*. Cognitive Linguistics, 27, 4, s. 493 – 505.
- Boström, Nick 2017. *Superintelligence (Až budou stroje chytřejší než lidé)*. Praha: Prostor. 512 s.
- Camillus, John C. 2008. *Strategy as a Wicked Problem*. Dostupné na: <https://hbr.org/2008/05/strategy-as-a-wicked-problem>.
- Čermák, František 2017. *Korpus a korpusová lingvistika*. Praha: Karolinum. 270 s.
- Čermáková, Anna – Mahlberg, Michaela (eds.) 2018. *The Corpus Linguistics Discourse. In honour of Wolfgang Teubert*. Studies in Corpus Linguistics, 87. Amsterdam: John Benjamins Publishing Company. VI, 261 s.
- Černík, Václav – Viceník, Jozef (eds.) 2004. *Problém rekonštrukcie sociálnych a humanitných vied*. Bratislava: Iris. 232 s.
- Černík, Václav – Viceník, Jozef (eds.) 2005. *Zákon, explanácia a interpretácia v spoločenských vedách*. Bratislava: Iris. 288 s.
- Černík, Václav – Viceník, Jozef 2011. *Úvod do metodológie spoločenských vied*. Bratislava: Iris. 386 s.
- Černík, Václav – Viceník, Jozef 2012. *Empirické a teoretické postupy spoločenských vied*. Bratislava: Iris RR. 279 s.
- Černík, Václav 2014. *Prakticko-projektívne postupy spoločenských vied*. Bratislava: Iris RR. 344 s.
- Clark, Andy – Chalmers, David 1998. *The Extended Mind*. Analysis, s. 7 – 19.
- Coeckelbergh, Mark – Funk, Michael 2018. *Wittgenstein as a Philosopher of Technology: Tool Use, Forms of Life, Technique, and a Transcendental Argument*. Human Studies, 41, 2, s. 165 – 191.
- Coeckelbergh, Mark – Funk, Michael – Koller, Stefan, 2018. *Wittgenstein and Philosophy of Technology*. Techné: Research in Philosophy and Technology, 22, 3, s. 287 – 295.
- Coeckelbergh, Mark 2017. *Using Words and Things. Language and Philosophy of Technology*. New York/Oxon: Routledge. 296 s.
- Col, Gilles – De Angelis Rosanna – Poibeau, Thierry 2017. *Continuity in the Interactions Between Linguistic Units*. In: Compagno (ed.). Quantitative Semiotic Analysis. Springer International Publishing, s. 29 – 48.
- Compagno, Dario (ed.), 2018. *Quantitative Semiotic Analysis*. Springer International Publishing. IX, 189 s.
- Cvrček, Václav – Kovářiková, Dominika 2011. *Možnosti a meze korpusové lingvistiky*. Naše řeč, 94, 3, s. 113 – 133.
- Cvrček, Václav – Fidler, Masako (eds) 2018. *Taming the Corpus. From Inflection and Lexis to Interpretation*. Springer International Publishing. VI, 225 s.
- Deleuze, Gilles 2015. *Záhryb*. Praha: Hermann & synové. 240 s.

- Desagulier, Guillaume 2018. *Corpus Linguistics and Statistics with R. Introduction to Quantitative Methods in Linguistics*. Springer International Publishing. XIII, 353 s.
- Dhar, Vasant 2013. *Data science and prediction*. Communications of the ACM, 56, 12, December 2013, s. 64 – 73.
- Dolník, Juraj 2008. *Egolingvistika*. Slovenská reč, 73, 5, s. 257 – 270.
- Dolník, Juraj 2009. *Všeobecná jazykoveda. Opis a vysvetľovanie jazyka*. Bratislava: Veda. 2009, 376 s.
- Dolník, Juraj 2010a (ed.). *Morfologické aspekty súčasnej slovenčiny*. 1. vyd. Bratislava: Veda. 2010. 466 s.
- Dolník, Juraj 2010b. *Jazyk – človek – kultúra*. Bratislava: Kalligram. 224 s.
- Dolník, Juraj – Orgoňová, Oľga – Bohunická, Alena – Faragulová, Alena – Piatková, Kristína 2015. *Cudzosť, jazyk, spoločnosť*. Bratislava: Iris. 316 s.
- Dolník, Juraj 2018a. *Jazyk v pragmatike*. Bratislava: Veda. 248 s.
- Dolník, Juraj 2018b. *Lexikálny význam v recepcii*. Jazykovedný časopis, 69, 3, s. 302 – 315.
- Eichinger, 2007. *Linguisten brauchen Korpora und Korpora Linguisten: Wege zu wohl dokumentierten und verlässlichen Aussagen über Sprache*. In Werner Kallmeyer – Gisela Zifonun (eds). *Sprachkorpora: Datenmengen und Erkenntnisfortschritt*. Berlin – New York: Walter de Gruyter, s. 1 – 8.
- Ferenc, Jakub Albert 2019. *Věda o interface v teorii rozšířené kognice na pozadí McLuhanovy mediální teorie*. Univerzita Palackého v Olomouci, Filozofická fakulta, fakulta. Dostupné na: Dostupné na: <https://www.academia.edu/38517387/>.
- Filkorn, Vojtech 2009. *Obrazy vedy. Veda o vede*. Bratislava: Iris. 202 s.
- Findra, Ján 1996. *Jazyk, reč, človek*. Bratislava: Q111. 111 s.
- Flanderková, Eva 2019. *Čeština v afázii: teorie a empirie*. Praha: Karolinum. 158 s.
- Franssen, Maarten – Lokhorst, Gert-Jan – van de Poel, Ibo 2018. *Philosophy of Technology*. The Stanford Encyclopedia of Philosophy. Edward N. Zalta (ed.). Dostupné na: <https://plato.stanford.edu/archives/fall2018/entries/technology/>.
- Garabík, Radovan – Gianitsová, Lucia – Horák, Alexander – Šimková, Mária 2004. *Tokenizácia, lematizácia a morfológická anotácia Slovenského národného korpusu*. Bratislava: SNK JÚLŠ.
- Garabík, Radovan – Karčová, Agáta – Žáková, Adriána 2018. *Kvantitatívna analýza polyfunkčného kompotentu sa na báze Slovenského národného korpusu*. Slovenská reč, 83, 1, s. 23 – 34.
- Ghosh, Paramita 2019. *The Digital Mesh: How is it Changing Enterprises?* Dostupné na: <https://www.dataversity.net/the-digital-mesh-how-is-it-changing-enterprises/>.
- Gleick, James 2013. *Informace*. Praha: Argo – Dokořán. 400 s.
- Gries, Stefan Th. – Stefanowitsch, Anatol (eds.), 2006. *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax And Lexis*. De Gruyter Mouton. 352 s.
- Gries, Stefan Th. 2010. *Corpus linguistics and theoretical linguistics A love-hate relationship? Not necessarily*. International Journal of Corpus Linguistics 15, 3, s. 327 – 343.
- Gries, Stefan Th. 2012. *Corpus linguistics, theoretical linguistics, and cognitive/psycholinguistics: Towards more and more fruitful exchanges*. Language and Computers 75, XX, s. 41 – 63.
- Haig, Brian D. 2018. *The philosophy of quantitative methods*. London: Oxford University Press. 168 s.
- Harispe, Sébastien et al., 2015. *Semantic Similarity from Natural Language and Ontology analysis*. Morgan & Claypool (Synthesis Lectures on Human Language Technologies, Graeme Hirst (ed.), volume 27), XV, 238 s.
- Havel, Ivan M. 2018. *Zápisky introspektora*. Praha: Oikymenh. 114 s.
- Hayles, N. Katherine 2012. *How We Think. Digital Media and Contemporary Technogenesis*. Chicago – London: The University of Chicago Press. 296 s.

- Hayles, N. Katherine 2017. *Unthought. The Power of the Cognitive Nonconscious*. Chicago – London: The University of Chicago Press. 263 s.
- Heidegger, Martin 1993. *Básnický bydlí člověk*. Praha: Oikoymenh. 185 s.
- Heidegger, Martin 2013. *Věk obrazu světa*. Praha: Oikoymenh. 64 s.
- Horáček, Jiří – Kesner, Ladislav – Höschl, Cyril – Španiel, Filip et al. 2016. *Mozek a jeho člověk. Mysl a její nemoc*. Praha: Galén. 382 s.
- Horecký, Ján 1974. *Základné problémy terminológie*. In *Kultúra slova*, 1974, 8, 5, s. 129 – 134.
- Horecký, Ján 1979. *Spôsob jestvovania jazyka*. Jazykovedný časopis, 30, 2, s. 97 – 101.
- Horecký, Ján 2000. *Človek a jeho jazyk. Esej čiže pokus o charakteristiku vzťahu*. Jazykovedný časopis, 51, 1, s. 3 – 6.
- Husserl, Edmund 2004. *Ideje k čistej fenomenológii a fenomenologické filosofii I*. Praha: Oikoymenh. 340 s.
- Chlumská, Lucie 2017. *Překladová čeština a její charakteristiky. Studie z korpusové lingvistiky 26*. Praha: Nakladatelství Lidové noviny. 149 s.
- Ihde, Don 2003. "If Phenomenology is an Albatross, Is Post-phenomenology possible?" In Don Ihde and Evan Selinger (eds.), *Chasing Technoscience: Matrix for Materiality*, Indianapolis: Indiana University Press, s. 15 – 26.
- Introna, Lucas 2017. *Phenomenological Approaches to Ethics and Information Technology*. The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.). Dostupné na: <<https://plato.stanford.edu/archives/fall2017/entries/ethics-it-phenomenology/>>.
- Jackendoff, Ray 2003. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. London: Oxford University Press. 477 s.
- Jarošová, Alexandra 2018a. *Špirála poznania alebo funkcionalizmus a kognitivismus (v lexikológii) cez optiku dvoch kongresových jubileí*. Slovenská reč, 83, 2, s. 123 – 142.
- Jarošová, Alexandra 2018b. *Pohľad na pomenovanie cez prizmu teoretických rámcov a slovníkové hesla*. Jazykovedný časopis, 69, 3, s. 277 – 301.
- Jurafsky, Daniel – Martin, James H. 2018. *Speech and Language Processing*. Stanford University Press. 558 s. Dostupné na: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>.
- Juszczyk, Stanisław 2003. *Metodológia empirických výskumov v spoločenských vedách*. Bratislava: Iris. 137 s.
- Kaku, Michio 2013. *Fyzika budúcnosti*. Praha: Argo – Dokořán. 348 s.
- Konopka, Marek 2018. *Korpuslinguistik, Grammatiktheorie, Grammatikographie*. In Angelika Wöllstein, Peter Gallmann, Mechthild Habermann, Manfred Krifka (eds.). *Grammatiktheorie und Empirie in der germanistischen Linguistik*. Berlin – Boston: Walter de Gruyter. s. 151 – 184.
- Kořenský, Jan 2014. *Proměny myšlení o řeči na rozhraní tisíciletí*. Olomouc: Univerzita Palackého v Olomouci. 251 s.
- Koukolík, František 2008. *Sociální mozek*. Praha: Karolinum. 272 s.
- Kováříková, Dominika 2017. *Kvantitativní charakteristiky termínů*. Praha: Nakladatelství Lidové noviny. 136 s.
- Kristeva, Julia 2004. *Jazyk lásky*. Praha: One Woman Press. 250 s.
- Kroes, Peter 2012. *Technical Artefacts: Creations of Mind and Matter. A Philosophy of Engineering Design*. Springer Netherlands. X, 206 s.
- Kusá, Daniela (ed.) 2011. *Ekológia ľudských kognícií*. Bratislava: Veda. 163 s.
- La Mantia, Francesco – Licata, Ignazio – Perconti, Pietro (eds.) 2017. *Language in Complexity: The Emerging Meaning*. Springer International Publishing. XXVI, 199 s.
- La Mantia, Francesco 2017. *From Topology to Quasi-Topology. The Complexity of the Notional Domain*. In: La Mantia, Francesco – Licata, Ignazio – Perconti, Pietro (eds.). *Language in Complexity: The Emerging Meaning*. Springer International Publishing. s. 119 – 139.

- Louw, Bill 2011. *Philosophical and literary concerns in Corpus Linguistics. Interview with Bill Louw*. In: Viana, Vander – Zyngier, Sonia – Barnbrook, Geoff (eds.). *Perspectives on Corpus Linguistics*. Studies in Corpus Linguistics, 48. Amsterdam: John Benjamins Publishing Company, s. 171 – 196.
- Manning, Christopher 2015. *Computational Linguistics and Deep Learning*. Computational Linguistics, 41, 4, s. 699 – 705.
- Markoš, Anton 2015. *Život, veda a metafora živého. Rozhovor s Antonom Markošom*. Filozofia, 70, 1, s. 66 – 71.
- Markoš, Anton 2017. *Evoluční tápání*. Praha: Pavel Mervart. 368 s.
- Maxwell, Nicholas 2018. *Science and Enlightenment. Two Great Problems of Learning*. Springer International Publishing, XII, 108 s.
- McCarthy, Tom 2005. *Remainder*. London: Alma Books. 286 s.
- McEnery, Tony – Hardie, Andrew 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: University Press. 312 s.
- Meloun, Milan – Militký, Jiří – Hill, Martin, 2012. *Statistická analýza vícerozměrných dat v příkladech*. Praha: Academia. 756 s.
- Merleau-Ponty, Maurice 2017. *Proměna vnímání a zkušenost pravdy*. Praha: Oikoymenh. 72 s.
- Michálek, Jiří 2017. *Být místem Bytí. Tři příspěvky ke studiu Martina Heideggera*. Praha: Karolinum. 56 s.
- Miko, František 1970. *Text a štýl. K problematike literárnej komunikácie*. Bratislava: Smena. 167 s.
- Moisl, Hermann 2015. *Cluster Analysis for Corpus Linguistics. Quantitative Linguistics* 66. De Gruyter Mouton. XV, 381 s.
- Munn, Katherine – Smith, Barry 2008. *Applied Ontology. An Introduction*. Ontos Verlag. 342 s.
- Nelson, Barnaby – Nelson, Josef – Sass, Louis A., 2014. *Disturbance of Minimal Self (Ipseity) in Schizophrenia: Clarification and Current Status*. Schizophrenia Bulletin, 40, 3, s. 479 – 482.
- Niel, Luis 2017. *Husserl's Concept of Urstiftung: From Passivity to History*. In Roberto Walton, Shigeru Taguchi, Roberto Rubio (eds.): *Perception, Affectivity, and Volition in Husserl's Phenomenology*. Phaenomenologica 222. Springer International Publishing, s. 137 – 161.
- Nitsche, Martin 2016. *Metodická přednost spleti. Tranzitivně-topologický model fenomenologie*. Praha: Togga. 260 s.
- Ong, Walter J. 2006. *Technologizace slova*. Praha: Karolinum. 238 s.
- Paitlová, Jitka 2018. *Kritický racionalismus H. Alberta a K. Poppera*. Praha: Karolinum. 204 s.
- Paolucci, Claudio 2017. *System and Structure. Semiotics as Encyclopaedic Theory of Complexity*. In La Mantia, Francesco – Licata, Ignazio – Perconti, Pietro (eds.). *Language in Complexity: The Emerging Meaning*. Springer International Publishing, s. 83 – 94.
- Pearce, Jack 2017. *Fundamentals for the Anthropocene*. Walter de Gruyter. VI, 135 s.
- Pecina, Pavel 2009. *Lexical Association Measures: Collocation Extraction*. Praha: Ústav formální a aplikované lingvistiky UK. 143 s.
- Peregrin, Jaroslav – Vlasáková, Marta 2018. *Filosofie logiky*. Praha: Filosofia. 403 s.
- Petitot, Jean 2018. *The Formalization of Elementary Semiotic Structures*. In Compagno, Dario (ed.), 2018. *Quantitative Semiotic Analysis*. Springer International Publishing, s. 33 – 54.
- Petříček, Miroslav 2018. *Filosofie en noir*. Praha: Karolinum. 323 s.
- Puc, Jan 2017. *Jak vzniká něco nového? Založení a transcendence*. Filozofia, 72, 4, s. 259 – 271.
- Půlpán, Zdeněk 2012. *Odhad informace z dat vágní povahy*. Praha: Academia. 200 s.
- Ritchey, Tom 2011. *Wicked Problems – Social Messes. Decision Support Modelling with Morphological Analysis*. Berlin – Heidelberg: Springer-Verlag. VI, 106 s.
- Romano, Claude 2018. *Fenomenologie v dialogu s analytickou filosofií*. Reflexe 55, 1, s. 127 – 145.
- Rorty, Richard 2001. *Filozofia a zrkadlo prírody*. Bratislava: Kalligram. 339 s.

- Saks, Elyn R. 2008. *The Center Cannot Hold: My Journey Through Madness*. New York: Hachette Books. 368 s.
- Sarti, Alessandro – Bamberi, Davide 2018. *Neuromorphology of Meaning*. In Compagno (ed.), 2018. *Quantitative Semiotic Analysis*. Springer International Publishing, s. 55 – 74.
- Sass, Louis A. 2013. *Self-disturbance and schizophrenia: Structure, specificity, pathogenesis (Current Issues, New Directions)*. Schizophrenia Research, 152, 1, s. 5 – 11.
- Shulman, Helen 2012. *Living at the Edge of Chaos: Complex Systems in Culture and Psyche*. Dæmon Publications. 256 s.
- Šimková, Mária – Gajdošová, Katarína – Kmeťová, Beáta – Debnár, Marek 2017. *Slovenský národný korpus: texty, anotácie, vyhľadávania*. Praha: Mikula. 168 s.
- Stahl, Reinhold – Staab, Patricia 2018. *Measuring the Data Universe. Data Integration Using Statistical Data and Metadata Exchange*. Springer International Publishing. VII, 117 s.
- Stefanowitsch, Anatol 2018. *Corpus Linguistics: A guide to the methodology (draft)*. Language Science Press. Dostupné na: https://www.researchgate.net/publication/313328874_Corpus_Linguistics_A_Guide_to_the_Methodology.
- Stewart, Ian 2014. *Matematika života*. Praha: Academia. 392 s.
- Štícha, František 2005. *Možnosti a meze české gramatiky*. Praha: Academia. 304 s.
- Štícha, František – Fried, Mirjam (eds.) 2008. *Grammar & Corpora/Gramatika a korpus. Sborník příspěvků ze stejnomenné konference 25. – 27. 9. 2007, Liblice/Selected contributions from the conference Grammar and Corpora, Sept. 25 – 27, 2007, Liblice*. Praha: Academia. 444 s.
- Svítek, Miroslav 2013. *Víc než součet částí: Systémový pohled na proces lidského poznání*. Praha: Academia. 228 s.
- Teubert, Wolfgang 2005. *My version of corpus linguistics*. International Journal of Corpus Linguistics, 10, 1, s. 1 – 13.
- Tognini-Bonelli, Elena 2001. *Corpus Linguistics at Work*. Studies in Corpus Linguistics, 6. John Benjamins Publishing Company. xii, 224 s.
- Tukey, John 1980. *We need both exploratory and confirmatory*. American Statistician, 34, s. 23 – 25.
- Turner, Raymond – Angius, Nicola 2019. *The Philosophy of Computer Science*. The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.). Dostupné na: <https://plato.stanford.edu/archives/spr2019/entries/computer-science/>.
- Tvrď, Filip 2014. *Turingův test. Filozofické aspekty umělé inteligence*. Praha: Togga. 213 s.
- Viana, Vander – Zyngier, Sonia – Barnbrook, Geoff (eds.) 2011. *Perspectives on Corpus Linguistics*. Studies in Corpus Linguistics, 48. Amsterdam: John Benjamins Publishing Company. xvi, 256 s.
- Višňovský, Emil 2004. *Filozofia a sociálne myslenie*. In Černík, Václav – Viceník, Jozef (eds.). *Problém rekonštrukcie sociálnych a humanitných vied*. Bratislava: Iris, s. 91 – 139.
- Wallis, Sean 2007. *Annotation, retrieval and experimentation, Or: you only get out what you put in*. In Anneli Meurman-Solin – Arja Nurmi (eds): *Studies in Variation, Contacts and Change in English. Volume 1. Annotating Variation and Change*. Dostupné na: <http://www.helsinki.fi/varieng/series/volumes/01/wallis/>.
- Wilczek, Frank 2011. *Lehkost bytí aneb Bytí jako světlo*. Praha: Dokořán. 304 s.
- Wladawsky-Berger, Irving 2014. *Why Do We Need Data Science when We've Had Statistics for Centuries?* Dostupné na: <https://blog.irvingwb.com/blog/2014/04/why-do-we-need-data-science-when-weve-had-statistics-for-centuries.html>.
- Yoos, George E. 2015. *Simplifying Complexity. Rhetoric and the Social Politics of Dealing with Ignorance*. Walter de Gruyter. X, 178 s.
- Zouhar, Marián 2011. *Význam v kontexte*. Aleph. 354 s.
- Zouhar, Marián 2016. *Konceptuálna analýza v analytickej filozofii*. Filozofia, 71, 5, s. 410 – 424.