

Srovnání kolokací totalitního jazyka v bývalém Československu s jazykem současným¹

Věra Schmiedtová (Praha)

A COMPARISON OF COLLOCATIONS IN THE TOTALITARIAN LANGUAGE OF THE FORMER CZECHOSLOVAKIA AND THOSE IN CONTEMPORARY CZECH

The goal of this small sample study is to explore what possible effects a political situation and the ideologization of public discourse may have on language and its lexis. Using two corpora, one composed of the 1952, 1969 and 1977 texts characterizing the totalitarian language of the former Czechoslovakia and the other based on contemporary texts, the study compares the seven most frequent collocates of seven selected lemmas (four nouns and three adjectives) found in the Totality corpus and the control contemporary language corpus. The comparison shows that long-term use of ideologically motivated word combinations may influence the associative semantic component of otherwise neutral words and result in change of their semantic prosody, or at least in abuse of their prosody. Hence the delimitation of a totalitarian discourse vocabulary must take into account not only ideologically marked lexis, but also the syntagmatic aspects of the discourse.

KEYWORDS

totalitarian language, contemporary language, collocations, corpus-based comparison, ideologization, connotation, prosody

KLÍČOVÁ SLOVA

totalitní jazyk, současný jazyk, kolokace, korpusové srovnání, ideologizace, konotace, prozodie

O. ÚVOD

Smyslem studie je naznačit, k jakým posunům může v jazyce docházet vlivem politické situace a ideologizace veřejného diskurzu. Na jazyce totalitního období lze ukázat některé rysy příznačné pro tento proces. Lze se domnívat, že tyto rysy budou společné každému totalitnímu diskurzu bez ohledu na to, jaké ideologii slouží. V tomto smyslu má být tato studie nejen dílčí připomínkou naší nedávné historie, ale i příspěvkem k obecnému popisu totalitního diskurzu.

Práce na slovníku komunistické totality, z které studie čerpá, ukázala, že slovník totality nelze jednoduše ztotožnit s podmnožinou ideologicky zabarvených slov. Jde o mnohem rafinovanější záležitost, kdy i neutrální slova mohou být zneužívána tím, že se systematicky mění jejich kolokační okolí a tím někdy i komponenty jejich

¹ Tato studie vznikla v rámci Programu rozvoje vědních oblastí na Univerzitě Karlově č. P11 Český národní korpus.

významu. Ovlivněna může být významová složka konotativní/asociativní, afektivní, stylistická a zmíněná kolokační, přičemž podstatnou roli zde sehrává efekt frekvence. Právě na kolokační složku významu se tato studie zaměřuje.

1. SLOVNÍK KOMUNISTICKÉ TOTALITY

Překládaná studie navazuje na publikaci *Slovník komunistické totality* (2010), která je ojedinělým pokusem objektivně zachytit specifický charakter politického diskurzu v období 1948–1989 v naší republice. Protože studie čerpá z dat shromážděných ve *Slovníku*, je na místě uvést stručnou informaci o této publikaci. *Slovník*, který sestavil kolektiv autorů Ústavu Českého národního korpusu FF UK, si klade za cíl „přinést na datech korpusu statisticky podložený výzkum lexikonu, resp. jeho jednotek, a ukázat jejich užití“. Jak vyplývá i z názvu, šlo o „výzkum specificky orientovaný na jazyk využívaný a zneužívaný k propagandě a manipulaci“. Protože jde o nejdelší období v trvání naší republiky, omezili autoři výběr textů pro sestavení korpusu Totalita na roky 1952, 1969 a 1977, reprezentující tři zásadní etapy komunistického období. Korpus je tvořen dvěma soubory textů: první obsahuje část ročníků deníku *Rudé právo* (ústředního tiskového orgánu KSČ) z těchto tři let (dvě třetiny korpusu), druhý soubor obsahuje knihy a tiskoviny (třetina korpusu) vydané ve stejné době, tj. 1952 (23 titulů), 1969 (10 titulů) a 1977 (58 titulů). Dlouholetá příprava korpusu zahrnovala skenování a slovnědruhově tagování textů, jehož výsledkem byl korpus o rozsahu 15 350 741 slov (tokens), z toho je 422 762 různých slovních tvarů (types) a 164 815 lemmat. Úvodní část publikace obsahuje esej historika o proměnách komunistického režimu a stručnou lingvistickou studii shrnující základní poznatky o slovníku komunistické totality, jeho lexémech, nominacích a užitích. Jádrem publikace jsou dva slovníky, abecední *Specifický totalitní slovník* a *Základní slovník podle frekvence*. Publikaci uzavírá krátká příloha s ilustračními materiály a texty.

Metodologicky zajímavý je postup při sestavování obou slovníků, byť je zmíněn jen stručně. Zatímco obsah frekvenčního slovníku je primárně závislý na výběru textů v korpusu a technické přípravě vstupních dat, vygenerování abecedního *Specifického slovníku* je výsledkem složitého statistického filtrování. Tím ovšem není řečeno, že tvorba frekvenčního slovníku je snadná a automatická. Pro účely publikace bylo nutno zredukovat více než 160 tisíc lemmat zjištěných v korpusu Totalita na únosnou délku. Frekvenční slovník se tak omezuje na 9994 lemmat s frekvencí vyšší než 63. Vedle celkové absolutní frekvence je u lemmatu uvedena i jeho relativní frekvence v každém textovém souboru ze tří zvolených let. Kompletní frekvenční slovník je pak obsažen na CD přiloženém k publikaci. Abecední *Specifický totalitní slovník* (počet hesel není uveden) nevznikl frekvenčním ohraničením, nýbrž na základě selekce trojího typu, přičemž „specifický“ znamená „v porovnávání s dnešním jazykem“ (reprezentovaným korpusem SYN2009PUB). Zahrnuje za prvé lexémy, které se již v dnešním jazyce nevyskytují (ve slovníku označeny křížkem), za druhé slova statisticky významně častější v korpusu Totalita než v současném korpusu (a dobově příznaková) a za třetí slova vyskytující se v obou korpusech, ale v korpusu Totalita

ve specifických kontextech a kolokacích. Právě poslední skupina běžně neutrálních, bezpříznakových slov je předmětem výzkumu této studie.

2. CÍL A METODA

Z hlediska sémantiky je patrně nejzajímavější skupina neutrálních slov, u nichž se specifická projevuje kontextuálně, odlišným spektrem kolokací, do nichž vstupují s charakteristickou frekvencí v obdobích s různým společensko-politickým klimatem. Jazyk užívaný v konkrétním dobovém společensko-politickém kontextu reprezentovaný korpusem Totalita a kontrolním korpusem bude označován jako totalitní diskurz a současný diskurz. Termínem kolokace je zde míněn jakýkoli gramaticky a sémanticky přirozený souvýskyt dvou slov (bigram) bez ohledu na to, zda toto spojení má nebo nemá kompozicionální význam. Výchozím bodem je tu předpoklad, že význam a užití slova lze charakterizovat jeho kolokacemi. Ale protože cílem je charakteristika dvou srovnávaných temporálních variant diskurzu, u kolokací nás bude spíše než síla vazby mezi oběma slovy (jejich vzájemná expektance) zajímat jejich absolutní frekvence, která o diskurz něco vypovídá. Z tohoto hlediska je nejpříznačnější kolokace ta, která je nejčastější. Mění-li se u slova okruh jeho typických kolokací v čase, lze předpokládat, že dochází k posunu v diskurzu (a potenciálně i v některých složkách jeho významu). Cílem studie je ilustrovat na malém vzorku slov pomocí srovnání jejich tehdejších a současných kolokací nejen tento posun, ale i to, že dobová politická situace může překvapivě zasahovat i do vnitřní sémantické struktury jinak neutrálních slov. Počet nejfrekventovanějších kolokací zvolených slov byl z praktických důvodů omezen na sedm. Vzorek slov byl zadán do cílového a kontrolního korpusu a jejich kolokace zjišťovány pomocí vyhledávacího programu. U substantiv byly vyhledávány kolokace s adjektivy, u zkoumaných adjektiv kolokace se substantivy. Vlastní analýza spočívá v porovnání okruhu kolokací téhož slova v cílovém a kontrolním korpusu a interpretaci odlišností.

2.1 VZOREK ZKOUMANÝCH SLOV

S ohledem na metodu kolokačního popisu, na které je studie postavena, vycházel výběr slov z těchto faktorů: frekvence, slovního druhu a (ideologické) neutrálnosti, bezpříznakovosti (v obecném jazyce a z pohledu dnešního). Z frekvenčního hlediska šlo o autosémantická slova z prvního sta lemmat v *Základním slovníku podle frekvence* obsaženém ve *Slovníku komunistické totality*. Mělo jít o slova frekventovaná, u nichž lze předpokládat hojný počet kolokátů. Ze stejného důvodu padla volba na substantiva a adjektiva, u nichž lze kolokace snáze sledovat a porovnávat než u sloves nebo příslovcí. Navíc substantiva a adjektiva se jako nositelé pojmů a příznaků zdají být pro účely studie nejvhodnější.

Z rozboru prvních sta slov ve frekvenčním slovníku vyplynulo, že téměř polovinu, 46, tvoří funkční slova (předložky, spojky, částice, k nimž jsou pro zjednodušení počítána i zájmena, slovesa modální a sponová a číslovky), která kompletně

obsadila prvních 20 pozic podle frekvence. Autosémantika tvoří zbytek, tedy 54, položek. Volbu mezi těmito 54 položkami určovala snaha vybrat pojmenování pokud možno bezpříznaková, neutrální (vyloučena byla proto slova jako *sovětský* a *socialistický*). Zároveň bylo přihlíženo k distribuci v rámci prvních sta slov, tj. vybírána byla lemmata na začátku, v polovině a na konci tohoto rozmezí. Fakt, že první slovo, *nový*, se objevilo až na 30. místě mezi prvními sty slovy, je dán tím, že první autosémantikem *mít* se vyskytlo teprve na 21. místě a *nový* je sedmým nejfrekventovanějším autosémantikem.

V rámci těchto parametrů byla volba do určité míry arbitrární. Výsledný vzorek obsahuje čtyři substantiva, *život*, *člověk*, *organizace*, *boj*, a tři adjektiva, *nový*, *vysoký*, *americký*. Zajímavé je porovnání pořadí těchto sedmi slov podle frekvence s jejich frekvenčním umístěním v synchronním korpusu SYN2010², které přináší Tabulka 1:

korpus Totalita		korpus SYN2010	
lemma	pořadí	lemma	pořadí
nový	30	člověk	42
člověk	41	nový	65
organizace	50	život	96
boj	73	vysoký	126
vysoký	88	americký	360
život	89	organizace	520
americký	90	boj	664

TABULKA 1. Pořadí zkoumaných slov podle frekvence v korpusu Totalita a SYN2010

Ukazuje se, že v synchronním korpusu se v první stovce nejfrekventovanějších výrazů objevují pouze tři ze sedmi sledovaných slov, *člověk*, *nový* a *život*. Víceméně shodné umístění má v obou korpusech pouze lemma *člověk*. Nejbližší k nim v pořadí má ještě adjektivum *vysoký*, zatímco tři zbylá slova se z hlediska frekventovanosti posunula výrazně níž. Zejména markantní je to u slova *boj*. Vzhledem k tomu, že korpus SYN2010 se liší od korpusu Totality svým rozsahem i složením, nelze z rozdílu v pořadí slov spolehlivě vyvozovat dalekosáhlé závěry. Nicméně je nepochybné, že k určitému posunu ve veřejném diskurzu zde z hlediska frekvence užívání tohoto slova určitě došlo. Konkrétně pořadí výskytu slov *americký*, *organizace* a *boj* signalizuje, že jejich význam poklesl.

3. POUŽÍVANÉ KORPUSY A SOFTWARE

Pro srovnání kolokací, které se vyskytly u týchž slov v různých obdobích, byly použity dva korpusy. Jako cílový korpus slouží výše popsáný korpus Totalita (viz oddíl 1),

² Žánrově vyvážený stomilionový reprezentativní korpus, který obsahuje psané synchronní texty z let 2005–2009.

na základě něhož vznikl *Slovník komunistické totality*. Druhý, označovaný jako kontrolní korpus, byl nově vytvořen pro účely této studie ze současných novinových/publicistických textů, které reprezentují období od roku 1989 dodnes. Je sestavený z textů z různých let a měsíců tak, aby vznikl korpus srovnatelný s korpusem Totality co do rozsahu a typu textů. Obsahuje vzorky textů z předních českých deníků, a to *Hospodářských novin* a *Lidových novin*, dále z regionálních deníků *Moravia* a *Bohemia* a z bezplatného deníku *Metro* s největším průměrným počtem distribuovaných výtisků v zemi, zejména z období 2000–2009. Kontrolní korpus čítá 15 692 278 slov (tokens), 519 603 různých slovních tvarů (types) a 217 860 lemmat. Je tedy o něco málo větší (o 341 537 slov) a obsahuje více lemmat (o 53 045), což ovšem z hlediska studie není podstatné.

Při vyhledávání kolokací v obou korpusech byl použit program Sketch Engine, korpusový vyhledávací systém Adama Kilgarriffa a Pavla Rychlého (vyvíjený od r. 2003 firmou Lexical Computing Ltd.), zahrnující tzv. word sketches, jednostránkové, automaticky počítačově generované souhrny gramatického a kolokačního chování slova v daném korpuse.

4. ANALÝZA KOLOKACÍ VYBRANÝCH NEUTRÁLNÍCH SLOV

4.1 SUBSTANTIVA A JEJICH KOLOKACE

Při analýze kolokačního chování čtyř zkoumaných substantiv, *život*, *člověk*, *organizace* a *boj*, je u každého substantiva nejprve vyhledáno prvních sedm kolokací s nejvyšší frekvencí v korpuse Totalita. Frekvence těchto sedmi kolokací je pak zjišťována v kontrolním korpuse. V další fázi je vyhledáno sedm nejčtenějších kolokací zkoumaného substantiva v kontrolním korpuse současného jazyka a frekvence těchto sedmi kolokací je zjišťována v korpuse Totalita. To umožňuje jednak porovnávat nejfrekventovanější kolokace v jednom i druhém korpuse, jednak porovnat distribuci kolokací, které jsou nejfrekventovanější v jednom období, s jejich distribucí v druhém ze sledovaných období. Od srovnání nejfrekventovanějších kolokací v korpuse Totalita s nejfrekventovanějšími kolokacemi v kontrolním korpuse lze očekávat, že naznačí, k jakým změnám priorit v diskurzu obou období došlo. Nejfrekventovanější adjektivní kolokace představují atributy nejčastěji spojované s daným substantivem v příslušném korpuse. Tyto atributy indikují nejdůležitější kontexty, v nichž se o tomto substantivu mluví. Dohromady vytváří těchto sedm kolokací určitý referenční rámec, profil substantiva v jednom i druhém časovém období. Interpretace rozdílů v distribuci vychází z předpokladu, že změna ve frekvenci výskytu kolokace v jednom období oproti druhému je indikátorem významu přikládaného dané kolokaci v diskurzu příslušného období. Je přirozené, že interpretace výskytů nemůže být exaktním procesem, jde spíše o hledání nápadných tendencí a kontrastů. Na druhé straně se závěry opírají o 112 frekvenčních údajů o kolokacích (56 nejfrekventovanějších kolokací a 56 hodnot jejich výskytu v druhém korpuse).

Výchozí korpus: Totalita			Výchozí: Kontrolní korpus				
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence						
společenský	785	→	51	celý	299	→	352
politický	693	→	27	lidský	164	→	344
kulturní	608	→	53	běžný	91	→	15
šťastný	517	→	11	každodenní	65	→	105
nový	451	→	55	vlastní	63	→	61
veřejný	419	→	49	osobní	61	→	66
hospodářský	354	→	7	nový	55	→	451

TABULKA 2. Frekvence výskytu kolokací substantiva život v cílovém a kontrolním korpusu

4.1.1 SUBSTANTIVUM ŽIVOT

Adjektivní kolokáty lemmatu život v korpusu Totalita je možné rozdělit do dvou skupin: na ty, které se týkají veřejné sféry, *společenský*, *politický*, *kulturní*, *veřejný*, *hospodářský*, a ty, které popisují kvalitu života, *šťastný* a *nový*. Fakt, že adjektiva kvality života se frekvenčně pohybují zhruba uprostřed, může značit, že mezi těmito kvalitami a veřejným životem společnosti existuje jistá (kauzální) souvislost. V kontrolním korpusu naproti tomu dominují adjektiva vztahující se k soukromému životu jedince: *celý*, *lidský*, *běžný*, *každodenní*, *vlastní*, *osobní*. Zároveň se tu objevuje jedno adjektivum společné oběma sedmičkám kolokací, *nový*. Lze spekulovat, že na rozdíl od nového společenského života tu jde o nový osobní život. Je ovšem třeba říci, že kolokace *kulturní* (53), *společenský* (51) a *veřejný* (49) se v kontrolním korpusu vyskytují jen těsně za sedmým nejfrekventovanějším adjektivem *nový*. Přesto je zde důraz na osobní život patrný. Nápadně nízký je v kontrolním korpusu výskyt tří ze sedmi nejčastějších adjektiv v korpusu Totalita: *politický*, *hospodářský* a zejména *šťastný*. Potvrzovalo by to určitou korelaci v diskurzu totality mezi šťastným životem a společensko-politickým systémem jako jeho zárukou.

Celkově se tedy zdá, že kolokace v obou korpusech vykazují dvě základní tendence: (kvantitativně) výraznou orientaci totalitního diskurzu na nadosobní život společnosti (v jeho různých podobách, především *společenské* a *politické*), od něhož se implicitně odvíjí příslib šťastného nového života (společnosti). Současný veřejný diskurz se naproti tomu primárně zaměřuje na každodenní soukromý život jedince. Zatímco kolokace v korpusu Totalita dobře odpovídají ideologickým představám o podřízenosti jedince a jeho soukromého života zájmům společnosti, kolokace v kontrolním korpusu vypovídají o prioritě soukromé sféry.

4.1.2 SUBSTANTIVUM ČLOVĚK

Mezi kolokacemi lemmatu člověk (*lidé*) v korpusu Totalita se příznačně objevují dvě ideologicky zatížená adjektiva, *sovětský* a *socialistický*. První z nich, *sovětský*, je dokonce nejfrekventovanějším kolokátem lemmatu člověk. Naproti tomu minimální výskyt

Výchozí korpus: Totalita			Výchozí: Kontrolní korpus				
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence		frekvence		frekvence		frekvence
sovětský	1450	→	3	mladý	782	→	1246
mladý	1246	→	782	starý	293	→	210
pracující	556	→	5	místní	126	→	10
nový	329	→	75	obyčejný	95	→	23
socialistický	319	→	1	jiný	82	→	76
starý	210	→	293	slušný	81	→	16
poctivý	203	→	12	nový	75	→	329

TABULKA 3. Frekvence výskytu kolokací substantiva *člověk* v cílovém a kontrolním korpusu

obou těchto adjektiv v kontrolním korpusu, tedy v současném diskurzu, příliš nepřekvapuje: *sovětský* (3) a *socialistický* (1). Zajímavější jsou ostatní ze sedmi nejčastějších adjektiv spojovaných s *člověkem* v korpusu Totalita. Ve svém úhrnu tyto kolokace nastiňují jakési paradigma „nového“ člověka: mladý a pracující (a poctivý). Jeho předobrazem je podle všeho sovětský člověk a s poměrně velkým odstupem socialistický člověk. Z tohoto obrazu vybočuje pouze adjektivum *starý*. V obou korpusech je nápadné zastoupení adjektiva *mladý*: v cílovém korpusu na druhém místě, v kontrolním korpusu je to dokonce nejčastější přídavné jméno. Ovšem zatímco v korpusu totality se s ohledem na ostatní kolokace zdá být spojení *mladý člověk/mladí lidé* součástí ideologického diskurzu, v kontrolním korpusu představuje spolu s ostatními kolokacemi spíše jen věkovou skupinu nebo sociální kategorii. Nasvědčuje tomu fakt, že druhým nejčastějším adjektivem je zde *starý*, které se sice vyskytuje mezi sedmičkou nejfrekventovanějších i v korpusu Totalita, ale až na předposledním místě (a má dokonce nižší absolutní frekvenci než v kontrolním korpusu). Tři další adjektiva v kontrolním korpusu nápadně kontrastují svou neutrálností s adjektivy korpusu Totalita: *místní*, *obyčejný*, *jiný*.

Zajímavé je i srovnání adjektiv morálních kvalit, která se v obou skupinách nejčastějších adjektiv objevila. Zatímco totalitní diskurz operuje pojmem *poctivý* (které má jen zanedbatelnou frekvenci v kontrolním korpusu, 12), současný diskurz užívá adjektivum *slušný* (které má pro změnu zanedbatelnou frekvenci v korpusu Totalita, 16). Ačkoli denotace těchto adjektiv mají k sobě velmi blízko, jejich asociace jsou zjevně odlišné. Lze se domnívat, že zatímco měřítkem poctivosti byl (primárně, často) vztah a oddanost k vlasti, režimu, straně, slušnost se chápe jako norma chování k jiným lidem. Z tohoto důvodu bude prototypické spojení *poctivý komunista*, nikoli *slušný komunista*. Přítomnost adjektiva *nový* by si zasloužila hlubší analýzu kontextů, v jakých je používáno v obou diskurzech. Za zmínku stojí, že v kontrolním korpusu se těsně za adjektivem *nový* (75) objevují adjektiva *postižený* (72) a *nemocný* (71), jejichž výskyt je v korpusu Totalita ve spojení s lemmatem *člověk* minimální (10, 21). To opět podporuje závěr, že zatímco totalitní diskurz probíhá spíše na rovině ideologické, propagační (tomu odpovídá zjevně generické spojení *sovětský/socialistický člověk*), na niž postižení a nemocní nenajdou příliš uplatnění, kolokace současného diskurzu se více vztahují ke konkrétním skupinám.

Výchozí korpus: Totalita			Výchozí: Kontrolní korpus				
Korpus / kolokace	Totalita	→	Kontrolní frekvence	Korpus / kolokace	Kontrolní	→	Totalita frekvence
	frekvence				frekvence		
stranický	2889	→	13	ziskový	137	→	0
základní	1570	→	40	místní	102	→	258
odborový	959	→	97	odborový	97	→	959
společenský	919	→	17	mezinárodní	96	→	283
závodní	594	→	1	obchodní	95	→	129
masový	451	→	0	příspěvkový	95	→	30
vesnický	377	→	0	ekologický	66	→	0

TABULKA 4. Frekvence výskytu kolokací substantiva *organizace* v cílovém a kontrolním korpusu

4.1.3 SUBSTANTIVUM ORGANIZACE

Celkově se ukazuje, že z hlediska frekvence je slovo *organizace* ze všech čtyř zkoumaných slov v totalitním diskurzu nejužívanější. Naopak v současném diskurzu je hned po slově *boj* nejméně frekventovaným lemmatem. S tím souvisí i značný nepoměr ve výskytu slova *organizace* v obou korpusech (stačí srovnat frekvence prvních kolokací v pořadí: 2889 a 137). To má samo o sobě důležitou výpovědní hodnotu. Kolokace lemmatu *organizace* v korpusu Totalita odkazují bez výjimky na politicky, resp. ideologicky profilovanou organizaci. Dvě nejčastější adjektiva, *stranický* a *základní*, se týkají přímo KSČ. Třetím nejčastějším spojením je *odborová organizace*, ovšem v situaci, kdy všechny podniky jsou státní, představují odbory stranicky řízenou strukturu. Zbýlé kolokace vymezují působnost organizace (celospolečenská a lokální, *masová*, *závodní*, *vesnická*). V kontrolním korpusu frekvence, a tedy důležitost těchto adjektiv (s výjimkou *odborový* a *základní*) markantně klesá. Okruh nejčastějších kolokací zde svědčí o výrazné depolitizaci slova *organizace*. Kolokáty *ziskový*, *obchodní*, *příspěvkový* jsou z oblasti profesní, ekonomické, a mění se i vymezení působnosti (*mezinárodní*, *místní*). Zejména adjektivum *mezinárodní* implikuje, že daná organizace představuje (nestranický) subjekt mezinárodního práva. Zajímavá je shodně třetí pozice adjektiva *odborový* v obou korpusech, která však v případě současného korpusu svědčí o faktické důležitosti odborů jakožto čistě profesních sdružení hájících zájmy zaměstnanců (nikoli státu). Adjektivum *ekologický* je pro změnu příznačné jenom pro současný diskurz (přestože slovo jako takové existuje už dávno), ovšem jde o jiný druh politické angažovanosti než u adjektiva *stranický*. Kolokace s nejvyšší frekvencí v korpusu Totalita mají, jak bylo řečeno, v kontrolním korpusu velmi nízkou frekvencí, s výjimkou zmíněného adjektiva *odborový*. Dvě z adjektiv, *masový* a *vesnický*, se v kontrolním korpusu nevyskytla vůbec. Naopak dvě z těchto sedmi nejčastějších kolokací, *zisková* a *ekologická organizace*, zcela absentují v korpusu Totality. Ziskovost a ekologické aktivity v popředí tehdejšího veřejného zájmu rozhodně nestály, podobně jako *masové* a *vesnické organizace* nefigurují v současném diskurzu.

Obecně řečeno tedy srovnání kolokací ukazuje, že v současném korpusu ztrácí slovo *organizace* asociaci „stranické angažovanosti“, kterou v převážné většině mělo

v korpusu Totalita, a místo toho asociuje ekonomickou, zájmovou nebo účelovou jednotku, útvar.

4.1.4 SUBSTANTIVUM BOJ

Podobný obraz jako *organizace* skýtá srovnání kolokací lemmatu *boj*. V korpusu Totalita má pět nejčastějších adjektiv explicitně ideologický charakter, *třídní, osvobozenecký, politický, revoluční, ideologický*. Dvě zbývající adjektiva jsou v podstatě rovněž ideologicky motivované atributy presuponující kolektivnost a oprávněnost politického boje (*společný a hrdinný*). Slovo *boj* se v totalitním diskurzu stalo nedílnou součástí politického slovníku a bylo téměř nemožné si ho představit v jiném kontextu. Přesto v kontrolním korpusu dochází k radikálnímu přehodnocení. Především celková frekvence slova *boj* je nápadně nižší (viz *třídní* 1128, *vyřazovací* 77). *Boj* přestal být programovým cílem, a tak lemma *boj* ztrácí z větší části své ideologické asociace: všechny kolokace jsou z oblasti obchodní (*konkurenční*), volební (*předvolební*; jde tu o politickou, ale nikoli ideologickou záležitost) a sportovní (*vyřazovací*). Z těchto oblastí jsou i zbývající přívlastky, *tvrdý, urputný, tuhý, i jarní* (týká se sportovních soutěží konaných na jaře). Budiž ale řečeno, že o sedmé místo se v kontrolním korpusu s adjektivem *tuhý* dělí i adjektivum *politický*. I přesto můžeme konstatovat, že stejně jako u slova *organizace*, zde dochází k deideologizaci významu slova *boj*, tzn. ke změně jeho asociační složky.

Výchozí korpus: Totalita			Výchozí: Kontrolní korpus				
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence		frekvence				frekvence
třídní	1128	→	11	vyřazovací	77	→	1
osvobozenecký	390	→	2	konkurenční	53	→	52
politický	318	→	23	tvrdý	36	→	68
revoluční	263	→	3	urputný	33	→	69
ideologický	260	→	2	předvolební	29	→	24
společný	245	→	4	jarní	30	→	1
hrdinný	152	→	0	tuhý	23	→	31

TABULKA 5. Frekvence výskytu kolokací substantiva *boj* v cílovém a kontrolním korpusu

Prakticky všech sedm nejfrekventovanějších kolokací v korpusu Totalita vykazuje mimořádně nízkou frekvenci v kontrolním korpusu, spojení *hrdinný boj* se v něm dokonce nevyskytuje vůbec. Podobně dvě ze sedmi nejfrekventovanější kolokací v kontrolním korpusu, *vyřazovací boj* a *jarní boj*, se v korpusu Totalita vyskytly pouze jednou. Kontrast mezi ideologickou rétorikou korpusu Totalita na straně jedné a enormním zájmem o sport, tedy o zábavu, obchodní a volební aktivity v současném korpusu na straně druhé (*vyřazovací a jarní boj* je zde boj o postup do ligy, o mistrovský titul apod.) je velmi markantní.

Souhrnně můžeme říci, že srovnání kolokací substantiv *život, člověk, organizace* a *boj* ukázalo značné odlišnosti ve složení sedmi nejčastějších v korpusu Totalita a v kont-

rolním, současném korpusu (a jejich postavení v druhém korpusu). Pouze ve čtyřech, resp. pěti případech se u téhož substantiva vyskytlo mezi sedmi nejčastějšími kolokacemi stejné adjektivum v obou korpusech: *kulturní život*, *nový život*, *nový člověk*, *odbořovací organizace* (a *politický boj*). Všechny odlišnosti přitom mají společného jmenovatele. V korpusu Totality se nejčtenější kolokace zvolených substantiv z větší části jeví jako (přímo či nepřímě) ideologicky motivované: život společnosti je upřednostňován před životem konkrétního jedince, prototyp člověka nové společnosti (*sovětský, mladý, pracující*) je upřednostňován před konkrétním člověkem (*mladý, starý, místní*) a slova *organizace* a *boj* jsou z pohledu kolokací pojímána jako ideologické a mocenské nástroje. Lze se domnívat, že ideologická asociace není jediná složka významu, kterou se liší užívání těchto slov v korpusu Totalita od jejich užití v kontrolním korpusu. Tato složka asociativní je totiž často vázána na složku evaluativní. Ve své době nejobvyklejší spojení *straničká organizace* a *třídní boj* se vžila v tehdejší rétorice do té míry, že představují prakticky kvazi terminologická sousloví. Původně neutrální slova *organizace* a *boj* získala za desetiletí užívání v oficiálním diskurzu vágně pozitivní rys něčeho společensky potřebného a sankcionovaného. V následujícím období odlišných priorit se zpátky vrátily evaluativně nezátížené kolokace a tento rys se vytratil.

4.2 ADJEKTIVA A JEJICH KOLOKACE

Rozbor kolokačního chování tří zkoumaných adjektiv, *nový*, *vysoký* a *americký*, je pojat obdobně jako u substantiv. Bylo vyhledáno sedm nejčtenějších kolokací v cílovém korpusu totalita a jejich distribuce v kontrolním korpusu. Potom bylo zaznamenáno sedm nejčtenějších kolokací v kontrolním korpusu a jejich distribuce v cílovém korpusu Totalita. Výsledný obraz je následující.

4.2.1 ADJEKTIVUM NOVÝ

Při pohledu na kolokace adjektiva *nový* upoutá na první pohled rozdíl mezi nejfrekventovanější kolokací v korpusu Totalita, *nová válka*, a nejfrekventovanější kolokací v kontrolním korpusu, *nový rok*. Protože není pravděpodobné, že by se počet nových válek

Výchozí korpus: Totalita				Výchozí: Kontrolní korpus			
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence		frekvence		frekvence		frekvence
válka	895	→	3	rok	754	→	333
stanovy	595	→	3	zákon	615	→	214
metoda	570	→	59	starosta	270	→	1
forma	471	→	45	vedení	251	→	279
život	451	→	55	system	250	→	233
technika	428	→	32	vláda	242	→	370
úspěch	378	→	2	majitel	194	→	8

TABULKA 6. Frekvence výskytu kolokací adjektiva *nový* v cílovém a kontrolním korpusu

v obou obdobích výrazně lišil, lze frekvenci této kolokace v korpusu Totalita přičíst na vrub politickému aspektu diskurzu (účelové zdůrazňování trvalého napětí a ohrožení). Také poslední ze sedmi nečastějších kolokací v korpusu Totalita, *nový úspěch*, lze interpretovat jako součást politického diskurzu (účelové a neustálé zdůrazňování úspěchů režimu, s tím možná souvisí i četný výskyt kolokace *nový život*, viz výše). Naproti tomu v kontrolním korpusu nacházíme na sedmém místě spojení *nový majitel*, signalizující zvýšený zájem o soukromé vlastnictví. O diametrálním rozdílu v zaměření obou diskurzů svědčí fakt, že *válka* a *úspěch* se v kontrolním korpusu prakticky neobjevily (četnost 3 a 2), *majitel* se v korpusu Totalita s adjektivem *nový* vyskytl jen 8×. V kontrolním korpusu lze na základě frekvence pozorovat, že dochází častěji k obměně legislativních a exekutivních prvků (*zákon*, *starosta*, *vedení*, *vláda*), u kterých jsou v diskurzu totality změny méně časté. V korpusu totality se změny týkají spíše realizačních a méně významných jevů (*stanovy*, *metoda*, *forma*, *technika*).

4.2.2 ADJEKTIVUM VYSOKÝ

Také kolokace adjektiva *vysoký* vypovídají do značné míry o změněné situaci v obou diskurzích, a to bez ohledu na společnou nejčastější kolokaci *vysoká škola*. V korpusu Totalita jsou příznačné pro tehdejší diskurz kolokace *nejvyšší orgán* (stranický), *vysoká pec* (těžký průmysl), *nejvyšší sovět* (a do značné míry i *vysoká efektivnost*, která prakticky absenteje v kontrolním korpusu). Současný diskurz charakterizuje sportovní terminologie (*nejvyšší soutěž*), nezávislé soudnictví (*nejvyšší soud*), ekonomika (*vysoké ceny*) a nový politický styl (*nejvyšší představitel*). V obou korpusech se shodně mezi nejčastějšími sedmi kolokacemi objevuje *vysoká úroveň*, i když v korpusu Totalita hned na druhém místě (v kontrolním korpusu je to *nejvyšší soutěž*), kdežto v kontrolním korpusu až na pátém.

Výchozí korpus: Totalita			Výchozí: Kontrolní korpus				
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence		frekvence		frekvence		frekvence
škola	1570	→	1060	škola	1060	→	1570
úroveň	966	→	277	soutěž	691	→	216
orgán	541	→	24	soud	434	→	190
pec	521	→	29	cena	304	→	125
stupeň	512	→	104	úroveň	277	→	966
sovět	400	→	1	představitel	171	→	125
efektivnost	242	→	3	počet	155	→	60

TABULKA 7. Frekvence výskytu kolokací adjektiva *vysoký* v cílovém a kontrolním korpusu

4.2.3 ADJEKTIVUM AMERICKÝ

Možná nejzajímavější je srovnání kolokací adjektiva *americký*. Odhlédneme-li od kolokací *Spojené státy americké* a *americká vláda*, společných oběma korpusům, ozna-

Výchozí korpus: Totalita				Výchozí: Kontrolní korpus			
Korpus / kolokace	Totalita	→	Kontrolní	Korpus / kolokace	Kontrolní	→	Totalita
	frekvence		frekvence		frekvence		frekvence
imperialista	1311	→	0	president	392	→	178
imperialismus	631	→	2	dolar	287	→	43
stát	412	→	159	společnost	171	→	134
monopol	285	→	0	stát	159	→	412
agresor	248	→	0	trh	131	→	9
letadlo	239	→	42	firma	123	→	45
vláda	232	→	116	vláda	116	→	232

TABULKA 8. Frekvence výskytu kolokací adjektiva *americký* v cílovém a kontrolním korpusu

čují všechny ostatní kolokace v korpusu Totalita jevy společensky nepřijatelné a potenciálně nebezpečné, ať už přímo, *americký imperialista*, *imperialismus*, *agresor*, nebo nepřímo (ekonomická a vojenská hrozba), *monopol*, *letadlo* a *vojsko* (z hlediska frekvence, 231, je sice *vojsko* až na osmém místě, ale od lemmatu *vláda* s frekvencí 232 ho dělí jediný výskyt). V kontrolním korpusu tato tendence zcela chybí. Vedle kolokací odkazující na státnost (*prezident*, *státy*, *vláda*, *společnost*) jsou tu neutrální kolokace z ekonomické oblasti, *dolar*, *trh* a *firma*. Pohled na frekvence negativních kolokací v kontrolním korpusu (*imperialista* — 0, *imperialismus* — 2, *agresor* — 0, *monopol* — 0, ale i *vojsko* — 5) prozrazuje, že zde nejsou prakticky vůbec zastoupeny. Lze říci, že způsob užívání adjektiva *americký* a četnost těchto frekvencí vedla v oficiálním diskurzu totality ke vzniku specifické sémantické prozodie tohoto adjektiva, a to v míře vyšší než u všech ostatních zkoumaných slov. Sémantická prozodie je sémanticko-pragmatický termín korpusové lingvistiky, který definoval Louw (1993, s. 157) jako „consistent aura of meaning with which a form is imbued by its collocates“. Míni se tím, že slova, kolokáty, které se s daným lemmatem často a pravidelně vyskytují a tvoří s ním širší lexikální jednotku, předávají lemmatu část svého významu. V daném případě *americký* (a jakékoli spojení s ním) začalo automaticky implikovat „špatný, odsouzeníhodný, nepřátelský“, získalo negativní sémantickou prozodii.

Analýza kolokací vybraných neutrálních slov ukázala, že co se týče okruhu nejčetnějších kolokátů zkoumaných slov, vykazují oba korpusy značné rozdíly. Z kvantitativního hlediska z celkového počtu 98 nejfrekventovanějších kolokací sedmi zkoumaných slov v obou korpusech byla shoda mezi korpusy pouze v 8, resp. 9 případech (4×, resp. 5× u substantiv a 4× u adjektiv). Ovšem hlavní rozdíl je v celkovém vyznění kolokací v korpusu Totalita a kontrolním korpusu. Zatímco za kolokacemi totalitního diskurzu je cítit konzistentní ideologický tlak, který určuje priority a hodnotící kritéria do té míry, že to ovlivňuje sémanticko-pragmatické složky běžně bezpříznakových slov, na kolokacích kontrolního korpusu nic nápadného na první pohled zřejmé není. Zdálo by se, že substantiva, která jsou nositeli pojmů, budou relativně snadněji využitelná k záměrné ideologizaci obsahu než adjektiva, ale příklad adjektiva *americký* ukazuje, že se tento proces nevyhne ani jim. Je to dáno tím, že „americký“ je na

rozdíl od ostatních adjektivem relačním, tzn. vyjadřuje primárně vztah k základovému substantivu „Amerika“. Vazba tohoto adjektiva na substantivum se pak promítá i do jeho kolokačního chování. Přestože lze namítnout, že rozbor a interpretace vycházejí pouze z kolokací a nepřihlíží až na výjimky k širšímu kontextu, rozdíly mezi totalitním a současným diskurzem signalizované nefrekventovanějšími kolokacemi (které nepochybně odrážejí odlišnou realitu) jsou těžko přehlédnutelné a odpovídají očekáváním.

5. DOBOVĚ PŘÍZNAČNÁ SLOVA

Pokud by se předchozí interpretace kolokací neutrálních slov jevila jako příliš zaujatá a tendenční, stačí si přiblížit totalitní diskurz na dobově příznačných slovech. Míra ideologizace diskurzu, která se může do jazyka promítnout, a s ní spojené idolatrie a fanatismus, je dnes pro generace, které dospívaly po r. 1989, možná těžko představitelná. K dokreslení obrazu historicky nedávného totalitního diskurzu stojí proto za připomenutí alespoň několik málo dobově příznačných výrazů a jejich spojení, které se sice v Základním slovníku podle frekvence obsaženém ve Slovníku komunistické totality neobjevují mezi prvními sty slovy, ale jsou velmi výmluvné a pro komunistické období typické. Vybrána byla čtyři slova, adjektiva *rodný* a *prohnilý* a substantiva *láska* a *šťěstí*, která jsou zajímavá tím, jak u nich totalitní úzus zasáhl do významové struktury.

U adjektiva *rodný* uvádí Příruční slovník jazyka českého (Díl IV, Část 2., 1944–48) tyto významy, (a) rodem, pokrevností, blízký; vlastní (*rodný otec, bratr*); (b) určující prostředí, kde jsme se narodili (*rodný dům, země*); (c) ve spojení *rodné jméno*, tzn. příjmení; (d) v rodě užívaný, obvyklý (*rodné slovo, písničky*). Další dva významy, rodový a rodící, označuje za zastaralé. Sémantická prozodie adjektiva *rodný* je zde výrazně kladná. Ve Slovníku spisovného jazyka českého (1960–71) nacházíme nový význam (s poznámkou (podle rus.)): expr. patříci pevně něj. společenství; milý 1, milovaný, drahý 3; r-á dělnická třída; r-á strana (o komunistické straně). To, k čemu zde došlo, lze hodnotit jako zneužití sémantické prozodie tohoto adjektiva k propagandistickým účelům. V korpusu Totality nacházíme mj. tyto příklady: *To bude můj dar k historickému sjezdu naší rodné bolševické strany. Každý z pracovníků brigády je hrdý na to, že její vedoucí byl delegátem XIX. sjezdu, že viděl a slyšel hovořit rodného Stalina.*

Příkladem na zneužití záporné sémantické prozodie je slovo *prohnilý*. Oba zmínované slovníky českého jazyka shodně uvádějí jako první význam „silně proniklý hnilobou; shnilý“ a jako druhý „zcela mravně porušený; zkažený“. V Korpusu totality nicméně frekvence kolokací s druhým významem je tak vysoká a frekvence kolokací s prvním významem tak nízká, že metaforický význam téměř vytlačil význam doslovný. Jinými slovy kolokace typu *prohnilý kapitalismus, nacionalismus, prohnilá buržoasní morálka* se na dlouhou dobu staly normou a běžným územ.

Také substantivum *láska* s výraznou emocionální složkou a pozitivní sémantickou prozodií bylo hojně využíváno, resp. nadužíváno k demonstrativnímu projevu loajality. Zatímco v kontrolním korpusu má slovo *láska* nízké zastoupení v publicistice,

vyskytuje se především v beletrii, v korpusu Totalita nacházíme četné výskyty jako: *Všichni, kteří vstupují do těchto místností, spojuje jediná láska, láska k soudruhu Gottwaldovi. Učitel má pěstovat v dětech city nejhlubší lásky, vděčnosti a oddanosti k Sovětskému svazu. Vyjadřující svou nezměrnou lásku a oddanost straně Lenina — Stalina.*

K velmi podobné situaci došlo u slova štěstí. Jeho první a hlavní význam definuje SSJČ jako „stav n. pocit blaha, úplného uspokojení; blaženost, blaho: lidské š.; mateřské, rodinné š.“, tedy jako štěstí osobní. V totalitním diskurzu však došlo k charakteristickému posunu, který byl zmíněn u kolokací slov život a člověk. Místo štěstí jedince převažují kolokace s generickým významem, jde o štěstí národa, lidstva, štěstí, které přijde v budoucnosti a je spojované s ideologií (podobně jako u náboženství). Příslib štěstí v korpusu Totality je častým argumentem propagandy při obhajování politiky a prosazování cílů: *Štěstí sovětského lidu je i naším štěstím. Marxisticko-leninské učení není jen teorie, je to návod, jak v praxi vybudovat socialismus a komunismus, blahobyť a štěstí lidstva.*

6. ZÁVĚRY

Srovnání kolokací totalitního jazyka s jazykem současným ukázalo masivní vliv tehdejší oficiální ideologie. Ze 49 nejfrekventovanějších kolokací sedmi zkoumaných slov jich lze dvě třetiny označit jako motivované více či méně otevřenou politickou propagací. Je třeba říci, že zkoumaný vzorek přitom obsahuje slova, která jsou z dnešního pohledu neutrální — přesto dvě třetiny jejich kolokací v korpusu Totalita neutrální nejsou. Korpusově založený výzkum kolokací navazuje na přístup J. R. Firthe (1957), který koncept kolokace rozpracoval a kterýrazil myšlenku, že význam a způsob užívání slova mohou být charakterizovány jeho nejtypičtějším kolokáty. Shrnul ji do známého dikta *Slovo poznáš po společnosti, v níž setrvává*, ale zároveň si byl vědom, že jde o určitou aproximaci na čistě lexikální rovině. Také v této studii šlo o přibližné nastínění toho, jak se systematické a cílené užívání nebo dokonce zneužívání jazyka může projevit na lexikální rovině. Z průzkumu vyplývá, že dlouhodobé užívání ideologicky motivovaných spojení skutečně může výrazně ovlivnit asociativní složku významu slov, v některých případech může ovlivnit sémantickou prozodii těchto slov nebo jejich prozodii může využít a parazitovat na ní. Studie tak na malém vzorku dokládá, že při ohraničování specifického slovníku totality se nelze omezit na ideologicky příznakovou skupinu lexikálních jednotek, ale je nutno přihlížet i k syntagmatickým aspektům totalitního diskurzu, které podstatně rozšiřují okruh toho, co je pro slovník totality specifické. Tematika složitých vztahů mezi jazykem a ideologií v širším slova smyslu je a bude velmi aktuální (srov. Barát — Studer — Nekvapil, 2013), a proto by si totalitní jazyk naší nedávné minulosti zasloužil systematictější a důkladnější výzkum, než jakého se mu zatím dostalo.

LITERATURA

- ČERMÁK, F. — CVRČEK, V. — SCHMIEDTOVÁ, V. (eds.) (2010): *Slovník komunistické totality*. Praha: Lidové noviny/Ústav českého národního korpusu.
- BAKER, M. — GILL, F. — TOGNINI-BONELLI, E. (eds.) (1993): *Text and Technology*. In Honour of John Sinclair. Philadelphia/Amsterdam: John Benjamins.
- BARÁT, E. — STUDER, P. — NEKVAPIL, J. (eds.) (2013): *Ideological Conceptualizations of Language: Discourses of Linguistic Diversity*. New York: Peter Lang Publishing.
- FIRTH, J. R. (1957): A synopsis of linguistic theory 1930–55. In: *Studies in linguistic analysis*, s. 1–32. Oxford: The Philological Society. Přetisk In: F. R. PALMER (1968), s. 168–205.
- HAVRÁNEK, B. a kol. (1960–1971): *Slovník spisovného jazyka českého*. Praha: ČSAV.
- KILGARRIFF, A. — RYCHLÝ, P. — SMRŽ, P. — TUGWELL, D. (2004): The Sketch Engine, EURELEX 2004, Lorient, France, s. 105–116.
- LOUW, W.E. (1993): Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In: M. BAKER et al. (1993), s. 157–176.
- PALMER, F. R. (ed.) (1968): *Selected Papers of J. R. Firth 1952–59*. London: Longmans.
- Příruční slovník jazyka českého (1935–1957)*: Praha: ČSAV, SN.
- RYCHLÝ, P. (2007): Manatee/Bonito — A Modular Corpus Manager. In: P. SOJKA — A. HORÁK (eds.), *First Workshop on Recent Advances in Slavonic Natural Language Processing*, RASLAN 2007. s. 65–70. Brno: Masaryk University. <http://nlp.fi.muni.cz/raslan/2007/papers/12.pdf>
- SCHMIEDTOVÁ, V. (2006): What did the totalitarian language in the former socialist Czechoslovakia look like? In: *The 1st Annual Conference of The Slavic Linguistics Society*. Indiana University. (Bloomington Indiana), September 2006. <http://www.indiana.edu/~sls2006/page2/page2.htm>
- SCHMIEDTOVÁ, V. (2007): Totalitní jazyk v bývalém Československu. Koncept slova práce. In: *Totalitarismus 3*, sborník z konference, Katedra antropologie, FF ZU, Plzeň, s. 110–116.
- SCHMIEDTOVÁ, V. (2008): Hodnotící prostředky v totalitním jazyce 1948–1989 v bývalém Československu. In: *Totalitarismus 4*, sborník z konference, Katedra antropologie FF ZU. Plzeň, s. 186–196.
- SCHMIEDTOVÁ, V. (2011): Die Sprache der Propaganda in der Tschechoslowakei 1948–1989. In: *Brücken. Germanistisches Jahrbuch Tschechien-Slowakei. Neue Folge 19/1–2 (2011)*. Praha: Nakladatelství Lidové noviny.
- Sketch Engine*, Lexical Computing Ltd, version SkE-2.39-2.78