

Slovo *to* v mluvených korpusech ČNK, jeho prefixace a reduplikace

Petra Klimešová, Zuzana Komrsková,
Marie Kopřivová, David Lukeš (Praha)

THE CZECH WORD 'TO' IN THE SPOKEN CORPORA OF THE CNC: PREFIXATION AND REDUPLICATION

The word form *to* (that.SG.N) traditionally tops frequency tables in corpora of spoken Czech: as a universal (gender- and number-neutral) exophoric (deictic) and endophoric (co-referential) device, it is crucial for spontaneous, unplanned discourse which requires reinforcing references to the context and co-text. Our estimate based on the ORAL series corpora is that about 70% of the instances of *to* in informal spoken language preserve strong referential semantics (i.e. the endo/exophoric function is preserved). In the remaining cases these functions are attenuated to different degrees: about 20% are part of what a phraseologically-oriented account would consider as indivisible chunks (often multi-word units expanding on the stub *to + je.be.3SG*), whereas the remaining 10% are purely pragmatic (chiefly serving as a turn-taking/keeping device). As the data show, *to* can even be prefixed, acting as a surrogate verb, and it is often reduplicated; both strategies indicate an attempt not to yield one's turn while searching for words. A comparison of their relative frequencies in spoken and written corpora reveals these constructions as characteristic of spoken language.

KEYWORDS

spoken language, Czech, that, reduplication, prefixation

KLÍČOVÁ SLOVA

mluvený jazyk, čeština, *to*, reduplikace, prefixace

1. ÚVOD

Specifika mluveného jazyka vycházejí z podstaty jeho produkce a percepce, mluvčí nemá možnost se vracet a korigovat již řečené před tím, než se sdělení dostane k adresátovi. Zejména pro spontánní projev je typická nepřipravenost, která má za následek větší množství dysfluencí a častější užívání prostředků (výplňková či zástupná slova atp.), jimiž mluvčí získává čas, aby sdělení pečlivěji promyslel, případně si vybavil vhodné slovo, vzpomněl si na nějaký fakt a přitom si udržel mluvní aktivitu. V zájmu situačního ukotvení a propojení průběžně budovaného celku, „lokalizace a identifikace osob, objektů, událostí, procesů a aktivit, o nichž se mluví nebo k nimž se referuje“, jsou častěji užívána deiktická slova (Lyons, 1977, citováno dle Hirschová, 2013).

Ve svém příspěvku se zaměřujeme na slovní tvar *to*, který ve spontánním mluveném projevu plní zmíněné funkce. Představíme nově navrženou, zatím hrubou klasifikaci těchto funkcí a blíže se budeme věnovat dvěma případům užívání slova *to*



v mluveném jazyce — prefixaci a reduplikaci. Jedná se o jevy, jež se pro mluvený jazyk v porovnání s jejich výskytem v psaných datech zdají typické. Oba jevy jsou navíc poměrně jednoduše dohledatelné na základě formálních kritérií a nejsou frekventované natolik, aby jim nemohla být věnována zevrubnější pozornost.

2. VÝRAZ TO V ODBORNÉ LITERATUŘE

Z hlediska slovnědruhového rozlišení bývá výraz *to* tradičně řazen k zájmenům (jako rodová varianta zájmena *ten*) anebo k částicím. Tomuto dvojímu rozdělení odpovídá i morfologické značkování psaných korpusů ČNK.¹

Částice *to* je v mluvnících často klasifikována jako citová, a to hlavně v mluveném jazyce (viz např. Havránek–Jedlička, 1960, s. 381: *to dnes prší*; Petr et al., 1986, s. 88). Schopnost vyjadřovat postoj mluvčího ke sdělované skutečnosti vyplývá z původní citoslovečné platnosti výrazu *to* (Havránek–Jedlička, 1960, s. 272). Podle Machka (1957, s. 525, heslo *ten*) se jedná o „zesilovací slabiku složeného tvaru *tento, toto* apod. ve funkci upozornovací [*tedy citoslovečné, pozn. autorů*], hle“. Možné zařazení k citoslovcím uvádí také SSJČ (heslo *to*), v citovaných příkladech se však jedná pouze o případy moravského oblastního užití s významem upozornění, např. *jsem trochu v nedbalkách, to, ani obuta nejsem*. Výraz *to* dále SSJČ klasifikuje jako zájmenné příslovce odkazující k určitému okamžiku, např. *k obědu se přičenete pozdě a to už obchází parta za plotem*.

Kromě tradičního slovnědruhového rozřazení lze užívání výrazu *to* klasifikovat podle kontextu a významu, kterého v něm nabývá: může odkazovat k jakémukoli slovu, frázi, propozici či celému textu, a proto jej chápeme jako univerzální odkazovací prostředek (viz Vychodilová, 2008; Čermák, 2010). Mohli bychom ho zařadit do skupiny tzv. krajních generalizátorů, vedle slov typu *skutečnost, fakt* ad., „která se naplňují různými obsahy v závislosti na kontextu. [...] Tyto krajní generalizátory nabývají funkce prostředků odkazových (referenčních) [...]“ (Jelínek, 1995, s. 763–764).

Funkčností výrazu *to* se detailněji zabývali např. Adamec (1998), Vychodilová (2008) a Kolářová (2014).² Shodují se v případech, kdy má slovo *to* funkci deiktickou, resp. anaforickou a odkazuje k různě hierarchizovaným strukturám jazykovým nebo ven z textu. Naopak odlišná je širší kategorie částic a kategorie zájmenných adverbíí. Např. věty typu *To leje!* Vychodilová vymezuje jako věty s částicí, zatímco Adamec je uvádí jako případy plnicí funkce adverbialní/okolnostní, kde se výraz *to* blíží emocionální částici, event. citoslovcí. Kolářová chápe výraz *to* jako emocionální částici ve funkci navazovací. Nicméně všichni vycházejí z tradičního dělení na zájmena a částice.

O kategorizaci navrhané Kolářovou a částečně i Vychodilovou platí, že je založená především na psaném jazyce, doklady z mluveného jazyka jsou u obou spíše sporadické. Adamec zdroje svých příkladů neuvádí. Naším výchozím materiálem je mluvený jazyk neformálních, spontánních konverzací a při klasifikaci nalezených případů se ukázalo, že případné podrobnější rozčlenění vyžaduje jiná kritéria, než užívají zmiňovaní

1 V korpusu SYN2010 je *to* jako zájmeno označkováno v 969 495 případech, jako částice v 3243 případech (z toho v 56 %, tj. 1813 případech, se vyskytuje na začátku věty).

2 Na tomto místě bychom rádi poděkovali PhDr. Ivaně Kolářové za laskavé zapůjčení rukopisu.

autoři, proto jsme se na základě materiálu pokusili vytvořit kategorizaci vlastní. Při procházení dokladů jsme narazili na dva jevy, které se zdají být pro mluvený jazyk charakteristické: prefixaci a reduplikaci. Těm jsme se věnovali podrobněji.

O prefixaci a reduplikaci výrazu *to* se v odborné literatuře nacházejí spíše kusé poznámky nebo jsou tyto jevy nahlíženy z jiné perspektivy — např. reduplikace bývá chápána především jako opakování záměrné, jímž mluvčí dosahuje určitého komunikačního cíle: např. intenzifikace (*typ je moc moc velký*, srov. Lakoff–Johnson, 2002, s. 143) nebo naopak signalizace nedůležitosti informace (*typ bylo to tam a tam*, srov. Adam, 2001–2002, s. 72).

3. DATA A METODOLOGIE

Data jsme vyhledávali jednak v korpusu mluvené češtiny ORAL, v zatím neveřejné pracovní verzi korpusu spojujícího data z korpusů ORAL2006, ORAL2008 a ORAL2013, jednak v korpusech psaného jazyka SYN2010 a SYN. Vzhledem k zaměření článku jsme primárně vycházeli z mluvených dat, korpus psaného jazyka byl použit ke vzájemnému srovnání vybraných jevů.

Pro získání představ o rozložení jednotlivých funkcí výrazu *to* v mluveném jazyce prošli čtyři anotátoři vybraný náhodný vzorek a pokusili se o kategorizaci podle dosavadní odborné literatury.³ Ta ovšem vychází především z psaného jazyka, proto byla vytvořena vlastní pracovní klasifikace, ve které byly zvláště vyčleněny jevy nacházející se v mluveném jazyce (viz 4.1.3). K této skupině patří i dva podrobněji zkoumané jevy: prefixace a reduplikace.

K tématu prefixace byla data v korpusu ORAL vyhledávána pomocí dotazu [word="{1,4}to"] a z nalezených tvarů byly vybrány takové, které představovaly prefigovaný výraz *to*. Z výsledných dat byla ručně vyfiltrována slova homonymní s prefigovaným zájmenem *to*, např. spojka *proto* nebo příslovce *přítom* přepsané dle realizované výslovnosti jako *přito*. Položky z takto vytvořeného seznamu byly následně vyhledány v korpusu SYN, přičemž výsledky byly také ručně filtrovány. Výjimkou jsou výrazy *zato* a *nato*, pro jejichž vyhledání byla použita kombinace konkrétního tokenu a tagu (X@.*). V tab. 1 u nich není zaznamenána žádná frekvence výskytu, což ovšem neznamená, že je jejich existence v korpusu vyloučena. Na rozdíl od ostatních prefigovaných výrazů jsme totiž neprošli dané výrazy ručně z důvodu vysoké frekvence (výrazy jsou homonymní s příslovcem *nato* a výrazem *zato*).

Jako materiál pro téma reduplikace pak byly v korpusu ORAL vyhledávány případy dvou bezprostředně vedle sebe stojících výrazů *to*, které se vyskytují v replice jednoho mluvčího.⁴ Abychom odstranili případy, kdy se před hledaným reduplikovaným spojením *to to* vyskytl další výraz *to*, použili jsme negativní filtr (odfiltrovali jsme další výskyty *to* nacházející se na pozicích před a za KWIC, tj. použili jsme negativní filtr v rozsahu -2 až +2 bez zahrnutí KWIC).

Pro srovnání byly dále hledány i výskyty *ten ten* a *ta ta*.

3 Vzhledem k široké problematice společné anotace zde neuvádíme podrobnosti.

4 Zjištěno pomocí dotazů [word="{(?)to"}{2} within <s/> v SYN2010 a [word="{(?)to"}{2} within <sp/> v korpusu ORAL.

4. VÝSLEDKY

4.1 VLASTNÍ KATEGORIZACE SLOVA *TO*

Vzhledem k nedostatečné teoretické opoře pro klasifikaci slova *to* v mluveném jazyce a zároveň dostupnosti poměrně velkého množství mluvených dat jsme se pokusili o vlastní kategorizaci založenou na datech z mluvených korpusů řady ORAL. Jde o pracovní aproximaci, jejímž cílem bylo zvolit kritéria, která budou jednoduše a systematicky aplikovatelná (i na často nejednoznačných mluvených datech) a zároveň nepovedou ke vzniku spoře doložených kategorií.⁵ Jsme si však vědomi toho, že jednotlivé kategorie nemusí být pevně vymezeny, naopak jejich hranice budou velmi často prostupné a nejasné, na což upozornil již Adamec (1998).

4.1.1 KATEGORIE „ODKAZOVÁNÍ“

První skupina zahrnuje všechny výskyty výrazu *to* s deiktickou a/nebo koreferenční platností, např. *a tak sme to Milušce podali přes braňku* (ORAL2008); *vykřikovala, že pojedě eště to odpoledne domů* (ORAL2008). Do této skupiny řadíme i příklady, v nichž je *to* použito v adverbialním/cirkumstantovém významu (Adamec, 1998), např. *to už pak bude mamce zase končit dovolená* (ORAL2006).

Tato skupina se při srovnání výsledků anotace náhodného vzorku ukázala jako nejčetnější. Výraz *to* má ve zkoumaných konkordancích buď stejnou funkci jako při anafoře v psaných textech (odkaz do kontextu), nebo plní roli deiktika (odkaz do kontextu).

4.1.2 KATEGORIE „KOMBINACE“

Do druhé skupiny řadíme takové příklady, v nichž se výraz *to* kombinuje s minimálně jedním dalším komponentem a společně tvoří víceméně ustálenou jednotku, jež má význam jako celek (srov. Čermák, 2007, s. 31). Kromě frazémů a ustálených kolokací (včetně responzních frází, např. *to jo*) sem dále řadíme kombinaci *to + slovesa přírodních jevů a procesů* (např. *to prší*) a spojku *a to*.

V anotovaném vzorku byla co do frekvence tato kategorie na druhém místě. U relativně uzavřené množiny responzních frází se anotátoři shodovali, naopak se poměrně lišili v označování frazémů a ustálených spojení.

4.1.3 KATEGORIE „MLUVENOST“

Třetí skupina sdružuje příklady užití výrazu *to* v jevech typických pro mluvený nepřipravený projev, které se neobjevují v jazyce psaném (s výjimkou snah zachytit mluvenost). Tyto jevy vznikají nezáměrně v průběhu tvoření promluvy. Jedná se zejména o falešné starty (např. *no tak to to to bys tam jel docela na hovno teda*; ORAL2013),

⁵ Vzhledem k omezenému rozsahu tohoto článku představujeme vlastní kategorizaci jen stručně. Detailněji bude rozpracována v připravovaném samostatném článku.

nedokončené výpovědi (např. *ale máš prostě to ...*⁶; ORAL2008) a o užití slova *to* z důvodu neschopnosti vybavit si konkrétní slovo (např. *docela jako to jako mě pobavil kdysi kolega*; ORAL2008), obecně tedy jde o používání z pragmatických důvodů. Zejména posledně jmenovaným důvodem jsou motivovány specifické případy prefixace a reduplikace, kterým se věnujeme dále.

Do této kategorie bylo při anotaci zařazeno nejméně příkladů, což je způsobeno i tím, že hranice jednotlivých kategorií jsou prostupné a případy v nich zahrnuté plní více funkcí (viz např. dál zachovaná funkce odkazování v reduplikovaných příkladech).

4.2 PREFIXACE

Výraz *to* doplněný o předponu *se* v materiálu vyskytoval zejména v případech, kdy si mluvčí není v daném okamžiku schopen vybavit vhodný výraz. Přibližně ve 21 % případů došlo k následnému doplnění hledaného výrazu,⁷ který se nacházel v nejbližším pravém kontextu, např. *já sem si říkal, že si to normálně nato, naskenuju* (ORAL2008). Jiným případem je užívání prefigovaného *to* z tabuizačních důvodů (např. *poto, vyto, příp. i nato*). První dva výrazy jsou již do jisté míry lexikalizované.

V tab. 1 jsou uvedeny všechny varianty prefigovaného *to*, které se vyskytují v korpusech řady ORAL. Pro srovnání jsme tytéž varianty vyhledávali v korpusech psaného jazyka SYN. Ukázalo se, že prefigované *to* se vyskytuje pouze v beletrii a v publicistice, a to vždy v záznamu přímé či reprodukované řeči, např. Deníky Moravia 5. 3. 2009: *V úterý jste hráli pohár, ze kterého jste vypadli... Trošičku jsme to po... poto. (smích) Byla částečně proházená sestava.*

	ORAL	i.p.m.	SYN	i.p.m.
<i>neto</i>	238	40,02	19	0,01
<i>vyto</i>	62	9,92	2	0
<i>rozto</i>	31	5,21	1	0
<i>voto</i>	22	3,53	0	0
<i>zato</i>	20	3,40	—*	—*
<i>nato</i>	20	3,40	—*	—*
<i>přeto</i>	15	2,52	0	0
<i>přito</i>	13	2,19	0	0

	ORAL	i.p.m.	SYN	i.p.m.
<i>uto</i>	9	1,51	0	0
<i>doto</i>	4	0,67	0	0
<i>vodto</i>	3	0,50	1	0
<i>poto</i>	3	0,50	6	0
<i>odto</i>	2	0,34	0	0
<i>předto</i>	2	0,34	0	0
<i>vobto</i>	1	0,17	0	0
<i>proto</i>	1	0,17	—*	—*

TABULKA 1: Prefigované *to* v mluvených a psaných korpusech.

* Nehledáno, viz oddíl 3.

V psaném jazyce je na rozdíl od mluveného častější prefigovaný výraz *tento* (viz tab. 2). Od něj lze tvořit i slovesa, např. *potentovat, vytentovat, zatentovat* (Komárek, 2006, s. 73).

6 Značka ...: značí v korpusech řady ORAL neukončenou výpověď, na kterou už nebylo navázáno.

7 Počítáme pouze případy, kdy byl u doplněného slova použit stejný prefix.

	SYN	ORAL
<i>potento</i>	23	0
<i>vytento</i>	23	0
<i>netento</i>	17	1

	SYN	ORAL
<i>natento</i>	7	0
<i>přetento</i>	3	0
<i>zatento</i>	2	0

TABULKA 2: Prefigované *tento* v mluvených a psaných korpusech.

Výraz *neto* se od ostatních prefigovaných výrazů v tabulce liší jednak sémanticky (představuje vlastně negované *to*), jednak frekvenčně (v psaných i mluvených korpusech výrazně převažuje, viz tab. 1). Mezi jeho nejčastější bezprostřední levé kolokáty patří výraz *to*. V kolokaci *to neto* často výraz *to* odkazuje k tématu hovoru nebo zastupuje konkrétní slovo, zatímco výraz *neto* je užitím negovaného tvaru předpokládaným predikátem, např. *z jednou operací a náko se to neto a během tejdne další operace* (ORAL2013). Mezi nejčastějšími bezprostředně pravými kolokáty se objevují interpunkční značky (tečky, čárky), jejichž použití se sice ve sledovaných mluvených korpusech liší,⁸ přesto se vždy jedná o signály členící promluvu.

I u ostatních prefixů může naznačovat prefigované *to* slovesný predikát (zejména u *poto*, *vyto*), a to přímým doplněním v následujícím kontextu⁹ nebo začleněním do struktury složeného slovesného tvaru (např. *když už sem to měl tak . rozebraný tak sem to rovnou voto*; ORAL2013). Prefigovaný výraz může ale také zastupovat adjektivum, např. *tak naleze do těch prkýnek a sou vo* voto , vohnilý , no* (ORAL2008). Ve všech příkladech se přitom jedná o adjektiva verbální. V některých případech odkazuje prefigovaný tvar k celé propozici (např. *a televize to přeto ?*; ORAL2013) a je těžké blíže určit jeho funkci, pokud není v dalším kontextu doplněn.

Poměrně vysoká četnost prefigovaného *to* v autentických mluvených projevech naznačuje, že se jedná o jev hojně v komunikaci využívaný. Kromě tabuizační funkce umožňuje tento prostředek mluvčímu získat čas na vybavení si konkrétního slova, příp. dalším účastníkům komunikace naznačit jeho sémantiku. Dořčení hledaného slova, které jsme našli v pouhých 21 % příkladů, indikuje, že sémantika prefigovaného výrazu je v dané komunikační situaci dostatečně jasná.

4.3 REDUPLIKACE

Ve spontánní mluvené komunikaci se vedle výskytu „jednoduchého“ slova *to* objevuje i výskyt několikanásobný¹⁰. Podobně jako u prefixace se jedná o jev, který sice v rámci všech výskytů slova *to* není tak častý, ale jeho relativní frekvence je v mluveném ja-

⁸ V ORAL2006 a 2008 je použita větná interpunkce po vzoru psaného jazyka. V ORAL2013 značí tečky realizované pauzy (jedna tečka kratší pauzu, dvě tečky delší pauzu), čárka se vůbec nepoužívá.

⁹ Z celkových 21 % doplněných prefigovaných *to* bylo v 18 % doplněno sloveso.

¹⁰ Při vyhledávání dvou stejných tvarů vedle sebe pomocí dotazu 1:[] 2:[] & 1.word=2.word se nachází dvojice *to to* na 4. místě (za *no no*; *jo jo*; *ne ne*). V psaném korpusu SYN2010 se jedná o třetí dvojici za *se se a tak tak*, její frekvence je však řádově nižší.

zyce nápadně vyšší než v psaném (2,77 i.p.m. v SYN2010 proti 258,77 i.p.m. v korpusu ORAL). Zajímaly nás zde především případy nezáměrné reduplikace.

Ačkoli se nabízí zařazení všech výskytů reduplikovaného *to* do kategorie spojené s mluveností, není to klasifikace přímočará — i zde bylo možné najít v okolním kontextu různě hierarchizovaný konotát (slovo, celá propozice, předchozí sdělení, ať už mluvčího, či jeho adresáta), k němuž spojení *to to* odkazovalo.¹¹ Na základě kontextu jsme se u všech 1318 příkladů pokusili zobecnit, jaké role¹² může reduplikované užití *to to* v komunikaci mít:

- **odkazování:** Reduplikované spojení *to to* odkazuje k nějakému konkrétnímu konotátu: *sem vzal zas to to co máš ty* (ORAL2013); *podle mě ta* takovej ten . ten . to to pravoslaví* (ORAL2013).
- **hraniční bod pro změnu, přeformulaci předchozího sdělení, signalizující neukončení sdělení, získávání času**¹³: Násobné užití výrazu a s ním spojená potřeba (a možnost) získat čas k promyšlení následného sdělení nás vedou k domněnce, že *to to* bylo užito jako hraniční bod k přeformulování, doplnění či konkretizování předchozího sdělení. V případech přeformulací se objevují změny např. gramatických kategorií, perspektivy výpovědi. Doprovodným jevem (a tedy indikátorem váhání a nerozhodnosti v pokračování) je výskyt pauzy, intonačního předělu či delšího odmlčení: *nevopravujem prtože stejně to to dneska už . ta cena* (ORAL2013); *uznávám že to to . ale možná zhrším* (ORAL2013); *tak jak jako eee jak se to to . všeliké zařízení jak se modernizují* (ORAL2013); *prostě eee to to . no ale, jesi to to jesli .. se ten se musí zapnout* (ORAL2013); *takže to to mmm* (ORAL 2008, v posledních dvou případech je možno výraz *to to* chápat i jako výplňkové slovo).
- **posloupnost:** Posloupností míníme lineární vztah mezi oběma členy reduplikovaného spojení *to to*, kdy první *to* odkazuje k druhému *to*, jež zastupuje např. chybějící slovo (sloveso atd.). Můžeme říci, že násobné *to* bylo užito záměrně a nejedná se o opakování téhož výrazu. Nápadné jsou zde především zvukově percepční vlastnosti — druhé *to* je vysloveno s odlišným intonačním průběhem než *to* první, a to jak klesajícím, tak stoupajícím, či s nápadnou délkou slabiky (v našem případě celého slova). V jednom příkladu nalezneme dokonce v přepisu explicitní *to tó* (*a zalezu lákem , třeba když sem to tó*. [pak pokračuje druhý mluvčí]; ORAL2008).

V níže uvedených příkladech následuje po spojení *to to* pauza, resp. označení syntaktického předělu: *Vanessa a tak . a Oťák to to , zabalil* (ORAL2006); *než začneš chodit s nákým klukem , tak to to , musíme zkonzultovat* (ORAL2006).

Typem posloupnosti jsou i případy, kdy jednotlivé výrazy *to* náleží k různým syntaktickým konstrukcím: *počkej není to to , co nám tenkrát jak sme byli v Belgii [to + korelativum to, co]* (ORAL2013); *nešahej na to to mám pro děti k Vánocům [sahat na něco/to + mít něco/to]* (ORAL2013).

11 V anotovaném vzorku řazeno k odkazování.

12 Záměrně nepoužíváme slovo „funkce“, které má specifický význam, zde nám jde zjednodušeně o to, co se děje s textem v blízkém okolí výrazů *to to*, jaký je kontext.

13 Tyto doklady byly řazeny do kategorie mluvenosti.

- **částice:** Reduplikované spojení slouží i k vyjádření pragmatických postojů ze strany mluvčího, především k vyjádření expresivity. Gramatiky výraz *to to* v této roli klasifikují jako částici: *to to jí překvapíme* (ORAL2013); *no ježkovy voči to to tě prej byly byly krásy* (ORAL2013) (v tomto příkladu je expresivní složka posílena i užitým frazémem a emfatickým prodloužením první slabiky slova *ježkovy*); *no to to zas byly kecy teda* (ORAL2013).

4.3.1 TO TO V MLUVENÉM A PSANÉM KORPUSU

Specifický výskyt nejen reduplikovaného výrazu *to to*, ale obecně slova *to*, úkoly a funkce, které je schopno plnit, vedou k jeho dvojitmu popisu: jako rodové varianty zájmena *ten* a jako samostatného výrazu *to* (viz úvod). Toto specifikum potvrzuje i pohled na excerpce z konkordancí uvedené v následujících tabulkách, které zachycují typické výskyty reduplikovaného spojení *to to* v mluveném a psaném korpusu a porovnávají je s výskyty reduplikovaného *ten ten* a *ta ta*.

že	sem	to	to	.	pak
že	se	to	to	,	...
že	jako	to	to	,	prostě
čemu	by	to	to	,	že
čeho	bych	to	to	,	a

je	to	to	nejlepší	,	
je	to	to	,	co	
Je	to	to	nejlepší	,	
je	to	to	jediné	,	
Je	to	to	,	co	

TABULKA 3 A 4: Excerpce z konkordance *to to* v korpusu ORAL (vlevo) a SYN (vpravo).

že	jakoby	ta	ta	.	skupina
že	jako	ta	ta	věta	Paříž
škodovku	.	ta	ta	na	tu
říká	.	ta	ta	sociální	pracovnice
řikal	,	ta	ta	holka	že
že	ubrali	ten	ten	kus	toho
Žižkovskýho	vysílače	ten	ten	je	kousek
řek	ňákej	ten	ten	jo	a
četla	dneska	ten	ten	článek	..

TABULKA 5: Excerpce z konkordance výrazů *ta ta* a *ten ten* v korpusu ORAL.

Konkordance výrazů *to to* v mluveném korpusu ukazuje frekventované užití v místech váhání — následuje vždy pauza či označení syntaktického předělu. V psaných textech je naopak *to to* součástí složitějších konstrukcí.

Výrazy *ten ten* a *ta ta* se vyskytují pouze, respektive většinou v mluvených korpusech (*ta ta* se vyskytuje v SYN pouze jednou, a to ve spojení *ta ta žena*, výraz *ten ten* nemá v psaném korpusu žádný výskyt) a na rozdíl od varianty *to to* se v okolním kontextu objevuje substantivum v odpovídajícím rodě.

I na těchto příkladech vidíme, že se obecně výraz *to* odlišuje od slov *ten* a *ta*, vyskytuje se v jiných kolokacích, nesouvisí s rodovostí, ale poukazuje na svou univerzálnost.

5. ZÁVĚR

Tvar *to* je tradičně nejméně frekventovaným tvarem v korpusech mluvené češtiny (srov. Čermák, 2007). Jedná se o důležitý prostředek deixe a koreference (anafora), tj. odkazování do kontextu, resp. kotextu. Neformální spontánní jazyk sestává z nepřipravených projevů, jejichž organizace se vytváří postupně. Intenzivní odkazování na předměty hovoru v nich usnadňuje orientaci, neboť se opírá o znalost situace bezprostředně sdílenou všemi účastníky komunikace. Exploratorní sonda na základě korpusu ORAL ukazuje, že v nadpoloviční většině případů si užití výrazu *to* zachovává alespoň do jisté míry tuto odkazovací funkci. Bývá však i součástí nezáměrných falešných startů, nedokončených výpovědí a výplňků, prostředkem k tomu, aby mluvčí získal čas na přesnější formulaci svých myšlenek (hledání slov) a udržel si slovo. Zde je důležitý zejména univerzální charakter výrazu *to* a jeho značná vágnost, jež mu umožňuje zastupovat entity všech rodů i čísel nebo celé delší úseky textu. Že je v této roli *to* dobře etablované, dokazuje fakt, že snadno podstupuje prefixaci náhradou nejčastěji za slovesný tvar, na něž se mluvčímu nedaří vzpomenout (mluvčí má pak možnost zamýšlenou sémantiku naznačit právě alespoň výběrem prefixu). Reduplikace je pak motivována přímo váháním (*to se v takových případech tedy může stát i výplňkovým slovem*), přesto je zcela funkční — mluvčí jí dosahuje jednak času k promyšlení sdělení, případně k jeho přeformulaci, ale zároveň ji využívá i jako prostředku odkazování a usouvztažňování.

Stručný nástin rozsáhlého materiálu z jedné oblasti mluveného jazyka zve k dalšímu výzkumu, od oblasti odkazování přes ustálenost kolokací a jejich funkce v mluveném projevu až po další jevy typicky mluvené, jako jsou například responzní fráze.

POUŽITÉ KORPUSY

- BENEŠOVÁ, L. — KŘEN, M. — WACLAWIČOVÁ, M. (2013): *ORAL2013: reprezentativní korpus neformální mluvené češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z: <<http://www.korpus.cz>>
- KOPŘIVOVÁ, M. — WACLAWIČOVÁ, M. (2006): *ORAL2006: korpus neformální mluvené češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z: <<http://www.korpus.cz>>
- KŘEN, M. — BARTOŇ, T. — CVRČEK, V. — HNÁTKOVÁ, M. — JELÍNEK, T. — KOCEK, J. — NOVOTNÁ, R. — PETKEVIČ, V. — PROCHÁZKA, P. — SCHMIEDTOVÁ, V. — SKOUMALOVÁ, H. (2010): *SYN2010: žánrově vyvážený korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z: <<http://www.korpus.cz>>
- KŘEN, M. — ČERMÁK, F. — HLAVÁČOVÁ, J. — HNÁTKOVÁ, M. — JELÍNEK, T. — KOCEK, J. — KOPŘIVOVÁ, M. — NOVOTNÁ, R. — PETKEVIČ, V. — PROCHÁZKA, P. — SCHMIEDTOVÁ, V. — SKOUMALOVÁ, H. — ŠULC, M. (2014): *Korpus SYN, verze 3 z 27. 1. 2014*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z WWW: <<http://www.korpus.cz>>
- WACLAWIČOVÁ, M. — KOPŘIVOVÁ, M. — KŘEN, M. — VÁLKOVÁ, L. (2008): *ORAL2008: sociolingvisticky vyvážený korpus neformální mluvené češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z: <<http://www.korpus.cz>>

LITERATURA

- ADAM, R. (2001–2002): Kdy zdvojujeme deiktická slova. *Češtinář*, 12, s. 71–72.
- ADAMEC, P. (1998): K syntaxi a sémantice „malých slov“: slůvko to. In: *Přednášky z XLI. běhu LŠSS*. Praha: Univerzita Karlova, s. 55–58.
- ČERMÁK, F. & KOL. (2007): *Frekvenční slovník mluvené češtiny*. Praha: Karolinum.
- ČERMÁK, F. (2007): *Frazeologie a idiomatiky česká a obecná*. Praha: Karolinum.
- ČERMÁK, F. (2010): *Lexikon a sémantika*. Praha: Nakladatelství Lidové noviny.
- HAVRÁNEK, B. — JEDLIČKA, A. (1960): *Česká mluvnice*. Praha: SPN.
- HAVRÁNEK, B. a kol. (1989): *Slovník spisovného jazyka českého*. 2., nezměněné vydání. Praha: Academia.
- HIRŠCHOVÁ, M. (2013): *Pragmatika v češtině*. Praha: Karolinum.
- JELÍNEK, M. (1995): Stylistika. In: Karlík, P. — Nekula, M. — Rusínová, Z. (eds.), *Příruční mluvnice češtiny*. Praha: Nakladatelství Lidové noviny, s. 701–782.
- KOLÁŘOVÁ, I. (2014): *Význam a konkurence vybraných druhů odkazovacích zájmen v kontextu*. Brno: Pedagogická fakulta Masarykovy univerzity. Rukopis.
- KOMÁREK, M. (2006): *Příspěvky k české morfologii*. Olomouc: Periplum.
- LAKOFF, G. — JOHNSON, M. (2002): *Metafory, kterými žijeme*. Brno: Host.
- LYONS, J. (1977): *Semantics I, II*. Cambridge: University Press.
- MACHEK, V. (1957): *Etymologický slovník jazyka českého a slovenského*. Praha: Nakladatelství Československé akademie věd.
- PETR, J. & KOL. (1986): *Mluvnice češtiny 2*. Praha: Academia.
- VYCHODILOVÁ, Z. (2008): *Výrazové prostředky textové soudržnosti v češtině v porovnání s ruštinou (na materiále zájmených lexémů „ten/etot“, „to/eto“ a jejich funkčních ekvivalentů)*. Olomouc: Univerzita Palackého.

Petra Klimešová | Ústav Českého národního korpusu, FFUK | nám. J. Palacha 2, 116 38 Praha 1
petra.klimešova@ff.cuni.cz

Zuzana Komrsková | Ústav Českého národního korpusu, FFUK | nám. J. Palacha 2, 116 38 Praha 1
zuzana.komrskova@ff.cuni.cz

Marie Kopřivová | Ústav Českého národního korpusu, FFUK | nám. J. Palacha 2, 116 38 Praha 1
marie.koprivova@ff.cuni.cz

David Lukeš | Ústav Českého národního korpusu, FFUK | nám. J. Palacha 2, 116 38 Praha 1
david.lukes@ff.cuni.cz