

# Korpusová kritická analýza diskurzu: povaha, možnosti a limity (na příkladu analýzy jazykových ideologií v českém parlamentním diskurzu)<sup>1</sup>

Ondřej Dufek

## ABSTRACT:

**Corpus-assisted Critical Discourse Analysis: nature, possibilities and limits (Exemplified on the analysis of language ideologies in the Czech parliamentary discourse).** The paper provides a thorough review of the corpus-linguistic approach to critical discourse analysis. It briefly presents the core of critical discourse analysis (CDA) and examines the possibilities of applying corpus tools to it. In the next step, critical commentaries on CDA are summarized and at the same time, possible corpus-linguistic solutions are offered. The final part offers an illustrative application of corpus-assisted CDA focusing on language ideologies in the Czech parliamentary discourse.

## KLÍČOVÁ SLOVA / KEYWORDS

český parlamentní diskurz, jazykové ideologie, jazykový zákon, korpusová kritická analýza diskurzu, kritická analýza diskurzu, korpusová lingvistika, kritika kritické analýzy diskurzu corpus linguistics, corpus-assisted critical discourse analysis, critical discourse analysis, criticism of critical discourse analysis, Czech parliamentary discourse, language bill, language ideologies

## 1. ÚVOD

Studium užívání jazyka (a dalších sémiotických kódů) v sociální realitě v konfrontaci s ideologiemi a mocenskými vztahy má ve světové vědě tradici dlouhou již několik desetiletí — přinejmenším pod označením kritická analýza diskurzu (critical discourse analysis, CDA). Tento výzkumný proud obdržel za dobu své existence řadu kritických komentářů, které zčásti posílily jeho rozvoj. Ten je do značné míry (ovšem rozhodně ne výlučně) spjat s využíváním korpusové lingvistiky (corpus linguistics, CL), která umožňuje solidnější datové založení a využití dříve nedostupných způsobů poznávání textů.

Tento článek chce představit korpusovou kritickou analýzu diskurzu, ukázat její výhody a relevanci, diskutovat její limity a ilustrovat její využití na českém materiálu. Domnívám se, že takový přehled může být užitečný, neboť v českém prostředí zatím korpusová CDA byla zohledněna jen sporadicky. O zapojení korpusových nástrojů do analýzy diskurzu se nedávno zmínily Jindřiška Svobodová, Šárka Hamrusová a Barbora Hinková (2018), ovšem na ploše pouhé jedné a čtvrt stránky a v podobě, která je do značné míry redukuje na pouhé dosažení větší spolehlivosti a zpracování většího

---

1 Článek vychází z mé disertační práce (Dufek, 2018) — je přepracovanou verzí části její čtvrté kapitoly rozšířenou o vybrané pasáže z kapitoly šesté. Děkuji oběma recenzentům tohoto článku za kritické čtení vedoucí ke skutečně podnětným připomínkám.



množství dat. To je — jak ukážu v tomto článku — jen jedna z výhod korpusové CDA. Solidnější, byť též stručné seznámení nabízí Eva Lehečková (2016b). Jednotlivé případové studie podnikli Tess Slavíčková (2015), Irene Elmerot (2017) a Michal Hořejší (2019). Tím se, pokud vím, literatura vztažená k českému prostředí v podstatě vyčerpává.

Sluší se nicméně zmínit, že v českém prostředí / na českém materiálu vzniklo už několik studií, které se sice nehlásí ke kritickým východiskům, ale svou korpusově-diskurzivní povahou se dostávají do nemalé blízkosti toho, co chce představit tento článek. Jsou to např. Zasina (2018) analyzující kolokačně českou mediální reprezentaci političek nebo Fidlerová a Cvrček (2018) rozebírající, v čem spočívá alternativnost jazyka zpravodajského webu Sputnik Česká republika. Tyto studie s CACDA sdílejí část nástrojů i teoretických konceptů, opírají se zčásti i o kritickoanalytickou literaturu, nedeklarují však svou ideologickou/politickou pozici ve vztahu k rozebíraným jevům a neformulují takto pozičně ukotvené a hodnotící interpretace.

V tomto textu nejprve velmi stručně představím podstatu CDA a pak rozeberu možnosti jejího obohacení o korpusové prvky. Následně shrnu nejdůležitější kritické výtky vůči CDA a zhodnotím, nakolik je může pomoci vyřešit právě korpusový přístup. Poslední částí textu je ilustrativní analýza části parlamentního diskurzu o návrzích jazykových zákonů, která předestře praktické možnosti využití korpusových nástrojů v kritické analýze jazykových ideologií.

## 2. ZÁKLADY KRITICKÉ ANALÝZY DISKURZU

Definovat CDA není snadný úkol. Pro účely tohoto článku se omezím na několik stručných poznámek věnovaných samé podstatě CDA, jejím cílům, vnitřnímu členění na samostatné směry a povaze její kritičnosti.

### 2.1 CO JE CDA

CDA<sup>2</sup> je obecně vzato přístup ke studiu jazyka a společnosti, sociopoliticky angažovaný vědecký směr; není to však teorie ani metoda (ke zdrojům, z nichž CDA čerpá, podrobněji v následujícím oddílu). Označuje také síť lidí s podobnými hodnotovými východisky, cíli a způsoby jejich dosahování, ale není to jednotná škola; termín CDA se vžil jako zastřešující označení do značné míry samostatných přístupů. Za dobu své existence (resp. existence onoho označení, lze říct i značky) už CDA získala akademickou relevanci, má vlastní časopisy,<sup>3</sup> konference.<sup>4</sup>

2 Nověji se začíná prosazovat i označení *critical discourse studies* (CDS), které bylo určitou dobu synonymní s CDA, ale nyní podle Johna Flowerdewa a Johna E. Richardsona (2018) odráží stále výraznější posun od aplikovaných analýz k filozofickému, teoretickému, metodologickému i praktickému obohacování tohoto směru (ostatně si všímají změny názvu 3. vydání knihy Ruth Wodakové a Michaela Meyera z *Methods of Critical Discourse Analysis* na *Methods of Critical Discourse Studies* (2015)).

3 *Discourse & Society*, *Journal of Language and Politics*, *Critical Discourse Studies* aj.

4 Např. *Critical Approaches to Discourse Analysis Across Disciplines*.

Stručně řečeno, CDA je analýza diskurzu využívající vědecké metody, ovšem nikoli se snahou o nedosažitelnou objektivnost, nýbrž s přiznanými politickými cíli upozornit na nerovnost a přispět k jejímu odstranění.



## 2.2 CÍLE A TÉMATA

CDA se snaží rozkrývat a zviditelňovat mocenské vztahy ve společnosti a jejich jazykovou/diskurzní povahu; soustředí se na moc v diskurzu a moc nad diskurzem. Z toho důvodu coby materiál analýzy používá autentické promluvy, reálné texty. V tomto smyslu se řadí do souboru směrů reagujících na redukcionistický Chomského generativismus.

Vznikla primárně na lingvistickém základě, snaží se však zohledňovat i jiné sémiotické kódy, především však není popisem jazyka, nýbrž popisem jeho užívání. Využívá lingvistický deskriptivní aparát, zajímá ji především nadvětná, textová a diskurzní rovina; lexikon či gramatiku vždy vykládá ve vztahu k rovinám vyšším a usouvztažňuje s relevantním kontextem (významy kontextu viz níže).

Lingvistické analýzy zpravidla poukazují na ideologie a na nich postavenou nerovnost, nespravedlnost, uplatňování a/nebo zneužívání mocenského postavení ve společnosti pomocí rozkrývání jejich projevů v diskurzu.<sup>5</sup> Widdowson (1998, s. 136) ji označuje za *uvědomělou lingvistiku* („linguistics with a conscience“).

CDA není studium humanitní, nezkoumá člověka a jeho nástroje dorozumívání, nezjišťuje primárně, jak jazyk funguje v mozku či myslí<sup>6</sup> — je studiem sociálním, protože ji zajímá, jak jazyk funguje coby prostředek určitých sociálních aktů, jak jazykové prostředky manifestují interakci v rámci sociálních vztahů. Jinak řečeno, propojuje „lingvistický výzkum jazykových struktur a společenskou roli jazykové komunikace jako prostředku pro utváření, předávání a uchovávání různých sociálních praktik (*social practices*)“ (Lehečková, 2016a, s. 79).

Rozpětí vztahu CDA k jazyku lze postihnout mj. dvojicí výrazů *language in use* a *language in action* — jde o to, co kdo (s) jazykem činí.

## 2.3 SMĚRY

Když uvádím, že CDA není koherentní školou, patří se vysvětlit, z čeho různého se tedy skládá. Označení *critical discourse analysis* se (podle Breezeové (2011, s. 495)) obje-

5 Někteří badatelé jsou spatřují stopy mocenských vztahů i přímo v systému jazyka, viz např. Caldas-Coulthardová a Coulthard (1996, s. xi), podle nichž CDA zjišťuje, „how power and discriminatory value are inscribed in and mediated through the linguistic system“. Většinou však dnes panuje konsensus na tom, že jádrem zájmu CDA je *užívání jazyka*. Mocenské potenciály přímo v jazyce bychom našli např. v starší kritické lingvistice (např. Fowler et al., 1979; Fowler, 1996; či Kress & Hodge, 1979), ale také např. v lingvistice genderové (v českém prostředí viz např. Valdřová, 1997; Čmejrková, 1995; 2002).

6 Srov. však zaměření na kognitivní rozměr reprezentace diskurzu např. u Teuna van Dijka (např. 2000, s. 11n.), Paula Chiltona (2011, 2014) nebo Chrise Harta (2015, 2019) aj.; viz též odd. 2.3.



vilu poprvé v článku Normana Fairclougha (1985, s. 739), ve známost vešlo hlavně s vydáním jeho knihy *Critical discourse analysis* (1995). Zčásti fundující funkci mělo (jak uvádí Wodaková a Meyer (2009, s. 3)) sympozium v Amsterdamu v r. 1990; ve stejném roce vznikl i časopis *Discourse & Society*. Hlavní směry jsou dialekticko-relační přístup Normana Fairclougha, socio-kognitivní přístup Teuna A. van Dijka a diskurzně-historický přístup Ruth Wodakové, multimodální studia v rámci CDA rozvíjí Theo van Leeuwen. Někdy se uvádějí ještě francouzská škola (Michel Pêcheux), duisburská škola (Siegfried Jäger) aj. (viz Wodak, 2011).

V poslední době se rozvíjejí novější zajímavé varianty: kognitivní CDA (viz např. Hart, 2015, 2019; Cap, 2018), kombinace CDA a teorie metafor (Musolff, 2016; Li, 2016; Charteris-Black, 2004, 2014) aj.

## 2.4 V ČEM SPOČÍVÁ KRITICHNOST?

Přestože je kritičnost obsažena v samém názvu CDA, její podstata není nikterak samozřejmá ani nesporná. Ruth Breezeová (2011, s. 499–500) poukazuje na to, že může mít přinejmenším dva významy: Může jít o odkaz ke kritické teorii frankfurtské školy (ovšem v rámci toho ještě není zřejmé, zda se akcentuje kritický postoj ke statu quo, anebo k „authoritarian positivism of orthodox Marxism“ (*ibid.*)), stejně tak ale dává smysl *kritický* ve významu nepřijímající vše bez okolků, ověřující a promyšlejší.

Norman Fairclough (1992, s. 9) při charakteristice CDA napsal, že „Critical’ implies showing connections and causes which are hidden“. Příbuzná, byť ne tak vyhraněná, je pozice Ruth Wodakové a Michaela Meyera (2009, s. 2), kteří vysvětlují, že „[a]ny social phenomenon lends itself to critical investigation, to be challenged and not taken for granted“. Podle jejich slov je frankfurtská kritická teorie důležitý inspirační zdroj. Kromě toho (*ibid.*, s. 7) s odkazem na Thea van Leeuwen (2005, s. 293) vysvětlují, že „[n]aming oneself ‚critical’ only implies superior ethical standards: an intention to make their position, research interests and values explicit and their criteria as transparent as possible, without feeling the need to apologize for the critical stance of their work“.

Ostatně soudím, že tyto dvě strany kritického postoje nejsou v rozporu.

## 3. KORPUSOVÁ KRITICKÁ ANALÝZA DISKURZU

Na základě kritiky CDA se množství badatelů začalo ohlížet po korpusové lingvistice (*corpus linguistics*; CL), jelikož se zdá(lo), že některé výtky může alespoň částečně řešit. Zodpovídání otázky, proč takové zdání budí (a nakolik to je/není zdání iluzorní), bude předmětem následující kapitoly. Nejprve však učiním několik poznámek k variaci označování korpusové CDA (oddíl 3.1) a stručně vysvětlím podstatu kombinace CL a CDA (oddíl 3.2).

### 3.1 CORPUS-ASSISTED CRITICAL DISCOURSE ANALYSIS

Stejně jako CDA ani *corpus-assisted critical discourse analysis* (CACDA) není homogenní skupina či škola. Ostatně na rozdíl od CDA nemá ani označení, na němž by se

všichni shodli. V literatuře se, často bez jakýchkoli výsledovatelných pravidelností, objevují kombinace naznačené ve schématu 1 — a ani to nevyčerpává všechny možnosti, např. *corpus approaches to discourse studies* nebo *discourse-oriented corpus studies*.<sup>7</sup> Ostatně název tohoto oddílu je prostě jen dosud neobsazená kombinace, která se zdá být dostatečně výstižná. Výraz *assisted* je navíc výhodný tím, že nevstupuje do dichotomie *corpus-based vs. corpus-driven* (viz níže).

<i>corpus</i>	<i>assisted</i>	( <i>critical</i> )	<i>discourse</i>	<i>analysis</i>
<i>computer</i>	<i>based</i>			<i>studies</i>
	<i>driven</i>			<i>approach</i>
	<i>informed</i>			
	<i>aided</i>			

**SCHÉMA 1:** Kombinatorika složek označení přístupů kombinujících CL a CDA

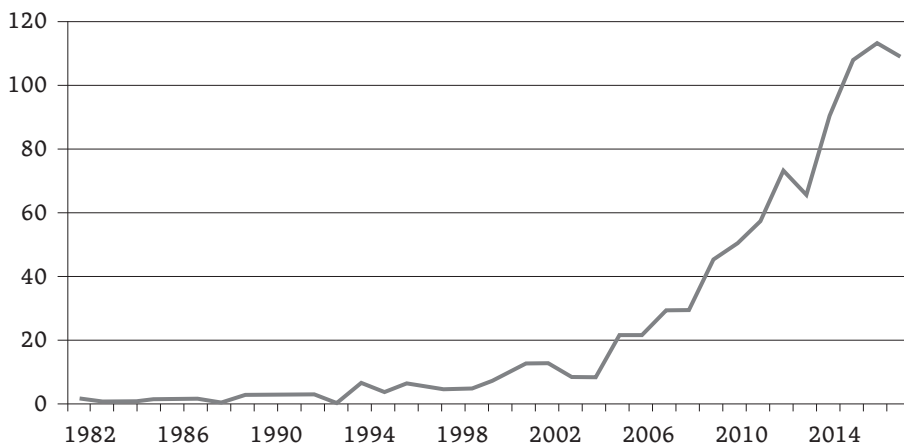
Jediná škola, která sama sebe označuje ustáleným názvem, je *corpus-assisted discourse studies* (CADS), spojovaná především s pracemi Alana Partingtona a jeho spolupracovníků (např. Partington, Duguid & Taylor, 2013). Proti tomu lze postavit podobný, leč v některých základních aspektech (přinejmenším deklarativně) odlišný přístup, který můžeme označit jako *corpus-based critical discourse analysis* (CBCDA) reprezentovaný lancasterskou školou, zejména Paulem Bakerem, Costasem Gabrielatosem aj. (např. Baker, Gabrielatos & McEnery, 2013). Alternaci *studies/analysis* zde můžeme považovat za málo významotvornou (srov. však poznámku 2); v druhém přístupu je klíčový výraz *critical*; nejproblematictější je pak rozlišení *corpus-assisted vs. -based*. První podobu si pro sebe rezervovala CADS, naopak lancasterská škola se proti pouhé asistenci korpusové lingvistiky vymezuje a prosazuje rovnoprávnost CL a CDA (Baker et al., 2008, s. 274).<sup>8</sup>

Kromě výše nastíněných pojmenovacích odlišností zmíním ještě variantu *corpus-informed*, kterou nalezneme u Kierana O'Hallorana (2007) nebo *computer-assisted* vyskytující se zejména ve starších pracích Michaela Stubbs (1994, 1996).

Už v devadesátých letech vznikla řada průkopnických studií, viz např. Louw (1993), Caldas-Coulthard (1993; 1995), Hardt-Mautner (1995), Krishnamurthy (1996), Stubbs (1996), Flowerdew (1997). Před necelými patnácti lety se počet článků kombinujících CL a analýzu diskurzu začal významně zvětšovat a od té doby téměř neustále roste — viz graf 1.

7 Dvě varianty názvu oborové bibliografie sestavované Costasem Gabrielatosem, donedávna dostupné na <<http://www.gabrielatos.com/CLDA-Biblio.htm>>.

8 Distinkce *assisted vs. based* ovšem není tak jednoznačná, jak bývá deklarováno. Paul Baker má v samotném názvu některých svých studií (viz např. Baker, 2014, 2015) označení *corpus-assisted*, naproti tomu Alan Partington (2006) hned v abstraktu svého článku říká, že rozpracoval „*corpus-based methodology*“.



**GRAF 1:** Vývoj počtu studií kombinujících korpusovou lingvistiku a (kritickou) analýzu diskurzu v databázi Scopus; dotaz: TITLE-ABS-KEY(corpus OR corpora OR „corpus linguistics“) AND TITLE-ABS-KEY(CDA OR „critical discourse analysis“ OR „discourse analysis“ OR „discourse studies“); výběr omezen na články, konferenční příspěvky a kapitoly v monografiích; manuálně nečištěno.

### 3.2 PODSTATA KORPUSOVÉ KRITICKÉ ANALÝZY DISKURZU

V tomto oddíle představím samotnou povahu korpusové kritické analýzy diskurzu, především to, jaké korpusové nástroje a jakými způsoby pro kritické rozbory využívá. Jak trefně poznamenává Debbie Orpinová (2005, s. 38), „[t]he major problem in combining a CDA approach with corpus methodology is deciding where to start“. Jedním z významných smyslů využívání CL je možnost zvládnout analyticky zpracovat větší množství dat, tj. takový jejich objem, který by člověk jednoduše nedokázal v rozumném čase / při rozumné námaze přečíst, natož pak intelektuálně uchopit a hledat výrazné prvky, pravidelnosti apod. To však rozhodně není jediná motivace — korpusový pohled na diskurz nabízí zároveň příležitost dostat se k běžným čtením nepostihnutelným pravidelnostem napříč texty.

Korpusový přístup (nejen v CDA), ostatně jako empirický výzkum jazyka obecně, nutně staví na práci s jednotkami vedoucí k jistému abstrahování od jejich zasazení v jazykovém, situačním a sociálním kontextu. Tato ztráta informací je však vyvážena ziskem na jiné straně. Michal Hořejší (2016, s. 5) k tomu poznamenává, že „[k]aždá korpusová analýza je z principu postavená na mechanismu nabývání informací z vertikály protínající velký počet jednotlivých textů (či korpusů) za cenu ztráty přirozené horizontály realizované linearitou textu a dialogickým rámcem diskurzu (intertextualitou)“. Corpus-assisted critical discourse analysis (CACDA) minimalizuje míru ztráty horizontály obvykle několika způsoby: opakovaným pohybem od počítačově generovaných průřezů korpusem k jednotlivým konkordancím až k širšímu kotextu originálních textů (a zpět) a zapojováním znalostí o situačním a sociálním kontextu na různých stupních analýzy. Přidaná hodnota korpusového přístupu je přitom nemalá. Jejím smyslem je „snaha se od této reality [jazykové situace; OD] odpoutat a do-

stat se za jazykovou praxi její fragmentací a mechanizací (sr. Stubbs 1996, Partington 2008). CL [korpusová lingvistika; OD] v CDA jinými slovy slouží k automatickému odhalování pravidelných (lexikálně-gramatických) schémat v rámci daného diskurzu za současného rozbití přirozené textové linearit (kotextu)“ (Hořejší, 2016, s. 3).

V takto nastaveném prostředí, kde máme velké množství textů zpracovaných do korpusově prohledávatelné podoby, se nabízejí dvě cesty.

První předpokládá, že analytik již alespoň do určité míry zná diskurz a/nebo že má již určité hypotézy, předpoklady, jež chce otestovat. V takovém případě může zadat dotaz na určité slovo či obecněji jazykový prostředek, který ho zajímá, a např. zkoumat, jak se chová v konkordancích. Jsou-li k dispozici taková data a je-li korpus vhodně strukturován, lze věnovat pozornost diachronii a sledovat vývoj funkce či pozice onoho jazykového prostředku v diskurzu. Na takovýto přístup se dobře hodí označení *corpus-based* — korpus a s ním spojené standardní nástroje pomáhají hledat a nacházet to, co analytik hledá.

Druhá možnost se hodí pro situace, kdy analytik buď nemá dobrou představu o povaze diskurzu, anebo se chce záměrně co nejvíce oprostít od svých intuicí; to bychom mohli označit jako *corpus-driven* přístup. Pak může být dobré začít něčím takovým, jako je frekvenčně seřazený seznam slov (který concordancery běžně umějí generovat). Aníž bych zabíhal do problematických detailů, můžu si snad dovolit tvrdit, že frekvence nějak souvisí s typičností, významem — míněno významností — v rámci diskurzu. Přirozeně takový frekvenční seznam bude na horních příčkách obsahovat jen těžko překvapivá, „zajímavá“ slova — s největší pravděpodobností tam nalezneme synsémantika, pak nejspíš pomocná slovesa, osobní zájmena apod., tedy slova, která se velmi často vyskytují přirozeně v každém textu bez ohledu na příslušnost k určitému diskurzu.

Pro ilustraci uvedme dvacet nejfrekventovanějších lemmat z první kapitoly knihy Paula Bakera (2006, s. 1–24) *Using corpora in discourse analysis* (viz tabulka 1).<sup>9</sup> V tabulce vidíme, že prvních osm pozic zabírají synsémantika<sup>10</sup> a sloveso *to be*, užitě pravděpodobně z velké části jako pomocné. Poté však narazíme na lemmata, která už mohou o zkoumaném diskurzu (resp. v našem případě malé části textu) něco vypovídat. Je patrně v pořádku, že mezi nejčastějšími autosémantikami najdeme lemmata *discourse*, *corpus* a *language*.

Čtenář by mohl zapochybovat nad užitečností takového seznamu, mohl by ho považovat za nepřekvapivý a nezajímavý. Avšak i tato ukázková tabulka nabízí podnětů mnohem víc, než se na první pohled zdá: Vztahuje se sloveso/substantivum *use* spíše k uplatňování nějakých postupů/nástrojů, anebo k užívání jazyka (*language use*)? Je tedy kapitola zaměřena spíše na techničtější, instrumentálnější výklad, anebo na konceptuální podloží analýzy diskurzu? Dále, proč je z osobních zájmen nejčastější *we* a nikoli *I* nebo *you*? Jaký plurál Baker používá — začleňuje do něj čtenáře, nebo

9 Seznam byl vytvořen v softwaru AntConc s využitím anglické lemmatizace nabízené na webu softwaru.

10 Vědomě nechávám stranou otázku, jak přesně je ošetřena lemmatizace výrazu *that*, tedy jaká část dokladů je ve funkci spojky a jaká ve funkci zájmena — pro můj výklad to není relevantní.



pořadí	frekvence	lemma	pořadí	frekvence	lemma
1	399	of	11	119	or
2	395	be	12	110	as
3	346	the	13	90	have
4	303	a	14	86	we
5	297	to	15	82	for
6	237	and	16	76	language
7	232	in	17	73	which
8	166	that	18	72	it
9	132	discourse	19	71	use
10	130	corpus	20	67	this

**TABULKA 1:** Začátek word listu první kapitoly knihy Baker (2006).

ne? A pokud ano, jde o kontaktovou strategii, nebo má jinou funkci? A do třetice, proč mezi frekvenční špičkou chybí poslední slovo z titulu celé knihy, u něž by bylo možno předpokládat taktéž jistou prominenci, tedy *analysis*? Toto lemma je ve skutečnosti 25. a je pouze důsledkem formálně založeného výběru, že se nedostalo před oči čtenáře (stejně tak dobře mohla mít tabulka 15 nebo třeba 30 řádků a důvod/vysvětlení by těžko mohl/o být lepší či horší). Poslední otázka však poukazuje na klíčový problém využívání CL v CDA: Co v korpusu není, to v korpusu nenajdeme.

Moderní korpusové softwary jako #Lancsbox (Březina, Timperley & McEnery, 2018) však při použití některých metrik velikosti účinku nabízejí kromě seznamů (pozitivních) klíčových slov také negativní klíčová slova, tj. výrazy, které se výrazně málo vyskytují ve zkoumaném korpusu v porovnání s korpusem referenčním. Taková negativní klíčová slova mohou analytika upozornit na jevy, jejichž přítomnost by očekával, které však v korpusu chybějí.

Nepřítomnost (nebo méně kategoriicky vzato nízká frekvence) může být srovnatelně důležitá a významná jako vysoká četnost. Zde je vhodné upozornit na neustálou nutnost kontroly a zvažování kroků, které analytik nechá na počítači a které by měl podniknout sám dále.

Takový frekvenční seznam, jaký představuje tabulka 1, je jistě podnětný, jenže — jak jsme viděli — i tak plný balastu pravděpodobně bez větší relevance pro rozkrývání na první pohled neviditelné povahy diskurzu. V tomto směru může lépe sloužit navazující seznam klíčových slov. Klíčová slova jsou totiž postavena na optice frekvenční významnosti určitých slov pro zkoumaný korpus v porovnání s korpusem referenčním, a tudíž opomíjejí to, co je oběma srovnávaným souborům společné.

Zde takový seznam nenabídnu, neboť není zcela triviální říct, s čím by se měla četnost porovnávat, oproti čemu by mělo být klíčové slovo četnější. Jedna možnost by samozřejmě byla porovnat předmět našeho zájmu s obecným korpusem, dejme tomu *Bank of English* nebo *British National Corpus*. Možná by však bylo zajímavější zkusit porovnat knihu Paula Bakera s něčím srovnatelnějším, např. s publikací *Patterns and Meanings in Discourse* Alana Partingtona et al. (2013). Klíčová slova budou představena a využita v následující kapitole.



Dalším krokem může být kolokační analýza. Může mít například smysl ptát se, s jakými adjektivy-přívlastky se určité klíčové slovo pojí, jak se (např. na škále pozitivně-negativně) konstruuje v diskurzu. U slov reprezentujících aktéry lze sledovat statisticky podložené pravidelnosti v uvádění charakteristik, činností apod. Evaluativní potenciál kolokací pokrývá koncept sémantické/diskurzivní prozodie — viz Monika Bednareková (2008) a Paul Baker, Andrew Hardie a Tony McEnery (2006, s. 58).

Obvyklou součástí CACDA bývá analýza konkordancí. Ta umožňuje sledovat pravidelnosti bezprostředního jazykového kontextu vyhledaného výrazu a zvýrazňuje tak vlastnosti korpusu/diskurzu v již vyváženějším poměru horizontálního a vertikálního pohledu.

Většina dosavadního výkladu se orientovala především na lexikální stránku diskurzu. Užitečnou součástí korpusových analýz diskurzu však bývají i rozbory gramatických pravidelností (viz např. Partington, 2002, s. 7–9; Baker, 2006, s. 151–174). Lze si představit např. studii zahrnování či vylučování určitých aktérů pomocí sledování zájmen a/nebo užití kategorie čísla. V úvahu přichází výzkum výskytu konkrétního výrazu v různých syntaktických funkcích (např. podmět/agens vs. předmět/patiens), Podmínkou uskutečnitelnosti takových sond je však do značné míry gramaticky (typicky morfologicky, případně i syntakticky) anotovaný korpus.

Zásadním principem CACDA, má-li si zachovat ohled na kontext, je (podle potřeby opakovaně aplikovaná) kombinace kvantitativního aspektu s kvalitativním. Konkrétní sled kroků zapojujících frekvenčně založené korpusové nástroje a detailní čtení širších jazykových kontextů spolu se zohledněním kontextu situačního, sociálního, kulturního apod. je třeba samozřejmě ušít na míru příslušnému předmětu zkoumání a analyzovanému materiálu. Bez kvantitativní složky bychom však studii v rámci CDA těžko mohli označit za korpusovou, bez kritického kvalitativního zaměření na kontextualizovaný detail bychom podnikali korpusovou studii bez kritické analýzy diskurzu.

#### 4. KRITIKA CDA A JEJÍ KORPUSOVÁ ŘEŠENÍ

Poté, co se kritická analýza diskurzu ustálila a začala být vnímána jako (alespoň do určité míry) smysluplný a reálně praktikovaný způsob vědeckého bádání jak samotnými jejími představiteli (tzn. že se sami identifikovali jako příslušníci směru, jež lze označit zkratkou CDA), tak akademiky stojícími mimo tento proud, začaly se pochopitelně objevovat na její adresu kritické poznámky. Některé byly vzneseny, protože se původní jednotlivosti v rostoucím objemu vznikající literatury ukázaly být pravidelnostmi, jiné zase proto, že se představitelé kritizovaných praktik postupně stali vůdčími osobnostmi a příslušníky kánonu, a jejich práce tedy částečně reprezentovala celou CDA.

V následujících odstavcích shrnu hlavní výtky — kritériem jejich výběru je kombinace frekvence a závažnosti. Děním je na dvě základní skupiny: (1) na ty, které se týkají samotné podstaty CDA, jejího založení, tedy výtky jdoucí k jádru přístupu jako takového; (2) na ty, které upozorňují na neblahou praxi, nedostatečné dodržování postupů považovaných za nezbytné pro vědecké počínání, jež je však spatřováno nikoliv



v několika jednotlivých studiích, nýbrž ve většině prací jako norma. Zároveň zvažím, zda korpusová lingvistika nabízí jejich (alespoň částečné) řešení. První skupinou se zabývají oddíly 4.1–4.2, druhou skupinou oddíly 4.3–4.7.<sup>11</sup>

#### 4.1 CDA NENÍ ANALÝZA, NÝBRŽ INTERPRETACE

Jedním z prvních výrazných kritiků CDA byl Henry Widdowson (1995, s. 159), který zpochybňuje oprávněnost samotného označení *analýza*. Tvrdí, že smysl analýzy spočívá v tom, že „it seeks to reveal those factors which lead to a divergence of possible meanings, each conditionally valid“, a že teprve po takovém rozboru může analytik svá zjištění interpretovat, argumentovat ve prospěch jednoho či druhého, zvažovat jejich relevanci atp. Upřednostnění určitého významu na základě subjektivní preference, bez přiznání (či uvědomění si) vlastní jednostrannosti je podle něj pouhá interpretace. CDA se většinou netají svou stranickostí, míněno nikoliv příslušností ke konkrétní politické straně, nýbrž obecněji politickou pozicí a též určitým politickým cílem, ovšem neblahým důsledkem je podle Widdowsona to, že analytik (resp. interpret) straní těm interpretacím, které vyhovují jeho ideovému nastavení.

Tato kritika se dotýká obecnějšího problému, který však nebývá zmiňován příliš často, totiž přípustnosti přiznaně neobjektivního přístupu. Na CDA lze nahlížet nejen jako na vědecký směr v pravém slova smyslu, nýbrž jako na sociálně angažovanou činnost využívající vědecké prostředky — pojmy, termíny, metody apod. Aby CDA byla ve Widdowsonových očích analýzou, měla by patrně zvažovat možné významy jazykových prostředků a jejich užití, jejich relevanci z hlediska sociální praxe atd. pokud možno neutrálně, a teprve poté se přiklonit k jejich výkladu. Stručně řečeno, otázka zní, zda CDA nachází díky studiu užívání jazyka ve společnosti nerovnost a ne-spravedlnost a oprávněně na ni upozorňuje, anebo má už předem jasno, jak se věci ve světě mají, a jen pro to hledá důkazy v textech.

Na kritickém postoji CDA ani korpusová lingvistika nic nezmění. Budeme-li ji však v duchu Widdowsonovy (1995) kritiky považovat za analýzu za předpokladu, že svou kritiku bude formulovat teprve tehdy, najde-li pro ni pokud možno nezpochybnitelnou oporu v datech, tj. nebude hledat argumenty pro předem jasnou interpretaci, může CL pomoci. Na rozdíl od dřívější, poněkud introspektivní tradice totiž nabízí mnohem snáze kontrolovatelné kroky všude tam, kde ji analytik uplatní. Konkrétněji viz v některých dalších bodech.

#### 4.2 TEORETICKÝ EKLEKTICISMUS

Poslední z fundamentálních výtek vůči CDA je absence jednoho komplexního teoretického modelu, který by se aplikoval na reálná data. CDA je vyčítáno, že si snad až trochu lehkovážně a utilitárně vybírá z rozličných teorií to, co se jí právě hodí, aniž by řešila, nakolik jsou jednotlivé části této teoretické směsice kompatibilní. Ruth

11 Výtek vůči CDA bylo vzneseno samozřejmě víc — např. že kořeny CDA jsou vzájemně nekompatibilní. Zde je však neprobírám — buď proto, že jen nepovažuji za stěžejní, nebo proto, že jsou z hlediska řešení prostřednictvím korpusové lingvistiky málo relevantní.

Wodaková (2001) však tuto skutečnost považuje naopak za výhodu — konkrétní téma, projekt či materiál podle ní vyžadují na míru sestavenou sadu nástrojů, neboť jde vždy o jiné texty, nadto v jiných kontextech, které zpřítomňují pokaždé jiné problémy. Výhodu v multidisciplinárním přístupu podobně vidí Teun A. van Dijk (1993, s. 252), který „chooses and elaborates theories, methods and empirical works as a function of their relevance for the realization of socio-political goals“. Widdowson (1998, s. 137) v tom však spíše než užitečnou možnost výběru a kombinaci analytických prostředků spatřuje jakousi *ad hoc* brikoláž bez ujasněných teoretických základů.

CACDA těží z CL tím, že používá její nástroje — velké korpusy jakožto elektronické prohledávatelné soubory textů a postupy hledání informací v těchto korpusích. Eklekticismus vyčítaný CDA se tím však spíše dále prohlubuje, byť ne snad na teoretické rovině, ale na rovině analytických prostředků.

K tomu poznamenávám, že eklekticismus je princip uplatňovaný v celé aplikované lingvistice — a CDA jakožto odnož studia diskurzu do aplikované lingvistiky nepochybně patřit může. Ani aplikovaná lingvistika, ani CDA se totiž nesnaží o postulování adekvátního koherentního modelu jazyka. Jde jí o co nejužitečnější, nejpronikavější popsání a vysvětlení toho, jak funguje konkrétní užití jazyka v kontextu. Pro takový záměr je teoretická koherence méně relevantní (v některých případech spíše irelevantní) a převažuje účelnost zohledněných metod, konceptů a přístupů.

### 4.3 NEREPREZENTATIVNOST ANALÝZ

Jedna z častých kritik směřuje k oprávněnosti závěrů a reprezentativnosti. Michael Stubbs (1997), Henry Widdowson (1998) nebo Jef Verschueren (2001) víceméně shodně upozorňují na to, že studie v rámci CDA často činí silné závěry založené na analýze velmi malého množství materiálu a/nebo dokládáné na jednotlivostech, o jejichž reprezentativnosti či významnosti vzhledem k diskurzu jako celku nemá čtenář obvykle žádné informace.

Haig (2004, s. 129–130) ovšem správně podotýká, že tato výtka platí spíše pro starší studie: „when critics write (in English, at least) about CDA they frequently equate CDA with earlier critical approaches to language such as Critical Linguistics (CL), either mistakenly or because the latter presents in many ways an easier target“ a že „when critics talk about CDA and really mean CDA, what they usually appear to have in mind is the work of Norman Fairclough“; já k tomu dodávám, že myslí-li Faircloughovy práce, tak patrně rané, nikoli ty revidované (viz též Breeze, 2011, s. 504). Sám Norman Fairclough později tyto metodologické nedostatky odstranil, např. jeho kniha *New Labour, New Language* (2000) je už založena na poměrně systematické analýze velkého množství materiálu. Ostatně — jak přiznává i Widdowson — už v době těchto výtek vznikaly práce, které se s fragmentárností/nereprezentativností a dalšími problémy dokázaly vypořádat díky využití korpusové lingvistiky, viz např. Krishnamurthy (1996).

Vynášela-li CDA příliš silné závěry na základě příliš kusých analýz, jsou metody CL jejím vhodným doplněním. Samy o sobě samozřejmě znamenají málo, neboť záleží především na tom, jak — tedy nakolik korektně a vhodně — se uplatňují, nicméně





potenciál je veliký. Dobře sestavený korpus dostatečné velikosti<sup>12</sup> může být reprezentativním vzorkem populacediskurzu, jež chce analytik zkoumat a jehož vlastnosti chce rozkrývat, nebo dokonce může být v jistém smyslu přímo populací,<sup>13</sup> jde-li o diskurz textově nepočtený. Reprezentativní vzorek spolu s validními a replikovatelnými analýzami dodává závěrům oprávněnou váhu. Pochopitelně i na korpusovém materiálu lze velmi snadno dělat fragmentární analýzy, letmý pohled do pár konkordancí vede zhruba tamtéž jako rozmáchlé rozbory několika málo vět v raných dobách CDA. To ovšem není v pravém smyslu využití korpusové lingvistiky.

#### 4.4 CHERRY-PICKING

S opomíjením velkého množství materiálu a dokládáním určitých tvrzení na příkladech s nejasným statusem vůči celku úzce souvisí i výhrada, že CDA si „vyzobává“ z dat jen ty části, které se jí hodí. Jinými slovy jde o sklon vybírat si jen ty texty či příklady, které jsou v souladu s interpretačním rámcem, perspektivou, předsudky či předpoklady badatele.

Tato kritika může pocházet ze dvou odlišných pohnutek. Buď čteme studii a v rozporu s tím, co tvrdí autor/autorka, víme o existenci protipříkladů, které prezentovanou interpretační linii rozrušují. Anebo je myslitelná obecnější situace, kdy čtenář jednoduše pociťuje, že závěry jsou nedostatečně podložené, že autor neučinil (nebo neukazuje, že učinil) kroky k nalezení protipříkladů, a zastává kritický postoj k důvěryhodnosti autorových/autorčiných tvrzení.

Cherry-picking trpí v literatuře jistou mírou terminologické neostrosti, není totiž vždy zřejmé, zda se tím myslí — idiomaticky řečeno — vybírání a ukazování „třešniček“ a zanedbávání „višniček“, anebo spíš vyvolávání dojmu, že takových „třešniček“ je dort plný, ačkoliv ve skutečnosti už žádné další k „vyzobání“ nejsou. Jinak řečeno, jde o upřednostňování jedné interpretace před druhou, anebo přehánění významnosti rysu využívaného pro interpretaci. Tak či onak je cherry-picking velmi blízce příbuzný s výše uvedenou nereprezentativností.

Samo zapojení CL do výzkumu pochopitelně ničemu takovému nezamezí. Jak však říká Victoria Kamasová (2017), „such approach helps to shift the researcher’s attention from what seems to be interesting and might confirm (unconscious) presumptions to what is demonstrably frequent, regular and forms some kind of a pattern“. Korpusu se lze ptát různými způsoby, ale v principu je díky němu obtížnější snažit se vědomě ohýbat data tak, aby působila určitým způsobem. Ke standardům CACDA patří uvádění frekvenčních seznamů slov či klíčových slov. Pokud je v takovém seznamu slovo, které by se zdálo stát proti analytickovým interpretacím, čtenář

12 Sama velikost však reprezentativnost nezajistí; kritickou reflexi viz např. u Chromého (2014).

13 Formulace „v jistém smyslu“ naznačuje, že jsem si vědom redukcionistického vyznění takového tvrzení. Diskurz, podle většinového chápání v CDA, není pouhý soubor textů, není tedy čímsi jako další jazykovou rovinou, nýbrž soubor textů, kontextů jejich vzniku (a možná i užívání), účastníků a pravidel, jimiž se jejich interakce řídí. To vše samozřejmě korpus neobsáhne, nemůže tedy být striktně vzato celou populací.

si ho snáze všimne. A pakliže by analytik takové slovo prostě bez dobrého vysvětlení vyřadil a zamlčel, mluvili bychom už nikoliv o v(ý)hodnosti CL pro CDA, nýbrž o frapantním porušování vědecké etiky.



#### 4.5 NEDOSTATEČNÉ ZOHLEDNĚNÍ KONTEXTU

Jak jsem již vyložil v oddílu 2.2, CDA vznikla z potřeby vysvětlit fungování, udržování, posilování atp. mocenských vztahů ve společnosti prostřednictvím užívání jazyka. Nejde tedy o studium jazyka samotného, jeho systému/struktury, nýbrž o sociální relevanci jeho konkrétních aplikací. Proto může být poněkud překvapivé, že se CDA vyčítá právě nedostatečná pozornost věnovaná kontextu. Je však třeba říct, že řada studií — a jde především o ty starší, a zejména britské — měla tendenci stavět své výklady na významu jednotlivého lexému či fráze bez většího ohledu na okolnosti užití, širší jazykový kontext atp. Jef Verschueren (2001, s. 60–79) tak např. ukazuje, jak Fairclough (1992) zanedbává situační kontext, přestože v něm jsou, nebo přinejmenším mohou být, přítomny i další pragmaticky relevantní faktory.

Na tuto výtku s úspěchem reaguje např. již zmíněný diskurzivně-historický přístup (*discourse historical approach*; DHA), možno též říci vídeňská škola CDA reprezentovaná především Ruth Wodakovou a jejími spolupracovníky.<sup>14</sup> DHA více než původní podoba ostrovní tradice dbá na zohlednění kontextu, a to nikoliv pouze jazykového či situačního, ale obecněji kulturního a historického. Považuje za důležité sledovat, co určitý koncept ve společnosti znamená, jaké jsou jeho vazby na další uzly kulturní sítě a jak se význam a užívání takového konceptu vyvíjí v čase. Tento diachronní aspekt jde proti neadekvátnímu ahistorismu, jenž může potenciálně zneplatnit závěry analýzy.

Jak známo, samotná tvorba (kompilace) korpusu s sebou nutně nese vytržení textů z jejich přirozeného kontextu, jakkoli je problematické s něčím takovým vůbec počítat, neboť každé nakládání s textem určitým způsobem mění soubor kontextových informací, jež se k němu váží. Principiálně je sice možné anotovat texty v korpusu tak, aby alespoň některé, vybrané informace o kontextu byly dohledatelné (autor, jeho sociodemografické či jiné charakteristiky, časové zařazení vzniku textu, jeho žánr atp.), avšak CDA by se měla snažit zohledňovat daleko bohatší kontextovou paletu. Toto zohlednění náleží kvalitativním fázím analýzy následujícím po korpusových krocích.

#### 4.6 PŘÍLIŠNÁ ZÁVISLOST NA ANALYTIKOVÍ

S přechozími souvisí i výtky, že CDA (zde ve smyslu samotného průběhu analýzy), analytické kroky a interpretace zjištění z nich vzešlých jsou příliš závislé na osobě analytika. Jef Verschueren (2001, s. 60–77) si stěžuje, že analytikovi (v jeho případě Faircloughovi) a jeho úsudku musíme prostě věřit. Je to totiž právě a jen analytik, který tvrdí, že něco je v diskurzu typické (nebo naopak exkluzivní, a proto

<sup>14</sup> V poslední době — už obohacená o korpusový přístup — se částečně transformovala v tzv. lancasterskou školu, viz oddíl 3.1.



významně) nebo že data vedou k tomu či onomu konkrétnímu kroku, závěru apod. Kritika tak částečně směřuje k obtížné replikovatelnosti výzkumu.

To je zjevně výtečná příležitost pro využití CL. Přirozeně zůstává na rozhodnutí analytika, kdy, kde a jaké nástroje CL využije. Smysl je v tom, že výsledky nejsou idiosynkratické, protože stejný krok vede, *ceteris paribus*, vždy ke stejnému výsledku. Jinými slovy, korektně užitá CL pomáhá zajišťovat replikovatelnost výzkumu.

#### 4.7 NEADEKVÁTNÍ TEORIE PUBLIKA

Mnoho dosud představených připomínek se vztahovalo k analýze textů — k tomu, jaké ideologicky příznakové významy podle CDA jeho autor do analyzovaného textu vepsal. Problém se ovšem podle kritiků ještě umocňuje ve chvíli, kdy se začne tvrdit, jak tyto významy působí. Ne nadarmo pojmenoval Michael Stubbs (1997) svůj článek *Whorf's children* — podle něj se CDA při svých závěrech spokojuje s naivním jazykovým (tedy vlastně diskurzním, pozn. OD) determinismem, ovšem v tom smyslu, že nikoliv možnosti jazyka, nýbrž diskurz (jeho obsah, povaha) předurčují myšlenkové procesy jeho uživatelů. Henry Widdowson (1998, s. 136) dodává, že kritickoanalytické práce působí dojmem, jako by snad jazyk byl tím jediným, co zodpovídá za význam. CDA se tím v očích kritiků nechtěně dostává zpět k přenosovému modelu komunikace s neproblematizovanou představou, že co se komunikuje, to publikum přijme. (Sluší se však poznamenat, že Shannon a Weaver (1949) byli ve skutečnosti v tomto ohledu dál, protože ve svém modelu počítali se šumem.)

Jak se tedy dobrat vlivu diskurzu na mysl čtenáře/posluchače/příjemce? Norman Fairclough (1989, s. 167) tvrdí, že analytik má oprávnění dělat intuitivní závěry, protože i analytik je v principu čtenář, který si je navíc vědom své vlastní pozice — a proto může být jeho čtení tím adekvátním. To nás přivádí k jistému rozporu, protože chce-li takový analytik tvrdit, že určitý lexém (nebo jiný jazykový prostředek) nějak ideologicky působí, říká tím zároveň, že i on je jím ovlivněn.

Jistěže existuje nějaký vztah mezi diskurzem a lidmi, kteří ho recipují, ale bylo by těžko obhajitelné předpokládat, že tento vztah je jednoduchý, přímočarý. Jak správně poznamenává Ruth Breezeová (2011, s. 508), můžeme sice říci, že užívání jazyka ovlivňuje kognici, avšak počítáme-li s užíváním jazyka jako s jediným zdrojem informací o ní, dostáváme se do bludného kruhu. Korektní by bylo podložit svá tvrzení o obsazích mysli příjemců výsledky zkoumání z jiných oblastí, než je jazyk. Bohužel ani Ruth Breezeová, ani Michael Stubbs (1997) nespecifikují, jaký model nakládání publika s obsahem diskurzu by byl podle nich hodný doporučení.

CL se obvykle vyjadřuje k produkční stránce užívání jazyka, neboť produkci, a nikoliv recepci, reprezentují data v korpusu obsažená.<sup>15</sup> Výzkumy čtení a obecněji recepce jazykových sdělení se provádějí různými metodami, kvalitativními i kvantitativními, využití korpusů v této oblasti mi však není známo. Na druhou stranu by podle mé intuice mělo být možné předpokládat, že opakované výskyty, tedy jedno-

<sup>15</sup> Něco trochu jiného je, že složení korpusu může být — a v případě přinejmenším prvních korpusů řady SYN ČNK i je — založeno recepčně (viz Čermák & Schmiedtová, 2004, s. 157).



duše vzato frekvence, mají nějakou souvislost s kognitivní „viditelností“. Jak říkají Costas Gabrielatos a Paul Baker (2008, s. 34) s odkazem na Michaela Hoeyho (2005), některé teorie v rámci korpusové lingvistiky tvrdí, že „people are unconsciously primed to infer various meanings due to the cumulative effect of all of their previous encounters with a word, collocation, or phrase“. Bylo by ovšem zkruslující pouze sčítat výskyty s nevyřčeným předpokladem, že každý z nich má stejnou hodnotu/váhu — texty mívají různý dopad s ohledem na své autory, způsoby/místa zveřejnění apod. (srov. Baker, 2006, s. 19–21).

Gabrielatosovo a Bakerovo (2008) tvrzení lze v principu vidět např. v zajímavé ukázce recepčního přístupu k analýze korpusových dat Fidlerové a Cvrčka (2015). Autoři porovnávají dva seznamy klíčových slov generované touž metrikou z téhož zkoumaného korpusu na pozadí dvou (především časově) rozdílných referenčních korpusů a jejich rozdílnost vztahují k rozdílným modelovým čtenářům. Referenční korpusy jsou pro ně reprezentace různých čtenářských zkušeností, a tudíž očekávání od dalších textů; tato očekávání podle nich tvoří klíč k interpretacím textu, potažmo diskurzu. Touto myšlenkovou konstrukcí se tudíž dostávají od jazykové produkce k recepci:

„We expect that [...] keywords extracted from a text with our method of keyword ranking (Difference Index—henceforth DIN) reflect how a text is interpreted by a typical reader of a specific period of time“ (Fidler & Cvrček, 2015, s. 199–200) — „Our foundational assumption is that a RefC [referenční korpus, pozn. O. D.] can approximate a model reader’s exposure to language patterns, which in turn reflect the typical reader’s point of view“ (*ibid.*, s. 204)

To je nepochybně podnětný přístup, přesto ho však považuji za značně reduktivní a v důsledku toho za omezeně validní. Předpoklad, že jazyková produkce, již je člověk vystaven, zásadně ovlivňuje jeho očekávání od textů, s nimiž se setká, jistě platí, jenže podle mého soudu jen omezeně.

Obávám se, že zatímco obecné tvrzení o primingu, jak ho formulují Gabrielatos a Baker (2008), je právě ve své vágnosti udržitelné, přímé ztotožnění korpusu s modelovým čtenářem jazykově žijícím v době zveřejnění textů obsažených v tomto korpusu je přece jen příliš odvážné.

V argumentaci Fidlerové a Cvrčka vidím tyto mezery: 1. Autoři svá zjištění pro SYN2010 vztahují k vágně vymezenému populárnímu názoru na zkoumané texty, doloženému jedním odkazem na publicistický článek. To je poměrně slabá opora pro tvrzení o shodě čtenářských postojů s klíčovými slovy. 2. Autoři argumentují mj. výzkumem Bermela et al. (2014), ten však s korpusovými frekvencemi porovnává posuzování gramatičnosti, nikoliv interpretace diskurzivních jevů. Mezi „čistě“ jazykovými jevy na jedné straně a interpretací obsahu, významu a hodnot na diskurzivní rovině je myslím přece jen příliš zásadní kvalitativní rozdíl na to, aby ji bylo možno zanedbat.

Shrnu-li uvedené kritické připomínky k CDA a jejich souvztažnost s CL, lze konstatovat, že nemalou část z uváděných problémů lze skutečně odstranit (nebo alespoň zmírnit) zapojením nástrojů korpusové lingvistiky do analytického procesu. Jak může takové zapojení vypadat v praxi, ukáže následující kapitola.

OPEN  
ACCESS

## 5. ANALÝZA JAZYKOVÝCH IDEOLOGIÍ V ČESKÉM PARLAMENTNÍM DISKURZU

### 5.1 DATA

Použití představeného přístupu teď předvedu na ilustrativní ukázce, která představuje jen úzký výřez plné analýzy, ale zachycuje všechny její fáze. Ukázka vychází z rozboru jazykových ideologií v českém parlamentním diskurzu (podrobněji Dufek, 2018). Východisko tvoří návrhy jazykových zákonů v Poslanecké sněmovně Parlamentu ČR od vzniku samostatné České republiky do současnosti (srov. Velčovský, 2015; Sojka, 2018).<sup>16</sup>

Zde analyzovaný materiál se skládá se dvou korpusů — jednoho zkoumaného a jednoho referenčního. Jádrem je korpus nazvaný JAZPARL obsahující sněmovní rozpravy o návrzích jazykových zákonů<sup>17</sup> a příslušné dokumenty (texty zákonů, důvodové zprávy, stanoviska vlády apod.) od vzniku České republiky do roku 2015 (viz tabulka 2). Texty zařazené do korpusu jsou zkopírovány z digitálního repozitáře stenoprotokolů rozprav volně přístupného ve Společné česko-slovenské digitální parlamentní knihovně.<sup>18</sup> Korpus je lemmatizován a analyzován pomocí softwaru #LancsBox (Březina, Timperley & McEnery, 2015).<sup>19</sup>

K tomuto zkoumanému korpusu byl pro potřeby analýzy klíčivosti stejným způsobem pořízen referenční korpus REFPARL, který obsahuje přepisy tematicky nespecifikovaných sněmovních rozprav — zvolil jsem vždy celou jednu předchozí a jednu následující schůzi po té, na níž se projednával některý z návrhů jazykových zákonů.

<sup>16</sup> V době započetí analýzy byl návrh zákona „o úředním jazyce České republiky a o změně zákona č. 200/1990 Sb., o přestupcích, ve znění pozdějších předpisů (zákon o úředním jazyce)“ (sněmovní tisk t0629) ze dne 21. října 2015 nejnovějším pokusem o legislativní regulaci jazyka. Poté ještě poslanci KSČM předložili podobný návrh v roce 2017, ale Poslanecká sněmovna ho — stejně jako zmíněný návrh z roku 2015 — do konce svého funkčního období neprojednala. Vliv jeho nepřítomnosti ve zkoumaném korpusu je tedy ve výsledku marginální.

<sup>17</sup> „Jazykový zákon“ je zde i dále označení neterminologické, zahrnující pro stručnost veškeré legislativní návrhy týkající se nějak úpravy věcí souvisejících s jazykem, tj. od dílčího pozměňovacího návrhu novely zákona o provozování rozhlasového a televizního vysílání přes samostatné návrhy zákona o státním/úředním/apod. jazyce až po návrhy ústavních zákonů.

<sup>18</sup> Viz <<http://www.psp.cz/eknih/index.htm>>.

<sup>19</sup> #LancsBox lemmatizuje data pomocí nástroje TreeTagger, který podporuje i češtinu, není však zcela spolehlivý. Proto se v analýzách např. jako samostatná klíčová slova objevují i nezákladní tvary, neboť jsou chybně nepřiráženy k (často) separátně existujícímu správně rozpoznávanému lemmatu (např. *etnický\_adjective* i *etnických\_adjective*). Dalším typem chyby, kterou TreeTagger generuje, je chybné slovnědruhovité tagování (např. *čl\_verb* — ve skutečnosti zkratka pro „článek“). To sice potenciálně znamená do určité míry zkreslení, na základě práce s daty popsané v následujícím oddíle ho však lze prohlásit za velmi malé. Při samotném nakládání s klíčovými slovy, jejich sémantickém sdružování a interpretaci už tyto chyby zohledňuji a v důsledku odstraňuji.



období	zákon
1995	ústavní zákon
1995/1996	zákon o státním jazyce České republiky a o jazycích národnostních menšin
1999/2000	zákon o jazyce České republiky
2002/2003	pozměňovací návrh k novele zákona o provozování rozhlasového a televizního vysílání
2003/2004	ústavní zákon
2015	zákon o úředním jazyce České republiky (neprojednáno)

**TABULKA 2:** Přehled návrhů jazykových zákonů.

Díky tomu, že se takto sestavený referenční korpus shoduje s korpusem zkoumaným z hlediska časového i žánrového, lze předpokládat, že jediná proměnná, kterou bude pomocí analýzy klíčových slov možné sledovat, bude téma, tedy v případě korpusu JAZPARL legislativní úprava jazykové oblasti.

Korpus JAZPARL dělím na dva subkorpusy — JAZPARL\_PRO a JAZPARL\_PROTI; viz tabulka 3. Tyto subkorpusy neodrážejí pouze diskurzivně konstituované skupiny aktérů, jejichž přesvědčení se projevují v promluvách — dělicí linie mezi nimi je demonstrována i reálně existujícími projevy, totiž hlasováním poslanců o příslušném návrhu zákona. Např. hlasování poslance pro jeho postoupení do dalšího čtení, resp. proti jeho zamítnutí (dohledatelné na webu poslanecké sněmovny) tedy znamená zařazení jeho promluvy do subkorpusu JAZPARL\_PRO.<sup>20</sup>

skupina	(sub)korpus	textů	tokenů	typů	lemmat
zkoumaný	JAZPARL	89	40 423	7 843	5 155
pro × proti	JAZPARL_PRO	38	28 671	6 249	4 112
	JAZPARL_PROTI	40	10 502	3 214	2 258
referenční	REFPARL	14	300 779	28 377	19 195

**TABULKA 3:** Velikost korpusů a subkorpusů.

## 5.2 ANALÝZA JAZYKOVÝCH IDEOLOGÍ V ČESKÉM PARLAMENTNÍM DISKURZU

Ideologie, včetně jejich konkrétních aplikací/realizací týkajících se jazyka, považuji za vlastní každému jedinci účastnícímu se sociálních interakcí a vstupujícímu do sociálních kontextů (viz např. Silverstein, 1979, s. 193; Woolard, 1998, s. 3. Tyto ideologie

<sup>20</sup> Do ani jednoho z těchto subkorpusů nebyly zařazeny promluvy poslanců, kteří se sice v průběhu projednávání k zákonu vyjádřili, nakonec však o něm nehlasovali nebo se zdrželi hlasování. V takovém případě je nebylo možno spolehlivě přiřadit k ani jedné skupině, v subkorpusech proto nejsou. V tabulce 3 proto subkorpusy JAZPARL\_PRO a JAZPARL\_PROTI mají dohromady jen 78 textů / 39 173 tokenů, přestože JAZPARL jako celek má o 11 textů / 1 250 tokenů víc. Z kritéria postoje projeveného hlasováním se vymykají dokumenty (texty zákonů, důvodové zprávy, stanoviska vlády apod.), které byly rozřazeny na základě převažujícího stanoviska na základě kvalitativního čtení.



mají potenciál formovat sociální akce, které se mohou střetávat se zájmy ostatních jedinců a skupin (např. McGroarty, 2008, s. 98). V tomto momentu se do úvah o ideologiích dostávají jako proměnné moc (příp. hegemonie apod.). Právě uplatňování moci s sebou nese riziko nespravedlnosti a utlačování. To je také perspektiva, která spoluutváří relevanci jakožto faktor při mém rozhodování a podrobení konkrétních částí diskurzu bližšímu zkoumání a jíž interpretuji výsledky analýzy.

Jazykové ideologie jsem ve zde představených datech zkoumal perspektivou komparace textů dvou diskurzních skupin — zastánců a odpůrců jazykových zákonů. Z prostorových důvodů se zaměřím pouze na druhou z nich, díky své reaktivní podstatě však zčásti osvětlím i povahu skupiny první.

K zodpovězení otázky, jaké jazykové ideologie jsou manifestovány v promluvách poslanců, použiju analýzu klíčových slov, kolokační analýzu, konkordanční analýzu a kvalitativní, interpretativní kritickou analýzu.

Klíčovými slovy rozumím lexikální jednotky, které mají ve zkoumaném korpusu významně vyšší frekvenci než v korpusu referenčním s ohledem na jejich disperzi. Výpočet klíčivosti kombinuje tři podstatné aspekty: zamítnutí nulové hypotézy (tj. že konkrétní slovo je ve zkoumaném korpusu frekventovanější jen náhodou) na určité hladině pravděpodobnosti, velikost efektu a zohlednění vnitrokorpusové variability. Proto využívá log-likelihood (která plní první požadavek) a Cohenovo  $d$  (druhý a třetí požadavek). Za klíčová budu tedy považovat taková slova, která budou mít kladnou hodnotu Cohenova  $d$  a zároveň hodnotu log-likelihood vyšší než 3,84 (což odpovídá  $p < 0,05$ ). Kritická reflexe konceptu klíčivosti viz Dufek (2018, s. 35–45).<sup>21</sup>

Jelikož jsou předmětem zájmu témata (koncepty, motivy), zaměřím se na plnovýznamová slova (bez zájmen a číslovek). Klíčová slova, která se zdají být významově související, vždy sdružím do sémantických skupin (toto sdružování dělám pouze na základě lemmat, ověřování v konkordancích je vždy předmětem následných kroků; názvy skupin jsou ilustrativní, rozhodně nikoliv jediné možné), poté uvedu zbývající klíčová slova seskupená podle slovnědruhové platnosti.<sup>22</sup>

21 Dufkův (2018) přehled způsobů výpočtu klíčových slov a jejich výhod a omezení ovšem nezahrnuje aktuální návrh Jesseho Egberta a Douga Bibera (2019). Autoři představují novou metodu určování klíčových slov — *text dispersion keyness*, která na rozdíl od *corpus frequency keyness* zcela opomíjí frekvenci a soustředí se pouze na disperzi slov v korpusu. Jejich kvalitativní zhodnocení seznamů klíčových slov generovaných jednou čistě frekvenčně založenou metodou, třemi metodami různými způsoby zohledňujícími disperzi a nakonec čistě disperzní metodou se zdá zřetelně favorizovat jejich vlastní *text dispersion keyness*. Je však třeba podotknout, že Egbert a Biber svou metodu nekonfrontují s přístupy používajícími metriky velikosti efektu. Důkladnější validizace jejich návrhu je tedy zatím otevřenou otázkou, přesto jde bezesporu o pozoruhodný a potenciálně zásadní příspěvek do debaty o adekvátnosti užívání klíčových slov v analýze diskurzu. Za upozornění na tuto práci děkuji recenzentovi.

22 Neuvádím chybně lemmatizované výrazy, tj. např. případy, kdy jsou jako lemmata uvedeny nezákladní tvary. Také přehlížím výrazy zjevně odkazující k sněmovnímu diskurzu, jako jsou *pan*, *kolega*, *předsedající*, *předkladatel*, *navrhovatel*, *paragraf* apod., kromě toho též rodná jména a příjmení.



Logika analýzy je následující:

Zprvė srovnám prozkoumám seznam klíčových slov vytvořený pro subkorpuse JAZPARL\_PROTI na pozadí subkorpuse JAZPARL\_PRO coby referenčního. To zvýrazní specifika zkoumaného subkorpuse vůči tomu druhému (specifika korpuse lze pak vztáhnout k specifickým subdiskurzům).

Zadruhé budu věnovat pozornost klíčovým slovům subkorpuse JAZPARL\_PROTI s referenčním parlamentním korpusem REFPARL. To ukáže specifika subkorpuse/subdiskurzu vůči obecnému sněmovnímu korpuse/diskurzu.

Zatřetí prověřím v návaznosti na druhý krok některá klíčová slova společná pro oba subkorpuse (tedy i JAZPARL\_PRO) pomocí kolokační analýzy, která pomůže ukázat, do jakých kontextů se klíčová slova prostřednictvím svých kolokátů zasazují.

Každý z těchto tří kroků bude podle potřeby průběžně podpořen detailnějším popisem užívání vybraných klíčových slov pomocí konkordanční analýzy a kvalitativního, kontextualizujícího kritického čtení.

Nejen v této ilustrativní sondě, ale i v plné analýze je přijatelné (nikoliv ideální, leč obhajitelné) nepodrobit všechna klíčová slova detailní analýze. Některá lze nechat stranou už jen na základě úvahy o jejich relevanci pro zkoumané téma, některá lze prověřit pohledem na výpis konkordancí, o některých může rozhodnout až podrobnější konkordanční analýza či přímo kvalitativní čtení širších kontextů. Podstatné je činit takové kroky transparentně.

V následujících oddílech věnuji pozornost zejména těm klíčovým slovům, která dobře plní ilustrativní funkci, a to buď ve vztahu k průběhu analýzy, nebo k zjištěným vzešlým z analýzy (jejichž užití je výstižně dokládají). Vzhledem k nutně omezenému rozsahu komentování postupu analýzy v tomto článku platí, že nejsou projevem strategie popsané v oddílu 4.4 jako cherry-picking, a že ta klíčová slova, která zde nepodrobují podrobnějšímu rozboru, byla vyhodnocena jako doklady vyjádření nepřispívajících k utváření jazykových ideologií, a to právě těmi postupy, které jsem načrtl. Kompletní analýza viz Dufek (2018, s. 53–98).

### 5.2.1 JAZPARL\_PROTI – JAZPARL\_PRO

Seznam klíčových slov subkorpuse JAZPARL\_PROTI vůči referenčnímu JAZPARL\_PRO obsahuje (opět mimo jiné) tato klíčová slova:

- **cizí:** slovensky, slovenština, Slovák, Morava;
- **legislativa:** řád, projev, zpracování, legislativně, přijmout;
- **hodnocení:** dostatečný, stávající, nutný, obvyklý, duplicitní, protiústavní, proklamativní, upravený.
- **substantiva:** závěr, titulěk, aplikace, pocit, výjimka, tvrzení, kvóta, soudruh, potíž;
- **adjektiva:** následující, faktický, jasný, chybný, konkrétní, stručný, pouhý, praktický;
- **slovesa:** souhlasit, mít, způsobit, být, omezovat, učit, vystupovat, spadat, odpovědět, hodit, naznačovat, vracet, přinést; —

OPEN  
ACCESS

— **adverbia:** jaksi, pouze, zvláště, fakticky, totiž, případně, bohužel, nikoliv, konceptně, přesně, shora, krásně, výjimečně.

Potenciál vyjadřovat postoje k projednávanému tématu nebo jeho protagonistům lze předpokládat u slov, jako jsou *soudruh*, *dostatečný*, *nutný*, *chybný*, *souhlasit* či *krásně*. Klíčové slovo *krásně* tento předpoklad jako jediné nepotvrzuje — jde pouze o dvě ocenění kvality projevu předřečníka. Ostatní slova však předpoklad potvrzují: *Soudruh* se sice objevuje pouze třikrát, ale jak ukazují příslušné texty, jeho funkce je delegitimizující. Všechna tři adjektiva vyjadřují většinově odmítavou pozici odpůrců jazykových zákonů: *dostatečná* je současná zákonná úprava, *nedostatečná* pak provázanost konkrétního návrhu s právním řádem, jeho určitost či jeho ochrana slovenštiny; polovina výskytu adjektiva *nutný* říká, že navrhovaná úprava není nutná, konceptně *chybný* je pak vždy návrh zákona.

Níže uvádím dva příklady projevů, které mj. pomocí uvedených (tučně zvýrazněných) klíčových slov vyjadřují signifikantní mínění jejich autorů stran jazyka a jeho legislativní regulace.

- (1) „Národní“ je ve své podstatě něco, co se dotýká jednotlivých národů, kmenů, a těch na našem území žil v průběhu dějin nemalý počet. Češi, Moravané, Slezané, Valaši, to vše jsou historické kmeny, které se naučily žít nejen pospolu, ale též s Němci, Poláky, Židy, **Slováky** a nevyjímaje ani Romy, Arabů a další. Proto by nebylo správné určovat češtinu za kulturní dědictví „jen“ Čechů, Moravanů a Slezanů. Čeština není jazykem jediného národa, je to jazyk občanů České republiky. [...] i díky cizím vlivům, se čeština v mluvené formě vyvíjí, a to do podoby, která vyhovuje všem, takže se stala jakýmsi univerzálním dorozumívacím nástrojem lidí žijících na území Čech, Moravy a Slezska. Jazyk je totiž jako živý organismus — žije, roste a vyvíjí se a žádná omezení nebo zákony na tomto nemohou nic změnit (to485-20040615-004-USDEU-Bielez).
- (2) Osobně vážím si snahy předkladatelů, hluboce však **nesouhlasím** s tím, že jazykovou kulturu u nás je možno zlepšovat zákonem, který nadto pouze apeluje a neukládá sankce. Dobré ovládnutí už je — a doufám bude — prestižní záležitostí zvyhodňující všechny, kteří dobrou češtinu používají. Český jazyk je poklad, který svobodně kultivujeme (to319-20000222-009-ODS-Nájemník).

V závěru prvního příkladu nacházíme projev jazykového organicismu (jazyk má podobné vlastnosti a projevy jako živé organismy), jinak je však celá citace zaměřena na obhajobu inkluzivnějšího pojetí jazyka, než je pojetí nacionální.

Druhý příklad je pozoruhodný tím, že obsahuje dva výrazy, které nejsou jednoznačně kompatibilní — *poklad* a *svobodně*. Interpretuji to tak, že *poklad* znamená cosi cenného a cenné věci je běžné chránit, což by předpokládalo spíše regulační opatření v tom smyslu, aby cennost nebyla ohrožena. Adverbium *svobodně* však jde významově proti takové představě. Na druhou stranu lze i svobodu pojímat jako hodnotu, která si zaslouží ochranu. Spojují se tak zde dva potenciálně, ale nejednoznačně odlišné pohledy na jazyk; jak uvidíme dále, není to případ ojedinělý.



1/3		2/3		3/3	
Pořadí	Lemma	Pořadí	Lemma	Pořadí	Lemma
5	<i>dámy_noun</i>	50	<i>nesouhlas_noun</i>	82	<i>zajišťovat_verb</i>
9	<i>pán_noun</i>	52	<i>titulek_noun</i>	83	<i>předkladatel_noun</i>
13	<i>kolega_noun</i>	54	<i>například_adverb</i>	84	<i>nerozumí_verb</i>
14	<i>předsedající_adjective</i>	56	<i>spotřebitel_noun</i>	86	<i>takový_pronoun</i>
16	<i>stručný_adjective</i>	57	<i>vedle_preposition</i>	88	<i>kučera_noun</i>
21	<i>slovensky_adverb</i>	60	<i>tlumočnicka_noun</i>	89	<i>nikterak_adverb</i>
22	<i>nadbytečný_adjective</i>	61	<i>předkladatelům_noun</i>	90	<i>rodném_adjective</i>
23	<i>projev_noun</i>	62	<i>ústavou_noun</i>	93	<i>spisovný_adjective</i>
24	<i>paragraf_noun</i>	64	<i>nepřijal_verb</i>	96	<i>živnostenském_adjective</i>
25	<i>často_adverb</i>	65	<i>slovenština_noun</i>	97	<i>kučery_noun</i>
27	<i>mít_verb</i>	66	<i>dostatečně_adverb</i>	98	<i>důvodová_adjective</i>
30	<i>hovořit_verb</i>	67	<i>velice_adverb</i>	100	<i>ověřeného_adjective</i>
32	<i>dostatečný_adjective</i>	68	<i>potřeba_noun</i>	102	<i>aplikační_adjective</i>
33	<i>řád_noun</i>	69	<i>česky_adverb</i>	103	<i>legislativy_noun</i>
34	<i>pouze_adverb</i>	71	<i>úředním_adjective</i>	105	<i>tolik_adverb</i>
35	<i>předkladatele_noun</i>	72	<i>předložený_adjective</i>	107	<i>občanství_noun</i>
37	<i>jaksi_adverb</i>	73	<i>legislativně_adverb</i>	108	<i>ustanovením_noun</i>
38	<i>předkladatelů_noun</i>	74	<i>upravující_adjective</i>	109	<i>bičků_noun</i>
39	<i>výjimka_noun</i>	75	<i>dobře_adverb</i>	110	<i>nějaký_pronoun</i>
40	<i>obsažený_adjective</i>	76	<i>navazovat_verb</i>	112	<i>právníckými_adjective</i>
42	<i>navrhovatelé_noun</i>	78	<i>řešit_verb</i>	113	<i>dobala_verb</i>
44	<i>navrhovatelů_noun</i>	ti	<i>65_numeral</i>		

**TABULKA 4:** Klíčová slova přítomná ve 100 neklíčovějších slovech subkorpusu JAZPARL\_PROTI vůči referenčnímu REFPARL a zároveň nepřítomná ve 100 neklíčovějších slovech subkorpusu JAZPARL\_PRO vůči referenčnímu REFPARL; šedě jsou uvedena klíčová považovaná za irelevantní, *kurzívou* chybně lemmatizovaná klíčová slova.

Už z prostého pohledu na klíčová slova a jejich skupiny lze získat hrubou představu o subkorpusu/subdiskurzu odpůrců návrhů jazykových zákonů — mimo jiné je orientován evaluativně. Ukazují to výše komentovaná adjektiva *dostatečný*, *nutný* a *chybný*. Blíže k jádru diskurzu se však dostaneme až v následujících dílčích analýzách, které se budou zabývat klíčovými slovy lépe odrážejícími podstatu subdiskurzu.

### 5.2.2 JAZPARL\_PROTI – REFPARL

V této části subkorpus JAZPARL\_PROTI vztáhnou k referenčnímu korpusu REFPARL. Vyjdu ze dvou seznamů, zde pro názornost dokumentovaných i tabulkami: Tabulka 4



ukazuje klíčová slova, která se vyskytují mezi sto nejvýznamnějšími v subkorpusu JAZPARL\_PROTI, a nejsou mezi první stovkou klíčových slov druhého subkorpusu (JAZPARL\_PRO). Po jejich analýze se dostanu k tabulce 5, která obsahuje klíčová slova sdílená mezi první stovkou oběma subkorpusy. Jinak řečeno, zaměřím se nejprve na komparativní specifika, poté na společné motivy.

Klíčová slova uvedená v tabulce opět pro větší názornost sdružím do skupin.

- **zahraničí:** slovenský, slovensky, slovenština; nerozumí;
- **český:** česky, rodném;
- **mluvení:** hovořit, projev;
- **hodnocení:** dobře, dostatečný, dostatečně, stručný, pouze, výjimka, nadbytečný;
- **legislativa:** dámy, pán, kolega, předsedající, paragraf, řád, předkladatel, předkladatele, předkladatelů, předkladatelům, navrhovatelé, navrhovatelů, ústavou, legislativy, legislativně, ustanovením, předložený, upravující, důvodová, nepřijal.
- **substantiva:** nesouhlas, titulek, spotřebitel, tlumočnicka, potřeba, občanství;
- **adjektiva:** obsažený, úředním, spisovný, živnostenském, ověřeného, aplikační, právníckými;
- **slovesa:** mít, navazovat, řešit, zajišťovat;
- **adverbia:** velice, často, jaksi, například, nikterak, tolik.

Lemmata *dobře*, *nadbytečný*, *potřeba*, *spisovný* a *hovořit* relevanci vzhledem k přesvědčením týkajícím se jazyka v analýze konkordancí a kontextů většinou nepotvrzují. Podobně ponechám stranou lemmata *hovořit*, *spisovný*, *dobře* a *potřeba*.

Adjektivum *nadbytečný* přesvědčení připisatelná předkladatelům jazykových zákonů odráží přece jen lépe: tři výskyty z pěti představují citace či parafráze jiných výroků (o nadbytečnosti zákonné úpravy jazykové situace), k nimž se připojuje nesouhlas, ve dvou výskytech se pak uznává nadbytečnost jednoho konkrétního článku návrhu zákona. Pro zajímavost připojuji jeden příklad ze subkorpusu JAZPARL\_PRO:

- (3) Nedomnívám se, že ohrožení českého jazyka neúměrně velkým působením cizích slov lze zabránit např. tím, že napíšeme nesouhlasný dopis tomu redaktorovi soukromé rozhlasové stanice, který hovoří jakýmsi „světovým českých jazykem“, jak o tom píše v Lidových novinách právě již zmíněný profesor filozofické fakulty bohemista Alexander Stich. Nesouhlasím ani s názorem ředitele Ústavu českého jazyka filozofické fakulty Univerzity Karlovy Oldřichem Uličným — jakkoli si obou bohemistů velmi vážím —, který mimo jiné nevíta tento zákon, protože, cituji: „Žijeme ve společnosti, která se nachází v hluboké morální depresi, ve společnosti bez ideálů, cílů, víry a vzorů, ale také bez respektu k zákonům. Od této společnosti lze stěží očekávat, že bude respektovat **nadbytečný** jazykový zákon.“ A já dodávám: to by také mohlo znamenat, že není-li v naší společnosti respekt k zákonům, nebudeme tedy přijímat zákony žádné (to319-19991102-009-KSČM-Svobodová).

Nyní přesunu svou pozornost k podobnostem obou subkorpusů.

		1/2				2/2			
JAZPARL_ PRO		JAZPARL_ PROTI		JAZPARL_ PRO		JAZPARL_ PROTI			
Cohenovo d	Pořadí	Lemma		Cohenovo d	Pořadí	Lemma		Cohenovo d	Pořadí
		Pořadí	Cohenovo d			Pořadí	Cohenovo d		
2,09	1	jazyk_noun	1	1,65	0,58	33	užívat_verb	18	0,50
1,03	2	český_adjective	7	0,61	0,57	34	úprava_noun	11	0,59
1,00	3	ústava_noun	58	0,37	0,57	35	kulturní_adjective	59	0,37
0,95	4	čeština_noun	4	0,71	0,57	36	styk_noun	47	0,40
0,88	6	národnostní_adjective	8	0,61	0,56	39	chránit_verb	19	0,45
0,88	7	zákon_noun	3	0,77	0,55	43	ochrana_noun	2	0,78
0,87	8	národní_adjective	70	0,36	0,55	44	menšina_noun	29	0,42
0,81	12	používání_noun	15	0,55	0,53	51	překlad_noun	26	0,42
0,77	13	jazykový_adjective	17	0,50	0,51	54	evropa_noun	111	0,32
0,73	14	úřední_adjective	10	0,59	0,51	56	francie_noun	55	0,38
0,67	17	právní_adjective	28	0,42	0,50	59	jiný_adjective	63	0,37
0,66	19	etnický_adjective	101	0,32	0,47	69	upravovat_verb	95	0,33
0,65	20	etnických_adjective	31	0,42	0,47	71	národnost_noun	20	0,44
0,64	21	menšin_noun	12	0,59	0,46	76	upravit_verb	43	0,41
0,63	25	cizí_adjective	6	0,62	0,46	78	slovenský_adjective	45	0,40
0,62	27	užívání_noun	51	0,39	0,44	91	předkladatelé_noun	41	0,41
0,61	28	identita_noun	53	0,38	0,43	100	listina_noun	104	0,32
0,60	29	příslušník_noun	49	0,39					

**TABULKA 5:** Klíčová slova přítomná ve 100 neklíčovějších slovech subkorpusu JAZPARL\_PRO vůči referenčnímu REFPARL a zároveň ve 100 neklíčovějších slovech subkorpusu JAZPARL\_PROTI vůči referenčnímu REFPARL; *kurzívou* chybně lemmatizovaná klíčová slova.

- **jazyk:** jazyk, jazykový, český, čeština;
- **atributy jazyka:** národní, úřední;
- **užívání:** užívat, užívání, používání;
- **zákony:** ústava, zákon, úprava, upravovat, upravit, listina;
- **menšiny:** národnost, národnostní, etnický, *etnických*, menšina, *menšin*, příslušník;
- **cizí:** Evropa, Francie, slovenský, cizí;
- **styk:** styk, překlad, jiný;
- **hodnoty:** kulturní, identita, chránit, ochrana.

Z klíčových slov obsažených v tabulce 5 je vidět, že se oba subkorpusy shodují ve zmiňování základních pojmů diskurzu o jazykových zákonech — kromě výrazů označujících legislativní stránku věci (*ústava, zákon, úprava, upravovat, upravit, listina*) zde najdeme lemmata přímo „jádrová“ — *jazyk, jazykový, český, čeština, národní, úřední; používání, užívání, užívat*. Obě strany se zjevně vyjadřují k právům menšin (*národnost,*

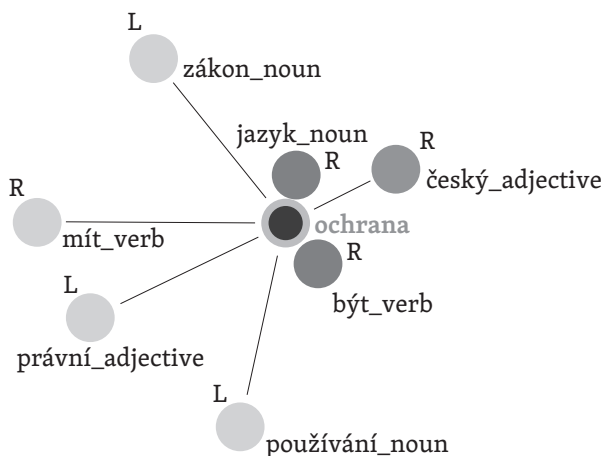


*národnostní, etnický(ch), menšin(a), příslušník*). Shodují se též ve vztahování se k zahraničním situacím, viditelném už v předchozím oddíle (*evropa, francie, slovenský, cizí*).

S ohledem na to, že podstatnou součástí jazykových ideologií jsou hodnoty, bude zajímavé podrobit bližšímu zkoumání skupiny klíčových slov, která do této skupiny řadím: 1) *chránit, ochrana*; 2) *kulturní, identita*.

### 5.2.3 KOLOKAČNÍ ANALÝZA

Obrázek 1 a tabulka 6 zachycují kolokace<sup>23</sup> lemmatu *ochrana* v subkorpusu JAZPARL\_PROTI. Podrobnější pohled do konkordancí prokázal, že lemmata z levého kontextu vypovídají o konceptualizaci *ochrany* v korpusu/diskurzu jen málo. Vpravo v subkorpusu JAZPARL\_PROTI však najdeme typicky (odhlédnuto od skloňování) sousloví *ochrana českého jazyka* (13 výskytů), pouze jednou jde o to, zda *český jazyk* potřebuje [takovouto] *ochranu*.



**OBRÁZEK 1:** Kolokace lemmatu *ochrana* v JAZPARL\_PROTI-REFPARL.

Pozice	Lemma	Log-likelihood	Frekvence kolokace	Frekvence kolokátu
vpravo	jazyk_noun	117,78	22	206
vpravo	být_verb	117,21	28	522
vpravo	český_adjective	80,49	15	125
vlevo	používání_noun	36,77	5	15
vlevo	zákon_noun	35,6	10	191
vlevo	právní_adjective	33,39	7	70
vpravo	mít_verb	16,77	5	100

**TABULKA 6** Kolokace lemmatu *ochrana* v JAZPARL\_PROTI-REFPARL

<sup>23</sup> Hodnota log-likelihood 6,63, frekvence kolokátu min. 5.





Jelikož je sousloví jako *ochrana českého jazyka* společné pro oba subkorpusy, mohli bychom usuzovat, že zastánci předkládaných jazykových zákonů ji budou prezentovat jako potřebnou, nutnou apod., zatímco odpůrci budou totéž zmiňovat, avšak odmítat, tj. že *ochrana českého jazyka* není žádoucí. Detailnější rozbor rozšířených kontextů a v některých případech i celých textů však vyznívá zřetelně méně jednoznačně. Uvedu několik příkladů ze subkorpusu JAZPARL\_PROTI.

- (4) **Český jazyk** je dostatečně chráněn speciálními právními předpisy pro jednotlivé obory a tento **zákon** není třeba (to319-20000331-004-ODS-Pejřil).
- (5) Nepochybuji o tom, že předkladatele návrhu vedl skutečně dobrý a ušlechtilý úmysl, tj. legislativně podpořit **ochranu českého jazyka**, který — jak všichni víme — je vystaven těžkým zkouškám, a to nejen v každodenním životě, v úředním styku, médiích, reklamě, a dovolíte-li, velice často i v tomto sále. [...] Mne překvapilo, že navrhovatelé **zákona** — přestože, jak říkám, důvod k jeho přípravě byl nepochybně ušlechtilý — se sami vzdali možnosti sankcionovat jednání, které bude v rozporu s literou či duchem **zákona**, a spoléhají na to, že použití češtiny bude v příslušných případech vynuceno jinými, již existujícími **zákony**. Jako příklad bych uvedl **zákon** na ochranu spotřebitele. Předložený **zákon** skutečně ani v jednom z paragrafů neuvádí, jaké sankce hrozí tomu, kdo nebude jednat v souladu s tímto **zákonem**, přičemž ani nestanoví, zda a v jakých situacích je čeština chráněna jako **jazyk** obecně, kdy je to spisovná čeština, kdy jsou to oprávněné dialekty atd. Ačkoliv já sám — stejně jako naprostá většina z nás, předpokládám — mám velmi často pocit, že čeština jako součást kulturní identity je devastována nad únosnou mez, nedomnívám se, že předložený **zákon** této devastaci v této podobě bude bránit účinněji než jiné aktivity a existující legislativa (to319-19991102-002-US-Matějů).
- (6) Smyslem návrhu **zákona** je — slovy předkladatelů, teď budu citovat: „rámcovým způsobem kodifikovat platné **právní** normy a zvyklosti založené v historickém vývoji tak, aby se v koncentrované normativní podobě zdůraznil morální, politický a **právní** příkaz zachovávat češtinu jako jeden ze základních prvků **české** kulturní identity“. Autoři **zákona** sami přiznávají, že smyslem **zákona** není zavádět nová opatření k **používání** a ochraně češtiny. Zákon si tedy klade za cíl češtinu chránit. Spadá tedy do skupiny **jazykových zákonů**, které mají bránit devastaci národního **jazyka**, zejména pronikáním různých -ismů — amerikanismů, germanismů atd. —, které souvisejí s kulturní a profesionální globalizací, a samozřejmě omezit **používání** cizích **jazyků** tam, kde se buď z praktických, či kulturních důvodů předpokládá použití národního **jazyka**, v našem případě češtiny (to319-19991102-002-US-Matějů).

Uvedené příklady ukazují, že zmínky o ochraně jazyka se rozpadají na několik skupin. Některé jsou neutrální nebo citátové (např. č. 6). Významná je skupina takových, které říkají, že češtinu není třeba chránit víc, než je tomu dosud (např. č. 4). Poslední



skupina tvrdí, že je potřeba češtinu chránit, a to i víc než dosud, ale nikoliv navrhaným způsobem, tedy předloženým zákonem (např. č. 5). Nikde jsem však nenalezl doklad zpochybnění potřeby ochrany jazyka vůbec, vždy jen ve vztahu ke konkrétní podobě předloženého návrhu nebo k zavádění zvláštní legislativy obecně.

Na příkladu (6) je zajímavé, že je povrchově poměrně neutrální, konstatační — nejen že nevyjadřuje žádnou námitku proti zákonu nebo jeho motivacím, on tyto motivace dokonce přijímá, když mluví o zákonech, které mají „bránit devastaci národního jazyka, zejména pronikáním různých -ismů — amerikanismů, germanismů atd.“. Probíhající nebo hrozící devastace se zde presuponuje jako podmínka braná za samozřejmou. Mluvčí reprezentuje nositele ideologie jazykového purismu.

Odpůrci se tedy se zastánci shodují víc, než by bylo možné předpokládat. Primární účel návrhů zákona nenegují, někdy jsou spokojeni s aktuální úrovní ochrany, a jindy dokonce podporují tezi, že je čeština ohrožena, ovšem rozcházejí se se zastánci v názoru na to, zda lze pro ně neblahý stav či trend zvrátit způsobem, který zastánci navrhuje. Dobře to ilustruje následující příklad:

- (7) Dámy a pánové, mně jde přece také o to, aby náš krásný národní jazyk byl chráněn, aby měl co nejvyšší prestiž a aby [byl; OD] rozvíjen. Pro mne ovšem otázka zní: pomůže tento zákon rozvoji našeho jazyka? Odpověď zní jednoduše — nepomůže. Bohužel zde je zákon slabou metodou k rozvoji jazyka (t0319-19991102-011-KDUČSL-Výborný).

Po rozboru konceptualizace ochrany přesunu svou pozornost k dvojici klíčových slov *kulturní* a *identita*. Nebudu zde předkládat výsledky kolokační analýzy, neboť žádné nejsou. Pro obě klíčová slova platí, že při nastavených kritériích nemají žádný významný kolokát. To však neznamená, že o nich nelze korpusovými nástroji nic zjistit. Konkordanční analýza podpořená v některých případech ověřováním přímo ve zdrojových textech ukazuje jejich poněkud ambivalentní vyznění.

Hned čtyři výskyty lemmatu *kulturní* najdeme v promluvě poslance Unie svobody Matějů z roku 1999:

- (8) Smyslem návrhu zákona je — slovy předkladatelů, teď budu citovat: „rámcovým způsobem kodifikovat platné právní normy a zvyklosti založené v historickém vývoji tak, aby se v koncentrované normativní podobě zdůraznil morální, politický a právní příkaz zachovávat češtinu jako jeden ze základních prvků české **kulturní** identity“. Autoři zákona sami přiznávají, že smyslem zákona není zavádět nová opatření k používání a ochraně češtiny. Zákon si tedy klade za cíl češtinu chránit. Spadá tedy do skupiny jazykových zákonů, které mají bránit devastaci národního jazyka, zejména pronikáním různých -ismů — amerikanismů, germanismů atd.-, které souvisejí s **kulturní** a profesionální globalizací, a samozřejmě omezit používání cizích jazyků tam, kde se buď z praktických, či **kulturních** důvodů předpokládá použití národního jazyka, v našem případě češtiny. [...] Ačkoliv já sám — stejně jako naprostá většina z nás, předpokládám — mám velmi často pocit, že čeština jako součást **kulturní** identity je devastována nad únosnou mez, nedomnívám se, že před-

ložený zákon této devastaci v této podobě bude bránit účinněji než jiné aktivity a existující legislativa (to319-19991102-002-US-Matějů).



První výskyt je citátový, ale zbylé tři ukazují, že Petr Matějů se s východisky navrhovatelů týkajícími se kulturního kontextu jazykové problematiky v zásadě ztotožňuje. Výstižný je např. poslední výskyt — *čeština jako součást kulturní identity* (míněno české, národní, „naší“ nebo podobně) je tu předkládána jako presupozice, platnost tohoto zapuštěného tvrzení se netematizuje a také širší kotext souzní s náhledem kladoucím češtinu do pozice ohrožené entity.

Ojedinelá je v rámci textů v subkorpusu JAZPARL\_PROTI připomínka poslance US-DEU Bieleseze, který odmítá spojovat češtinu s národní conceptualizací a říká, že „by nebylo správné určovat češtinu za kulturní dědictví, jen Čechu, Moravanů a Slezanů“; toto vyjádření v širším kontextu viz výše příklad (1).

Obraz toho, co je *kulturní* v očích odpůrců jazykového zákona, doplním stručnými poznámkami k *identitě* tamtéž. Polovina výskytů je citátová a byla již vzpomenuť výše (jde o českou *kulturní identitu*). Zbytek je z úst poslance ODS Pospíšila. Ten jednou shrnuje stanovisko vlády, podle něž „předložený návrh [ústavního zákona z roku 2003; OD] užívá nejasné a zároveň též zneužitelné pojmy a formulace, jako např. národní jazyk, státní a národní **identita**“ (to485-20040331-002-ODS-Pospíšil). Podruhé mluví za ústavně právní výbor, který „doporučuje, aby slovo **identita**, resp. **identity**, se nahradilo slovem *svébytnost*, resp. *svébytnosti*“ (to485-20040615-002-ODS-Pospíšil).

Celkově vzato odpůrci jazykových zákonů vykazují v oblasti *kulturní identity* podobné rysy, jaké jsem výše doložil pro motiv *ochrany*. Se svými kolegy z opačného názorového tábora se často shodují v premisách, odlišnost spočívá jen v postoji k navrženému závěru; jen někdy se mezi nimi objevuje fundamentálnější nesouhlas se spojováním jazyka s kulturou a identitou (jen) českého národa a zároveň poukaz na občanský, nikoliv národní rozměr soudobé České republiky.

## 5.2.4 SHRnutí VÝSLEDKŮ ANALÝZY

Podobný závěr, jaký jsem uvedl v předchozím odstavci, lze formulovat i pro podobu subdiskurzu odpůrců návrhů jazykových zákonů v české poslanecké sněmovně obecně. Poslanci, kteří hlasovali proti návrhům jazykových zákonů, se s podporovateli těchto návrhů navzdory protikladnému chování shodují víc, než by odpovídalo (reálným hlasováním podloženému) rozdělení na tyto dvě skupiny. Mnoha z nich byly společné jazykové ideologie jako např. axiologismus (jazyk má výjimečnou hodnotu), protektivismus (jazyk potřebuje neustálou ochranu) nebo purismus (domácí prostředky jsou lepší než cizí) (k terminologii viz Lanstýák, 2016). Charakteristické je pro ně úzké diskurzovní spojování motivů jazyka a národa, jazyka a (národní) kultury.

To je v ostrém rozporu s rámcem, do něž návrhy zákonů vstupovaly. Diskurz takto úzce spojující český jazyk s českým národem (což bylo v diskurzu patrné bez ohledu na to, jestli určitý konkrétní návrh připisoval jazyku atribut státní, národní nebo úřední) má podle mého přesvědčení několik problémů:



- 1) Konstruuje se jednonárodnostní stát, což hrubě opomíjí zdejší národnostní menšiny.
- 2) Implicitně vytváří skupiny s rozdílnými diskurzními právy a pozicemi — etnické Čechy, pro něž je čeština jejich jazyk s velkou symbolickou a identitní hodnotou, a ostatní, kteří se nehlásí k české národnosti a čeština jim přinejmenším částečně nepřínáleží i v případě, že je jejich primárním jazykem.
- 3) Spojování jazyka s národem neodpovídá českému ústavnímu rámci, který je založen na občanském principu, neboť preambule ústavy říká: „My, občané České republiky [...] přijímáme tuto Ústavu České republiky“.

To, že žádný z návrhů jazykových zákonů sněmovnou neprošel, nevešel v platnost a nereguloval tedy přímo životy lidí v Česku, nezpochybňuje potřebnost upozorňování na povahu diskurzu o těchto návrzích. Už samotné pronášení prezentovaných výpovědí a vzájemné ujišťování aktérů o relevanci a legitimitě jejich názorů představuje sociální praxi, která má potenciální praktické důsledky v jazykovém životě české společnosti. V diskurzu se vyjednávaly a upevňovaly ideologické konceptualizace, které přinejmenším měly potenciál marginalizovat nebo vylučovat jednu skupinu v porovnání s jinou, majoritní.

Nastíněná analýza jazykových ideologií tedy upozorňuje na nerovnost, ovšem v tomto případě netypicky nikoliv v postavení přímých aktérů diskurzu, tj. poslanců vyjadřujících se k návrhům jazykových zákonů, nýbrž spíše v zastoupení těch, jichž se potenciální důsledky dotýkají, tj. neetnických Čechů (nebo těch, kteří by nebyli zahrnuti do českého národa). Jinak viděno, analýza upozorňuje mimo jiné na překvapující míru sdílení jazykového purismu a úzkého diskurzního spojování motivů jazyka a národa, jazyka a (národní) kultury. Vzhledem k tomu, že sociální praxi (byť třeba ne explicitně záměrnou) směřující k odebírání symbolického kapitálu (tj. v tomto případě jazyka) (Bourdieu, 1998) určité skupině považují za projev uplatňování moci vedoucí k nerovnosti, interpretují převážnou povahu diskurzu o návrzích jazykových zákonů jako legitimní předmět kritiky v intencích kritické analýzy diskurzu.

## 6. ZÁVĚR

Kritická analýza diskurzu je v českém prostředí už relativně etablovaným rámcem zkoumání užívání jazyka v sociálním kontextu. Její výrazná současná podoba, korpusová CDA, zde však dosud zůstávala bez adekvátní teoreticko-metodologické reflexe. Tento text proto představil podstatu tohoto přístupu a zvážil přidanou hodnotu zapojení korpusových nástrojů jak obecně, tak ve vztahu k rozvoji CDA v návaznosti na její kritiky. Využití tohoto způsobu výzkumu sociální interakce pak prověřila ilustrativní kritická analýza jazykových ideologií v českém parlamentním diskurzu.

Věřím, že tento článek poslouží přinejmenším dvěma účelům. Zaprvé, že svou aplikační částí doloží užitečnost korpusového přístupu k analýze diskurzu, a zadruhé, že poskytne východisko pro další badatele v orientaci v bohatém a nesnadném poli solidně datově založené a metodologicky poučené analýze diskurzu. Stu-

die spadající do tohoto přístupu totiž mají potenciál pozitivně rozvíjet jak vědecké poznání jazykové interakce ve společenském kontextu, tak povahu těchto interakcí samotných.



## LITERATURA

- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London, NY: Continuum.
- Baker, P. (2014). “Bad wigs and screaming mimis”: using corpus-assisted techniques to carry out critical discourse analysis of the representation of trans people in the British press. In C. Hart & P. Cap (Eds.), *Contemporary Critical Discourse Studies* (s. 211–236). London: Bloomsbury.
- Baker, P. (2015). Does Britain need any more foreign doctors? Inter-analyst consistency and corpus-assisted (critical) discourse analysis. In: N. Groom, M. Charles & J. Suganthi (Eds.), *Corpora, Grammar and Discourse: In Honour of Susan Hunston* (s. 283–300). Amsterdam: John Benjamins.
- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). Sketching muslims: a corpus driven analysis of representations around the word ‘muslim’ in the British press 1998–2009. *Applied Linguistics*, 34(3), 255–278.
- Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Bednarek, M. (2008). Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*, 4(2), 119–139.
- Bermel, N., Knittl, L., & Russel, J. (2014). Absolutní a proporcionální frekvence v ČNK ve světle výzkumu morfosyntaktické variace v češtině. *Naše řeč*, 97(4–5), 216–227.
- Bourdieu, P. (1998). *Practical Reason: On the Theory of Action*. Stanford, CA: Stanford University Press.
- Breeze, R. (2011). Critical discourse analysis and its critics. *Pragmatics*, 21(4), s. 493–525.
- Březina, V., Timperley, M., & McEnery, A. (2015). #LancsBox v. 4.5. Dostupné z <<http://corpora.lancs.ac.uk/lancsbox/>>.
- Caldas-Coulthard, C. R. (1993). From discourse analysis to critical discourse analysis: the differential re-presentation of women and men speaking in written news. In J. M. Sinclair, M. Hoey, & G. Fox (Eds.), *Techniques of Description: Spoken and Written Discourse* (s. 196–208). London: Routledge.
- Caldas-Coulthard, C. R. (1995). Man in the news: the misrepresentation of women speaking in news-as-narrative discourse. In S. Mills (Ed.), *Language and Gender: Interdisciplinary Perspectives* (s. 226–239). Edinburgh Gate: Longman.
- Cap, P. (2018). Spatial Cognition. In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge Handbook of Critical Discourse Studies* (s. 92–105). London: Routledge.
- Čermák, F., & Schmiedtová, V. (2004). Český národní korpus — základní charakteristika a širší souvislosti. *Národní knihovna*, 3, 152–168.
- Dufek, O. (2018). *Kritická analýza jazykových ideologií v českém veřejném diskurzu*. Disertační práce. Praha: FF UK.
- Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analysis. *Corpora*, 14(1), 77–104.
- Elmerot, I. (2017). *These womens’s verbs. A combined corpus and discourse analysis on reporting verbs about women and men in Czech media 1989–2015*. Master thesis. Stockholm: Stockholms universitet. Dostupné z: <<http://www.diva-portal.org/smash/record.jsf?pid=diva2:1161582>>.
- Fairclough, N. (1985). Critical and descriptive goals in discourse analysis. *Journal of Pragmatics*, 9(6), 739–763.



- Fairclough, N. (1989). *Language and Power*. London: Longman.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity.
- Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*. London: Longman.
- Fairclough, N. (2000). *New Labour, New Language?* London: Routledge.
- Fidler, M., & Cvrček, V. (2015). A data-driven analysis of reader viewpoints: reconstructing the historical reader using keyword analysis. *Journal of Slavic Linguistics*, 23(2), s. 197–239.
- Fidler, M., & Cvrček, V. (2018). Going Beyond “Aboutness”: A Quantitative Analysis of Sputnik Czech Republic. In M. Fidler & V. Cvrček (Eds.), *Taming the Corpus. From Inflection and Lexis to Interpretation* (s. 195–225). Cham: Springer.
- Flowerdew, J. (1997). The discourse of colonial withdrawal: a case study in the creation of mythic discourse. *Discourse and Society*, 8(4), 453–477.
- Flowerdew, J., & Richardson, J. E. (2018). Introduction. In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge Handbook of Critical Discourse Studies* (s. 1–10). London: Routledge.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: a corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996–2005. *Journal of English Linguistics*, 36(1), 5–38.
- Haig, E. (2004). Some observations on the critique of critical discourse analysis. *Studies in Language and Culture*, 25(2), 129–149.
- Hammersley, M. (1997). On the foundations of critical discourse analysis. *Language and Communication*, 17(3), s. 237–248.
- Hardt-Mautner, G. (1995). ‘Only Connect’: *Critical Discourse Analysis and Corpus Linguistics* [online]. Lancaster: University of Lancaster. Cit. 20. 3. 2018. <<http://ucrel.lancs.ac.uk/papers/techpaper/vol6.pdf>>.
- Hart, C. (2015). Cognitive Linguistics and Critical Discourse Analysis. In: E. Dąbrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (s. 322–346). Berlin: Mouton De Gruyter.
- Hart, C. (2019). Spatial Properties of ACTION Verb Semantics: Experimental Evidence for Image Schema Orientation in Transitive vs. Reciprocal Verbs and Its Implications for Ideology. In: C. Hart (Ed.), *Cognitive Linguistic Approaches to Text and Discourse. From Poetics to Politics* (181–204). Edinburgh: Edinburgh University Press.
- Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge.
- Hořejší, M. (2016). Zapojení korpusových metod do kritické analýzy diskurzu. Dostupné z <[http://www.academia.edu/29421175/Zapojen%C3%AD\\_korpusov%C3%BDch\\_metod\\_do\\_kritick%C3%A9\\_anal%C3%BDzy\\_diskurzu](http://www.academia.edu/29421175/Zapojen%C3%AD_korpusov%C3%BDch_metod_do_kritick%C3%A9_anal%C3%BDzy_diskurzu)>.
- Hořejší, M. (2019). Karel Klostermann a diskurz o Šumavě. *Slovo a slovesnost*, 80(2), 105–127.
- Charteris-Black, K. (2004). *Corpus Approaches to Critical Metaphor Analysis*. Basingstoke: Palgrave-MacMillan.
- Charteris-Black, J. (2014). *Analysing Political Speeches: Rhetoric, Discourse and Metaphor*. Basingstoke & New York: Palgrave-MacMillan.
- Chilton, P. (2011). Still something missing in CDA. *Discourse Studies*, 13(6), 769–781.
- Chilton, P. (2014). *Language, Space and Mind: The Conceptual Geometry of Linguistic Meaning*. Cambridge: Cambridge University Press.
- Chromý, J. (2014). Korpus a reprezentativnost. *Naše řeč*, 97(4–5), 185–193.
- Kamasa, V. (2017). Corpus linguistics for critical discourse analysis: what can we do better? In P. Pězik & J. Waliński (Eds.), *Language, Corpora and Cognition* (s. 221–239). Frankfurt am Main: Peter Lang.
- Krishnamurthy, R. (1996). Ethnic, racial and tribal: the language of racism? In C. R. Caldas-Coulthard — M. Coulthard (eds.), *Texts and Practices: Readings in Critical Discourse Analysis* (s. 129–149). London: Routledge.
- Lehečková, E. (2016a). Analýza diskurzu a kritická analýza diskurzu. In K. Šebesta et al., *Aplikovaná lingvistika: Příručka pro studenty Bc. studia ČJL* (s. 78–95). Praha: Karolinum.

- Lehečková, E. (2016b). Kritická analýza diskurzu v kontextu korpusové a kognitivní lingvistiky. In J. Hasil (Ed.), *Přednášky z 59. běhu Letní školy slovanských studií* (s. 65–76). Praha: FF UK.
- Li, W. (2016). Rethinking Critical Metaphor Analysis. *International Journal of English Linguistics*, 6(2), 92–98.
- McGroarty, M. (2008). The political matrix of linguistic ideologies. In: B. Spolsky & F. M. Hult (Eds.), *The Handbook of Educational Linguistics* (s. 98–112). Oxford: Blackwell.
- Musolff, A. (2016). *Political Metaphor Analysis*. London: Bloomsbury Academic.
- O'Halloran, K. (2007). Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics*, 28(1), 1–24.
- Orpin, D. (2005). Corpus linguistics and CDA: examining the ideology of sleaze. *International Journal of Corpus Linguistics*, 10(1), 37–61.
- Partington, A. S. (2002). *The Linguistics of Political Argument: The spin-doctor and the wolfpack at the White House*. London: Routledge.
- Partington, A. S. (2006). Metaphors, motifs and similes across discourse types: corpus-assisted discourse studies (CADS) at work. In A. Stefanowitsch & S. Th. Gries (eds.), *Corpus-Based Approaches to Metaphor and Metonymy* (s. 267–304). Berlin: Mouton de Gruyter.
- Partington, A. S. (2008). The armchair and the machine: corpus-assisted discourse studies. In C. T. Torsello, K. Ackerley & E. Castello (Eds.), *Corpora for University Language Teachers* (s. 189–213). Bern: Peter Lang.
- Partington, A. S., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam: John Benjamins.
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana-Champaign, IL: University of Illinois Press.
- Silverstein, M. (1979). Language structure and linguistic ideology. In: P. R. Cline, W. F. Hanks & C. L. Hofbauer (Eds.), *The Elements: A Parasession on Linguistic Units and Levels: April 20–21, 1979: Including Papers from the Conference on Non-Slavic Languages of the USSR: April 18, 1979* (s. 193–247). Chicago, IL: Chicago Linguistic Society.
- Slavičková, T. (2015). Investigating nepřizpůsobivý as a key word in critical analysis of Czech press reports on Roma. *Korpus — gramatika — axiologie*, 11, 69–82.
- Sojka, P. (2018). Zakotvení statusu českého jazyka v našem právním řádu. *Nová čeština doma a ve světě*, 7(2), 14–20.
- Stubbs, M. (1994). Grammar, text and ideology: computer-assisted methods in the linguistics of representation. *Applied Linguistics*, 15(2), 201–223.
- Stubbs, M. (1996). *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Oxford: Blackwell.
- Svobodová, J., Hinková, B., & Hamrusová, Š. (2018). *Kritické čtení mediálních textů*. Olomouc: Vydavatelství Univerzity Palackého.
- van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse and Society*, 4(2), 249–283.
- van Dijk, T. A. (2000). *Ideology and discourse: A Multidisciplinary Introduction*. Dostupné z <<http://www.discourses.org/OldBooks/Teun%20A%20van%20Dijk%20-%20Ideology%20and%20Discourse.pdf>>.
- van Leeuwen, T. (2005). *Introducing Social Semiotics*. London: Routledge.
- Velčovský, V. (2015). Pokusy o zavedení jazykového zákona v České republice. *Nová čeština doma a ve světě*, 4(1), 41–49.
- Verschueren, J. (2001). Predicaments of criticism. *Critique of Anthropology*, 21(1), 59–81.
- Widdowson, H. G. (1995). Discourse analysis: a critical view. *Language and Literature*, 4(3), 157–172.
- Widdowson, H. G. (1998). The theory and practice of critical discourse analysis. *Applied Linguistics*, 19(1), 136–151.
- Wodak, R. (2001). The discourse-historical approach. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (s. 63–95). London: Sage.
- Wodak, R. (2011). Critical linguistics and critical discourse analysis. In J. Zienkowski, J.-O. Östman & J. Verschueren (Eds.), *Discursive*



- Pragmatics* (s. 50–69). Amsterdam: John Benjamins.
- Wodak, R., & Meyer, M. (2009). Critical discourse analysis: history, agenda, theory, and methodology. In R. Wodak & M. Meyer (Eds.), *Methods for Critical Discourse Analysis* (s. 1–33). London: Sage,
- Wodak, R., & Meyer, M. (2015). *Methods of Critical Discourse Studies*. London: Sage.
- Woolard, K. A. (1998). Language ideology as a field of inquiry. In B. B. Schieffelin, K. A. Woolard, & P. V. Kroskrity (Eds.), *Language Ideologies: Practice and Theory* (s. 3–47). Oxford: Oxford University Press.
- Zasina, A. J. (2018). Image of Politicians and Gender in Czech Daily Newspapers. In M. Fidler & V. Cvrček (Eds.), *Taming the Corpus. From Inflection and Lexis to Interpretation* (s. 167–194). Cham: Springer.

**Ondřej Dufek** | Ústav pro jazyk český AV ČR, Praha  
<dufek@ujc.cas.cz>