

Koreference a typy reference v češtině

Jiří Pergler

ABSTRACT:

Coreference and types of reference in Czech. The article deals with coreference relations in Czech and their classification based on the classification of reference types. First, we define the notion of coreference, putting stress on the distinction between coreference and endophora. Then the scheme of reference and coreference used in the annotation of the coreferential relations in the Prague Dependency Treebank (PDT) is presented, where two types of coreferential relations are applied: the type SPEC and the type GEN. Analyzing several examples, we argue that this scheme is not capable of describing more complicated reference relations, which results especially from the too broad definition of the type GEN. Hence we suggest an alternative classification of the types of reference, employing two independent criteria: the criterion of genericity and the criterion of specificity, so that four types of reference are established: specific individual, non-specific individual, specific generic and non-specific generic. We deal with each of these four types in detail, and finally, we show how this classification can be used when capturing coreference relations in text.

KLÍČOVÁ SLOVA / KEY WORDS:

endofora, generičnost, koreference, Pražský závislostní korpus, reference, specifičnost coreference, endophora, genericity, Prague Dependency Treebank, reference, specificity

0. ÚVOD

V tomto článku se budeme zabývat vztahy koreference v češtině a jejich klasifikací založenou na typech reference. Východiskem pro nás bude model koreference uplatněný při manuální anotaci koreferenčních vztahů v Pražském závislostním korpusu (PDT). Po krátkém uvedení do tohoto modelu upozorníme na některé problémy, které jsou spojeny s dichotomickou klasifikací reference, s níž pracuje koncepce PDT, a pokusíme se s využitím odborné literatury navrhnout alternativní model klasifikace reference, pomocí něhož by bylo možno uvedené problematické případy adekvátněji popisovat.¹

1. POJETÍ KOREFERENCE

Koreferencí rozumíme vztah mezi dvěma nebo více referenčními výrazy v textu, které referují k téže entitě, tj. jejichž referent je identický. Z definice se tedy jedná o vztah symetrický (pokud A je koreferenční s B, pak také B je koreferenční s A)

¹ Tomuto tématu se podrobněji věnujeme ve své bakalářské práci (Pergler, 2010).

a tranzitivní (pokud A je koreferenční s B a B je koreferenční s C, pak také A je koreferenční s C). Mimo jiné i z tohoto důvodu je nutné vztah koreference zřetelně odlišovat od vztahů vnitrotextového odkazování v textu, tedy od vztahů endofory (tj. anafory a katafory).^{2, 3} Vztahy koreference a endofory se v textech často vyskytují současně (jako je tomu ve větě (1),⁴ jedná se v takovém případě o koreferenční endoforu), není tomu tak ale vždy. Obvyklé jsou i případy endofory nekoreferenční, kde se endoforický vztah vyskytuje, zatímco vztah koreference nikoli (příklady (2)–(5)).

- (1) Máme *dodavatele*, s nimiž spolupracujeme.
- (2) Máme *dodavatele*, ale potřebovali bychom *jich* více.
- (3) Pozvali jsme deset *přátel*, ale dorazilo *jich* jen pět.
- (4) Počet *dopravních nehod* za loňský rok překročil 12 000. V roce předchozím *jich* bylo zaznamenáno 7 565 (PDT).
- (5) Podle vládního návrhu by *žijícím občanům poškozeným nacistickou perzekucí* nebo vdovám či vdovcům *po nich* měla být poskytnuta částka 2300 korun [...] (PDT).

Kromě toho se v textech vyskytují i případy vztahů koreference bez přítomnosti vztahu endoforického. Podmínky, za nichž se tak často děje, formuluje Nédolužko (2011, s. 69–70). Jedná se podle ní zpravidla o případy, kdy se oba koreferenční výrazy nacházejí v rématu, kdy jsou koreferenční výrazy v různých diskurzivních jednotkách (v rámci téhož textu) nebo kdy je mezi oběma výrazy velká textová vzdálenost.

K uvedenému vymezení koreference a zřetelnému odlišení koreference od endofory se teoreticky hlásí i model koreference uplatňovaný v PDT,⁵ ovšem prakticky tyto zásady nejsou v anotaci dodržovány zcela důsledně. Deklarovanou symetričnost koreference zpochybňuje především samotný fakt, že koreferenční vztahy jsou zde zachyceny pomocí orientovaných hran. Za další příklad nedodržování principu symetričnosti pokládáme skutečnost, že slovesa mohou v uvedeném modelu fungovat pouze jako počáteční člen koreferenčního řetězce, a nemohou tedy stát ve slovosledné pozici napravo od koreferenčního výrazu.⁶ U symetrické relace by takové omezení pochopitelně postrádalo smysl. Za důsledek nedůsledného odlišení koreference od endofory pokládáme též skutečnost, že do anotace koreference

2 Vymezení endofory se zde nebudeme podrobně věnovat, je však zřejmé, že na rozdíl od koreference se jedná o vztah asymetrický (pokud výraz A v textu endoforicky odkazuje na výraz B, pak neplatí, že by také výraz B odkazoval na výraz A). Podle našeho názoru je tento vztah vhodné pojímat také jako intranzitivní (tedy předpokládat, že endoforický výraz má v textu pouze jeden antecedent).

3 Na možnost odlišení koreference a endofory upozorňují např. již Hajičová et al. (1987, s. 199).

4 Pokud není uvedeno jinak, jsou příkladové věty použité v tomto článku vlastní, zpravidla převzaté z práce Pergler (2010).

5 Viz např. Nédolužko (2011, s. 69).

6 O tom viz dále, příp. viz také Nédolužko (2011, s. 84–85).

nebyla v popisované fázi anotace zapojena osobní zájmena pro první a druhou osobu (přestože tato zájmena pochopitelně referenci mají a jejich účasti na koreferenčních vztazích by nic nemělo bránit). Jako jeden z důvodů pro toto rozhodnutí Nédolužko (2011, s. 77) uvádí, že u těchto zájmen „nemohou fungovat anaforická pravidla“. Pokud by byla koreference od endofory důsledně oddělena, postrádal by tento argument relevanci.

2. KOREFERENCE V PDT

Vztahy koreference jsou v PDT manuálně anotovány, a to v několika fázích. V první fázi, kterou popisují např. Mikulová et al. (2005, s. 935–1055), byla provedena anotace u osobních a přivlastňovacích zájmen pro třetí osobu, všech zvrtných zájmen, zájmena *ten* v substantivní funkci a vztažných zájmen a příslovčí. Ve druhé fázi jsou pak anotovány ostatní případy koreference,⁷ podrobnější přehled o této fázi podává Nédolužko (2011). Později anotace procházela ještě několika méně významnými úpravami (např. doplnění anotace u zájmen pro první a druhou osobu), kterým pro účely tohoto článku není třeba věnovat zvláštní pozornost.

Anotace koreference je v PDT zapojena do tektogramatické (významové) roviny věty a je zachycována pomocí orientovaných hran mezi uzly tektogramatických stromů reprezentujícími příslušné koreferenční výrazy. Množiny uzlů navzájem propojených koreferenčními vztahy tvoří koreferenční řetězce. Do koreferenčních vztahů mohou vstupovat všechny referenční výrazy v textu, tedy především substantiva, zájmena a některá adjektiva, příslovce a číslovky.⁸ Kromě těchto slovních druhů se koreferenčních vztahů účastní také slovesa, avšak ta mohou vystupovat pouze jako první člen koreferenčního řetězce, tj. nemohou být ve vztahu koreference s výrazem, který je slovosledně nalevo od nich.⁹

Vztahy koreference jsou v PDT klasifikovány podle dvou nezávislých kritérií. Na základě prvního z nich se koreference dělí na koreferenci gramatickou a koreferenci textovou. Rozdíl mezi nimi spočívá v tom, že zatímco u gramatické koreference „je možné určit koreferovaný člen na základě gramatických pravidel“ (Mikulová et al., 2005, s. 937), u textové koreference něco takového možné není. Koreference gramatická se týká zvrtných zájmen, vztažných prostředků, doplnění s dvojí závislostí vyjádřených slovesnou formou (tj. např. koreference aktantů slovesných přechodníků),

7 Kromě toho se v rámci této fáze provádí ještě anotace tzv. asociační (bridging) anafory.

8 Samozřejmě ne všechny výrazy s touto slovnědruhovou platností mají referenční funkci. Za referenční se obvykle nepokládají např. substantiva v pozici predikátu.

9 Toto řešení pokládáme za nesprávné, neboť jeho důsledkem je skutečnost, že zatímco např. ve větách *Petr pracoval celé odpoledne. Jeho práce však nepřinesla žádné výsledky* je vztah koreference mezi výrazy *pracoval* a *práce* anotován, ve větách typu *Petrova práce nepřinesla žádné výsledky. Pracoval přitom celé odpoledne* žádný vztah mezi výrazy *práce* a *pracoval* anotován není. Domníváme se, že reference těchto výrazů se v závislosti na jejich slovosledném pořadí v textu nemění, a proto není důvod toto pořadí zohledňovat při anotaci koreference.

kontroly, tzv. kvazikontroly a recipročních konstrukcí.¹⁰ Ostatní koreferenční vztahy se chápou jako případy koreference textové.¹¹

Druhým členěním, o které nám v tomto článku půjde především, je členění koreferenčních vztahů podle typu reference výrazů, které se těchto vztahů účastní.¹² Na základě tohoto kritéria se rozlišují dva typy: typ SPEC a typ GEN.

Typ SPEC představují především vztahy mezi výrazy s tzv. specifickou referencí. Specifickou referencí se přitom rozumí „odkazování ke konkrétnímu existujícímu, reálnému referentu a objektu skutečnosti“ (Nědolužko, 2011, s. 104). Nejčastějším případem koreference typu SPEC je vztah mezi dvěma jmennými frázemi, z nichž obě mají specifickou referenci. Tento typ koreference je však anotován též v některých případech, kdy antecedent¹³ má referenci nespécifickou. Děje se tak tehdy, pokud jeho reference není generická a pokud zároveň postcedent obsahuje určitý identifikátor.¹⁴ Příkladem může být např. věta (6) převzatá z práce Nědolužko (2011, s. 107):

- (6) Přesto si značky mohl všimnout jen *někdo velice pozorný* [...] a ani *ten velice pozorný člověk* by jim patrně nepřikládal žádný význam (Nědolužko, 2011).

Antecedentem je zde výraz *někdo velice pozorný* a jeho reference je nespécifická, neboť nereferuje k žádnému vybranému (identifikovatelnému) člověku. V případě koreferenčního výrazu *ten velice pozorný člověk* se předpokládá, že jeho reference je důsledkem přítomnosti určitého identifikátoru *ten* specifická, neboť je dostatečně specifikována kontextem. Mezi oběma výrazy je tak anotován vztah koreference typu SPEC.

Druhým typem je typ GEN, kterého nabývají výrazy s nespécifickou nebo generickou referencí. Generická reference je v PDT vymezována jako reference k typu objektů, přičemž platí, že tato reference „se nemusí vztahovat na všechny objekty daného typu“ (Nědolužko, 2011, s. 99). Autorka uvádí dva testy generičnosti (Nědolužko, 2011, s. 108–109). Podle prvního z nich je jmenná fráze generická, „pokud X může být použit v konstrukci X jako (typický) Y, X jako druh (forma) Y“.¹⁵ Podle druhého testu

¹⁰ Viz např. Mikulová et al. (2005, s. 937–1039).

¹¹ Přísně vzato, netýká se tato „gramatičnost“ a „textovost“ vztahů koreference, nýbrž vztahů endofory (koreferenčních i nekoreferenčních). Existence pojmů gramatická a textová koreference je tak dalším příkladem nedostatečného rozlišování mezi koreferencí a endoforou.

¹² Toto rozdělení se v anotaci týká pouze koreference textové, ovšem teoreticky není důvod je omezovat jen na ni. V PDT lze toto omezení vysvětlit tím, že uvedené rozdělení začalo být uplatňováno až ve druhé fázi anotace (zatímco gramatická koreference byla anotována již ve fázi první).

¹³ Antecedentem budeme mít v tomto článku na mysli ten z dvojice výrazů (propojených vztahem koreference nebo endofory), který je slovosledně více vlevo. Výraz, který je více vpravo, pak budeme označovat jako postcedent.

¹⁴ Pojem „určitý identifikátor“ ovšem Nědolužko (2011) blíže nevymezuje. Lze předpokládat, že za určitý identifikátor se pokládá např. osobní zájmeno či spojení substantiva se zájmeny *ten, tento*.

¹⁵ Např. větu *Jezdí vlakem* je možné upravit takto: *Jezdí vlakem, protože je to nejlevnější dopravní prostředek* (Nědolužko, 2011, s. 108). Tento test přebírá autorka od Rachilinové a Krejdlina (1981).

je generická taková jmenná fráze, kterou lze nahradit spojením *ten/to, kdo/co je X*. Nespecifickou referencí se pak rozumí taková reference, která nesplňuje kritéria reference specifické. Podstatu nespecifičnosti ani vzájemný vztah nespecifičnosti a generičnosti však Nédolužko (2011) blíže neobjasňuje.

3. PROBLÉMY DICHOTOMICKÉHO MODELU PDT

Popsaný model naráží při analýze koreferenčních vztahů v textech na některé typy případů, které není schopen adekvátně popisovat. Pozornost zde soustředíme především na ty problematické jevy, které jsou spojeny s popsáním dichotomickým rozdělením reference na typy SPEC a GEN, nebudeme se zde tedy podrobněji věnovat problémům souvisejícím spíše s technickou stránkou anotace¹⁶ nebo těm, které vyplývají z pojetí tektogramatické roviny v PDT.¹⁷

Za první typ těchto problematických jevů pokládáme způsob, jakým jsou anotovány výrazy s nespecifickou a negenerickou referencí, které obsahují určitý identifikátor. Jak jsme již uvedli, v případě přítomnosti určitého identifikátoru je reference postcedentu v těchto případech chápána jako specifická a anotován je zde typ SPEC. Příkladem může být již uvedená věta (6) či dvojice vět (7):

(7) Chtěla by mít *dceru*. Dala by jí jméno *Natálie*.

V modelu PDT je reference zájmena *jí* ve větě (7) chápána jako specifická, přičemž se předpokládá, že její specifičnost je dána kontextem (tedy výrazem *dceru*), ke kterému poukazuje užití určitého identifikátoru.

16 Technickou stránkou anotace rozumíme např. skutečnost, že typy SPEC a GEN nejsou připsovány přímo referenčním výrazům, jak by bylo lze očekávat, nýbrž jednotlivým vztahům koreference, přičemž se rozhoduje vždy podle typu reference postcedentu. Důsledkem toho anotace neobsahuje informaci o typu reference prvního členu koreferenčního řetězce.

17 Příkladem takového problému může být např. problematika předložkových frází vyjadřujících zpravidla místní či časové určení, která mají dvojí referenci (např. *po válce, před domem*). V těchto případech referuje jednak celá předložková fráze, jednak její substantivní část (na tuto skutečnost poukazuje např. možnost dvojí různé „pronominalizace“ takových frází, např. spojení *po válce* lze substituovat jak výrazem *tehdy*, tedy pronominalizovat celou předložkovou frází, tak i výrazem *po ní*, kde se pronominalizuje pouze substantivní část). Jelikož na tektogramatické rovině nejsou předložky reprezentovány zvláštními uzly, bylo nutné se při anotaci rozhodnout, zda v podobných případech vycházet z reference celých předložkových frází, nebo jenom substantivních částí, přičemž obě možná rozhodnutí vedou k tomu, že informace o koreferenci nemůže být úplná. Při anotaci byla nakonec zvolena druhá možnost. Negativním důsledkem je jednak fakt, že jsou jako vzájemně koreferenční anotována např. spojení *před válkou* a *po válce*, jednak skutečnost, že např. výrazy *po válce* a *v poválečném období* jako koreferenční vůbec anotovány nejsou.

Toto řešení nepokládáme za správné z několika důvodů.¹⁸ Především se domníváme, že použití určitého identifikátoru v textu nemá vliv na referenční kategorii specifičnosti, nýbrž pouze na kategorii určitosti. Podle našeho názoru rys určitosti (tj. skutečnost, že se referent prezentuje jako známý) neimplikuje rys specifičnosti.¹⁹ Toto tvrzení opíráme o skutečnost, že výraz *jí* ve větě (7) zcela jistě nenaplnuje výše citovanou definici specifičnosti, podle které se v případě specifické reference musí jednat o existující referent a objekt skutečnosti, a reference tohoto výrazu je proto nesespecifická. Lze to vysvětlit i nesespecifickou referencí antecedentu — jelikož antecedent sám má referenci nesespecifickou, je logické, že poukazem k němu nemůže získat specifickou referenci postcedent. Určitý identifikátor tedy neukazuje na specifičnost příslušné reference, nýbrž pouze na přítomnost koreferenčního vztahu mezi postcedentem a antecedentem, přičemž tento koreferenční vztah specifičnost reference neimplikuje.

Dalším argumentem proti uvedenému pojetí PDT je vázanost typu reference na přítomnost formálního prostředku (tedy určitého identifikátoru). Lze si totiž představit kontexty, ve kterých je užití určitého identifikátoru fakultativní a jeho případné vypuštění nemění referenční strukturu textu:

- (8) Chtěli bychom s touto firmou navázat *spolupráci*.
- (9a) *Tato spolupráce* by trvala jeden rok.
- (9b) *Spolupráce* by trvala jeden rok.
- (9c) Trvala by jeden rok.

Na větu (8) lze navázat všemi třemi větami (9a–c), přičemž nepokládáme za oprávněně předpokládat, že by typ reference subjektivních výrazů v příkladech (9a–c) byl odlišný. Ze skutečnosti, že se uvedené varianty liší co do přítomnosti určitého identifikátoru, vyvozujeme, že vázat typ reference na přítomnost tohoto formálního prostředku není náležité.

Druhý typ problematických případů lze demonstrovat na větě (10):

- (10) Obratlovci mají *oči*.

Předmětem našeho zájmu je typ reference výrazu *oči*. Je zřejmé, že se tu nejedná o referenci specifickou, neboť množina prvků, k níž tento výraz referuje, není uzavřená, tj. nelze ji stanovit výčtem všech prvků. Podobné výrazy jsou tak v anotaci PDT zařazeny do typu GEN, který je vyhrazen především výrazům s referencí generickou. Jenže se zdá, že reference výrazu *oči* v příkladu (10) generická není, neboť v uvedené větě tento výraz rozhodně nerefereje k celému druhu „oči“.²⁰

18 Kromě dále uvedených důvodů lze za argument proti uvedenému pojetí pokládat také skutečnost uvedenou v poznámce 16, tj. že v anotaci není obsažena informace o prvních členech koreferenčních řetězců. Případy typu (7) jsou tak anotovány stejně jako běžné případy specifické reference.

19 Tímto pojetím vzájemného vztahu určitosti a specifičnosti se dostáváme do rozporu s pojetím určitosti (resp. určenosti) např. u Hlavsy (1975) či Adamce (1980).

20 Uváděli jsme sice výše citát z práce Někola (2011, s. 99), že generická reference „se nemusí vztahovat na všechny objekty daného typu“, předpokládáme ovšem, že objekty,

Problém by nastal, pokud by v příslušném textu za větou (10) následovala např. věta (11):

(11) Hmyz má také oči.

Pokud bychom připustili, že reference výrazu *oči* ve větě (10) je generická, museli bychom generickou referenci předpokládat také u výrazu *oči* ve větě (11). Vzhledem k tomu, že by tak oba výrazy referovaly ke stejnému druhu (tj. druhu „oči“), museli bychom mezi nimi předpokládat vztah koreference, a to přesto, že množiny, k nimž tyto výrazy referují, jsou disjunktní. Předpoklad koreferenčního vztahu mezi disjunktními množinami by vedl k tomu, že by pojem koreference získal značně neintuitivní povahu, nicméně sám o sobě by nemusel představovat zásadní problém, neboť v případě koreference generických výrazů nejde o identitu jednotlivých exemplářů, nýbrž o identitu druhů, ke které v uvedených případech skutečně dochází. Představme si ale situaci, že v textu po sobě následují v uvedeném pořadí věty (10), (11) a (12).

(12) *Tyto oči* však vypadají zcela jinak než *oči obratlovců*.

Na první pohled se zdá, že výrazy *tyto oči* a *oči obratlovců* ve větě (12) jsou ve vztahu koreference s výrazy *oči* ve větách (11) a (10) (v tomto pořadí). Pokud ale předpokládáme, že koreference je vztah symetrický a tranzitivní, tak z koreference výrazů *oči* (10) — *oči obratlovců* (12), *oči* (11) — *tyto oči* (12) a *oči* (10) — *oči* (11) vyplývá také vzájemná koreference obou referenčních výrazů ve větě (12). Věta (12) tak získává interpretaci „*x* vypadá zcela jinak než *x*“, což je ovšem samozřejmě interpretace nesprávná. Dichotomické rozdělení reference na typy SPEC a GEN tedy v tomto a podobných případech selhává.

Třetí skupinu problematických případů představují věty typu (13):

(13) Když je *stůl* velký, vejde se *na něj* spousta věcí.

(14) Na *stůl, který je velký*, se vejde spousta věcí.

Věta (13) je významově podobná větě (14), ovšem referenční struktura je zde jiná. Zatímco ve větě (14) najdeme referenční frázi *stůl, který je velký*, kterou lze pokládat za generickou,^{21, 22} ve větě (13) je třeba se zabývat referenčními výrazy *stůl* a *něj*.

k nimž se taková reference nevztahuje, jsou v jistém smyslu spíše výjimečné. Ve větě (10) se o takovou výjimku podle našeho názoru nejedná.

21 Např. v práci redigované Carlsonem a Pelletierem (1995) by ovšem analogická fráze v angličtině za generickou pokládána nebyla. K tomuto rozhodnutí vedou autory formální důvody: v angličtině může mít spojení určitého členu se substantivem v singuláru generickou interpretaci pouze tehdy, pokud referent představuje tzv. „well-established kind“. V češtině však podobné formální omezení zřejmě neexistuje, jak ukazuje i uvedená fráze *stůl, který je velký*.

22 I zde se však setkáme s jistými obtížemi týkajícími se vztažného zájmena *který*. Lze je ale vyřešit tím, že za referenční budeme pokládat tuto frázi jako celek.

Oba výrazy ve větě (13) jistě nemají referenci specifickou (neboť nereferují ke konkrétnímu objektu), model PDT jim tedy připisuje referenci typu GEN, přičemž se předpokládá, že jde o výrazy s generickou referencí. Položme si nyní otázku po vzájemném vztahu výrazů *stůl* a *něj*. Mohlo by se zdát samozřejmé, že mezi těmito výrazy existuje vztah koreference. Detailním rozbořením situace ale zjistíme, že model PDT při důsledné aplikaci všech pravidel neumožňuje jejich vztah jako koreferenci popsat.

Tím, že u výrazu *stůl* předpokládáme generickou referenci, tvrdíme, že referuje ke všem exemplářům tohoto druhu (s možnými výjimkami), tj. (téměř) ke všem myslitelným stolům. Zájmeno *něj* v hlavní větě ovšem referuje pouze k těm stolům, které jsou velké, rozhodně tedy nereferuje ke stejnému druhu jako výraz *stůl*. Tyto výrazy tedy nemohou být koreferenční. Uvedenou úvahu dokládá i situace, ve které by za větou (13) následovala věta (15):

(15) Když je ale *stůl* malý, nevejde se na něj skoro nic.

Při analýze tohoto případu bychom se nacházeli v podobné situaci, jakou jsme popsali u vět (10)–(12). Museli bychom předpokládat vzájemnou koreferenci obou výrazů *stůl*, a pokud bychom připustili koreferenční vztah zájmen *něj* se svými antecedenty (tj. výrazy *stůl*), museli bychom vzhledem k symetričnosti a tranzitivitě koreference připustit též vzájemnou koreferenci zájmen *něj*. Takové řešení je opět v příkrém rozporu s intuicí, neboť předpokládá totožnost množiny stolů, na něž se vejde spousta věcí, s množinou stolů, na které se nevejde skoro nic. Ani tyto případy tedy model PDT nedokáže adekvátně popsat.

4. NÁVRH KLASIFIKACE REFERENCE

Při návrhu vlastního modelu typů reference vycházíme z předpokladu, že příčinou výše popsaných problémů je především příliš široké vymezení typu GEN. Budeme proto usilovat o takovou klasifikaci, která by byla založena na podrobnějším členění reference do více typů.

Takové pojetí klasifikace reference není zcela samozřejmé, jak ukazuje např. přístup Palkův (1988), který navrhuje pouze dichotomické třídění na denotáty-konstanty a denotáty-proměnné. Podrobnější klasifikaci pro češtinu navrhuje Hlavsa (1975), který však vychází především z opozice určitosti (kterou v našem pojetí nepokládáme za zásadní), což vede k situaci, že z šesti navrhovaných typů reference jich pět odpovídá typu SPEC v modelu PDT a pouze jeden „problematickému“ typu GEN. Za inspirativní pokládáme článek Adamcův (1980), který kromě reference singulativní (tj. přibližně typu SPEC) a generické mluví též o tzv. referenci podmíněně singulativní, při které mají jmenné fráze „naději stát se singulativními teprve při realizaci obsahu predikátu“ (Adamec, 1980, s. 257–258). Příkladem může být např. výraz *auto* ve větě (16):

(16) Chci si koupit *auto*.

Pro náš návrh však budeme vycházet především z pojetí popsaného v práci editorů Carlsona a Pelletiera (1995; dále cit. jako GenBook), ze kterého přebíráme třídění reference podle dvou nezávislých kritérií — kritéria generičnosti a kritéria specifičnosti.

Kritérium generičnosti se týká povahy referentu — v případě, že je referentem nějakého výrazu individuální objekt nebo množina objektů,²³ budeme mluvit o referenci individuální, naopak o referenci generické budeme mluvit tehdy, bude-li referentem příslušného výrazu druh. Na kritériu generičnosti je logicky nezávislé kritérium specifičnosti, podle kterého budeme členit referenci na specifickou a nespecifickou. O specifickou referenci půjde tehdy, bude-li referentem jasně vymezená (tj. nikoli nekonečná) množina prvků tvořící součást světa textu, v opačném případě se bude jednat o referenci nespecifickou.

Kombinací těchto dvou kritérií získáme celkem čtyři typy reference — specifickou individuální, specifickou generickou, nespecifickou individuální a nespecifickou generickou. Toto rozdělení typů reference pochopitelně není a nemá být vyčerpávající klasifikací reference v češtině, jak již bylo naznačeno, nepočítáme zde např. s opozicí určitost/neurčitost, která je, zabýváme-li se referencí, jistě opozicí zásadní. Důvodem je, že naše klasifikace má sloužit jako prostředek k popisu koreferenčních vztahů a jako taková zohledňuje pouze ta kritéria, která jsou pro popis koreferenčních vztahů relevantní, tj. ta kritéria, která se týkají přímo samotné povahy referentu. Opozice určitost/neurčitost se přitom netýká přímo povahy referentu, nýbrž „pouze“ toho, zda je referent prezentován jako známý, a není proto v naší klasifikaci zohledněna.

Aby nedošlo k nedorozumění, je třeba ještě zdůraznit, že referenci chápeme vždy jako vlastnost jmenné fráze v jejím konkrétním užití, nikoli jako vlastnost jmenné fráze jako takové (tj. bez ohledu na aktuální kontext).²⁴ V některých výpovědích přitom může být jmenná fráze referenčně nejednoznačná, jako příklad lze uvést větu (i) zmíněnou jedním z anonymních recenzentů:

(i) Petr se chce setkat se dvěma politiky.

Uvedená věta je dvojnásobná: může znamenat, buď že (1) existují dva konkrétní politici, s nimiž se Petr chce setkat, nebo že (2) Petr se chce setkat s libovolnými dvěma politiky. Citovaný recenzent pokládá existenci podobných vět za argument proti naší klasifikaci reference, neboť tato klasifikace údajně nedovede dvojnásobnost podobných vět vysvětlit. Cílem naší klasifikace ovšem není explanace, nýbrž deskripce, přičemž jak nyní ukážu, z tohoto hlediska pro nás nepředstavují dvojnásobné věty tohoto typu žádný problém. V uvedené větě má zvládnutá jmenná fráze (zcela v souladu s níže uvedenými vymezeními) ve významu (1) specifickou individuální referenci, ve významu (2) nespecifickou individuální referenci. V konkrétním jazykovém projevu

²³ Individuální objekt je pochopitelně možno chápat jako jednoprvkovou množinu objektů.

²⁴ Je tedy samozřejmé, že dvě různá užití formálně shodné jmenné fráze mohou mít jiný typ reference, např. ve větě *Na druhém konci ulice stojí velké domy* má jmenná fráze *velké domy* specifickou individuální referenci, zatímco ve větě *Velké domy jsou dražší než malé domy* má formálně shodná jmenná fráze specifickou generickou referenci.

pak popsaná nejednoznačnost může nebo nemusí být odstraněna širším kontextem.²⁵ Pokud odstraněna není, zůstává součástí daného textu a jako taková musí být reflektována dvojím možným lingvistickým popisem.

V následujících částech představíme podrobněji všechny uvedené typy a zároveň se pokusíme o přesnější vymezení dosud jen zjednodušeně nastíněných kritérií.

4.1 SPECIFICKÁ INDIVIDUÁLNÍ REFERENCE

Pro specifickou individuální referenci (z hlediska kritéria generičnosti) platí, že jejím referentem je individuální objekt nebo množina objektů. Podívejme se však nyní podrobněji na kritérium specifičnosti.

Jak jsme již naznačili, je toto kritérium založeno na tom, zda je referentem jasně vymezená (tj. uzavřená) množina prvků, a také na tom, jestli referent tvoří součást světa textu. Tyto dvě podmínky je možné převést na dvě binární kritéria — kritérium vymezenosti (s hodnotami [+vymezenost] pro výraz referující k uzavřené množině a [-vymezenost] pro výraz referující k otevřené množině)²⁶ a kritérium reálnosti (s hodnotami [+reálnost] pro výraz s referentem tvořícím součást světa textu a [-reálnost]²⁷ pro výraz s referentem netvořícím součást světa textu, tj. referentem, který figuruje jako součást světa textu pouze hypoteticky). Příkladem jednotlivých typů mohou být výrazy *auto/auta* ve větách (17)–(20):

- (17) Mám *auto* [+vymezenost], [+reálnost].
- (18) Chtěl bych mít *auto* [+vymezenost], [-reálnost].
- (19) Lidé mají *auta* [-vymezenost], [+reálnost].
- (20) Lidé by chtěli mít *auta* [-vymezenost], [-reálnost].

Na základě kritérií vymezenosti a reálnosti je možné přistoupit k definici specifické reference: specifická reference je taková reference, která je vymezená a (zároveň) reálná.

Identifikace specifické individuální reference v textech není podle našeho názoru spojena s výraznějšími obtížemi. V zásadě platí, že se tento typ reference kryje s typem SPEC ve schématu PDT, výjimku tvoří pouze výše popsané případy ilustrované větou (7), kde u výrazů typu *jí* nepředpokládáme referenci specifickou, nýbrž referenci (individuální) nespécifickou, neboť takové výrazy evidentně mají rys [-reálnost].

Přesto však zde chceme upozornit na dva typy případů, ve kterých příslušnost výrazu ke specifické individuální referenci není zcela jasná. První typ případů sou-

²⁵ Např. pokud by následovala za větou (i) věta *Oba ale již schůzku odmítli*, znamenalo by to, že první věta je míněna ve významu (1), tj. že příslušná jmenná fráze má specifickou individuální referenci. Význam (2) (a tedy nespécifickou individuální referenci) by implikovalo např. pokračování *Je mu jedno, z jakých by byli politických stran*.

²⁶ Uzavřenou množinou je taková množina, kterou lze charakterizovat úplným výčtem jejích prvků. Otevřenou množinu naopak takto charakterizovat nelze, podrobněji viz níže v oddílu 4.3.

²⁷ Ve své bakalářské práci (Pergler, 2010) používáme pro rys [-reálnost] termínu hypotetičnost.

visí s problematikou distributivního přiřazování množin, jíž se věnuje např. Hlavsa (1975), který jako příklad uvádí větu (21):

(21) Každé dítě splnilo nějaký úkol (Hlavsa, 1975).

Otázkou je zde reference výrazu *nějaký úkol*. Hlavsa (1975, s. 98) pro podobné případy navrhuje chápání těchto výrazů jako denotátorů s operátorem Var (který přibližně odpovídá typu GEN v modelu PDT), a to z důvodů hlavně formálně jazykových. Je otázkou, do jaké míry je možné při analýze reference vycházet z formálně jazykových kritérií, jisté ovšem je, že je reference výrazu *nějaký úkol* individuální a že má rysy [+vymezenost] a [+reálnost]. Z tohoto důvodu preferujeme chápání reference podobných výrazů jako individuální specifické.

Druhou nejednoznačnou oblastí jsou výrazy, jejichž designát²⁸ je tvořen omezenou množinou prvků. Např. v textu o územní samosprávě v ČR se můžeme setkat s výrazem *kraj*, do jehož designátu patří pouze 13 (resp. 14, chápeme-li jako kraj i Prahu) prvků. Pokud se v takovém textu bude vyskytovat výraz *kraje* (referující ke všem 13, resp. 14 krajům), nemusí být na první pohled jasné, jedná-li se o referenci individuální (tj. referenci k uzavřené množině objektů), nebo generickou (tj. referenci k celému druhu „kraj“). V takových případech je podle našeho názoru nutné řídit se významem příslušného predikátu.

(22) Kraje mohou navrhovat zákony.

(23) Kraje minulý týden vyhlásily povodňovou pohotovost.

Ve větě (22) predikát vypovídá o obecné možnosti krajů, přičemž tato možnost by potenciálně platila i pro kraj, který by mohl nově vzniknout (jakkoli je takový vznik v praxi nepravděpodobný). Naopak ve větě (23) predikát vypovídá o konkrétní události, a pokud by v budoucnu vznikl nějaký nový kraj, tento predikát by o něm nevypovídal. Proto ve větě (22) pokládáme referenci výrazu *kraje* za specifickou generickou, zatímco ve větě (23) za specifickou individuální.

4.2 SPECIFICKÁ GENERICKÁ REFERENCE

Z hlediska kritéria specifičnosti se jedná o referenci specifickou, tj. výrazy s tímto typem reference mají rysy [+vymezenost] a [+reálnost]. Referentem však nejsou jednotlivé objekty nebo množiny objektů, nýbrž druhy.²⁹

Otázkou nyní je, jak přesně vymežit pojem druh a tím odlišit referenci generickou od reference individuální. Především je jasné, že generičnost není totožná s univer-

²⁸ Tj. množina prvků, k nimž mohou tyto výrazy v různých kontextech referovat. Termín přebírám od Hlavsy (1975), Palek (1988) používá v odpovídajícím významu termínu *extendum*.

²⁹ Vymezenost a reálnost se zde tedy týkají celého druhu, nikoli jednotlivých jeho reprezentantů (objektů). Např. ve větě (24) výraz *kočky* referuje k jedinému, tedy vymezenému (reálnému) druhu „kočka“.

zální kvantifikací, neboť např. věta (24) je chápána jako pravdivá, přestože na světě jistě existují i kočky, které ocas z nějakého důvodu (např. úrazu) nemají.

- (24) Kočky mají ocas.
 (25) Lvi mají hřívu (GenBook; přeložil JP).

Věta (25), převzatá z GenBook (s. 48), navíc ukazuje, že objekty, o nichž se ve větě s generickým výrazem vypovídá, ani nemusí tvořit většinu reprezentantů³⁰ daného druhu, neboť v případě druhu „lev“ platí, že hřívu mají pouze dospělí samci, nikoli samice a mláďata (a věta je přesto chápána jako pravdivá).

Přehled teorií generické kvantifikace podává GenBook (s. 45–63), žádná z těchto teorií však autory není hodnocena jako uspokojivá, a proto je v této publikaci (GenBook, s. 78–85) dále navržen výčet interpretací (založených na sémantice příslušného predikátu), kterých může generická jmenná fráze nabývat. Podle našeho názoru je tento výčet v zásadě aplikovatelný i pro jazykový materiál češtiny, a proto jej bereme za základ našeho vymezení generické reference (v opozici k referenci individuální). Jednotlivé typy interpretací generických výrazů zde ilustrujeme větami (26)–(33), které ve vlastním překladu přebíráme z uvedené publikace.

- (26) *Blboun nejapný* je vyhynulý.
 (27) *Němečtí zákazníci* si loni koupili 11 000 nových BMW.
 (28) *Mladí lidé* se dívají na televizi šest hodin denně.
 (29) *Brambora* obsahuje vitamin C.
 (30) *Holandané* jsou dobří námořníci.
 (31) Na Aljašce jsme natáčeli *medvědy grizzly*.
 (32) *Člověk* poprvé stanul na Měsíci v roce 1969.
 (33) Jak postupujeme k severu, jsou *vlci* větší a větší.

Věta (26) představuje tzv. vlastní referenci k druhu, kdy příslušný predikát je možné použít pouze ve spojení s druhem jako celkem, nikoli s jednotlivými objekty. Ve větě (27) vypovídá predikát o souhrnu reprezentantů druhu (k němuž referuje výraz *němečtí zákazníci*). U typu (28) predikát vypovídá o průměrném reprezentantu příslušného druhu. Ve větě (29) je generická reference interpretována jako reference k jednotlivým reprezentantům příslušného druhu. Ve větě (30) se jedná o tzv. predikaci odlišující vlastnosti. Tento případ se na první pohled podobá typu (29), liší se však tím, že predikována je zde vlastnost, která odlišuje příslušný druh od jiného druhu a která zdaleka nemusí platit o všech typických reprezentantech příslušného druhu.³¹ Za pozornost stojí též typ (31), kde predikát vypovídá pouze o některých reprezentantech příslušného druhu (tj. v této konkrétní větě o těch medvědech, kteří byli natáčeni). K tomuto typu uvádí GenBook (s. 83), že se jednotlivé jazyky liší v míře, do

³⁰ Reprezentantem druhu rozumíme objekt či skupinu objektů, které náleží k příslušnému druhu.

³¹ Formálně je tento rozdíl mezi typy (29) a (30) zřejmě manifestován menší přijatelností singulárního výrazu u typu (30).

kteřé tento typ připouštějí. V češtině by kritériem pro odlišení tohoto typu od nespécifické individuální reference mohla být (ne)možnost použít v příslušném kontextu singulár, je však otázkou, zda je vůbec oprávněné existenci tohoto typu v češtině předpokládat.³² Typ (32) představuje tzv. průkopnickou interpretaci, kdy se predikát vztahuje pouze k prvnímu reprezentantu příslušného druhu. U posledního typu (33) příslušný predikát vyjadřuje vzájemné porovnání reprezentantů daného druhu.

Zvýšenou pozornost je nutné věnovat typu (29), tedy referenci k reprezentacím druhu. Kromě toho, že o něm předpokládáme, že se v českých textech vyskytuje ze všech uvedených typů nejčastěji, je důvodem skutečnost, že se tento typ na první pohled podobá nespécifické individuální referenci s rysem [-vymezenost], protože jej lze chápat jako referenci k otevřené množině prvků. Je tedy třeba uvést kritéria, na základě kterých je možné tento typ případů od nespécifické individuální reference odlišit.

Především je nutno upozornit, že tento typ generických výrazů se může vyskytovat pouze v tzv. generických propozicích, tj. propozicích, jejichž význam je založen na zobecnění nad situacemi nebo objekty.³³ Jak ukazují autoři GenBook (s. 23–36), je možné popisovat význam těchto propozic pomocí sémantického zápisu s využitím tzv. generalizačního operátoru GEN. Např. význam věty (29) je možné zapsat takto: GEN[x; y](IS(x, brambora); IS(y, vitamin C) & x obsahuje y).³⁴ Restriktor (tj. část před středníkem) vyjadřuje předmět zobecnění v dané propozici, v tomto případě je v restriktoru pouze proměnná x, zastupující reprezentant druhu „brambora“, což lze interpretovat tak, že daná propozice je zobecněním nad všemi v daném ohledu typickými reprezentanty druhu „brambora“.

Vzhledem k tomu, že typ (29) se vyskytuje pouze v generických propozicích a že všechny generické propozice lze zapsat pomocí uvedeného sémantického zápisu, je možné definovat podmínky pro tento typ generické reference pomocí pozice referenčního výrazu (resp. jemu odpovídající proměnné) v sémantickém zápisu: pokud se v restriktoru sémantického zápisu generické věty nachází proměnná reprezentující příslušný referenční výraz a pokud platnost této proměnné není v restriktoru nijak omezena, má tento referenční výraz generickou referenci.³⁵

Chceme dále upozornit, že typ (29), u něž je generická reference interpretována jako reference ke všem typickým reprezentantům druhu, se může vyskytovat také v adnominálních pozicích. Příkladem může být např. reference výrazu *čokolády* ve spojení *milovník čokolády*. I zde je možno generičnost reference vysvětlit pomocí sémantického zápisu s generalizačním operátorem GEN, neboť vztah mezi výrazy *milovník* a *čokolády* lze také vyjádřit pomocí generalizačního operátoru GEN: GEN[x; y](IS(x, čokoláda); IS(y, milovník čokolády) & y miluje x). Podstatné na této formalizaci

32 Jiným řešením by bylo tyto případy chápat jako referenci nespécifickou individuální.

33 Blíže o tomto typu propozic viz GenBook (s. 19–63). Stručněji o nich pojednáváme též ve své bakalářské práci (viz Pergler, 2010, s. 25–27).

34 Relace IS je definována takto: IS(x, y) \leftrightarrow (x = y R(x, y)). Tedy pokud x je v relaci IS s y, pak buď jsou x a y totožné, nebo objekt x je reprezentantem druhu y.

35 Pokud se ale proměnná reprezentující referenční výraz nachází v sémantickém zápisu mimo restriktor nebo se nachází v restriktoru, avšak je zde omezena nějakou doplňující podmínkou, pak se o generickou referenci typu (29) nejedná (o těchto případech viz dále).

je, že proměnná *x* (reprezentující čokoládu) se zde nachází v restriktoru, a proto má výraz *čokolády* v uvedeném spojení specifickou generickou referenci.

4.3 NESPECIFICKÁ INDIVIDUÁLNÍ REFERENCE

Jde o referenci individuální, referentem tedy není druh, nýbrž objekt či množina objektů. Jak jsme již uvedli výše, nespecifičnost reference je v našem pojetí implikována přítomností alespoň jednoho z rysů [-reálnost], [-vymezenost]. Zaměříme se nejprve na nespecifickou referenci s rysem [-reálnost].

Rys [-reálnost] znamená, že referent příslušného výrazu není součástí světa textu, resp. je ve světě textu přítomen pouze potenciálně. Tato potenciálnost vyplývá z příslušného predikátu, který pojmenovává eventualitu, jejíž realizace je buď podmíněná, nebo je přímo vyloučená. Predikátové sloveso tak může mít např. formu kondicionálu, imperativu, futura, může být v dosahu negace, případně může být spojeno s modálním slovesem nebo částicí apod. Podmíněnost eventualy pojmenované slovesem může být umožněna také tím, že příslušné sloveso je v neurčitém tvaru a samo představuje doplnění jiného slovesa. Příklady výrazů s rysem [-reálnost] představují věty (34)–(39):³⁶

- (34) Karel by si koupil *auto*, kdyby měl peníze.
- (35) Kup si *auto*!
- (36) Koupím si *auto*.
- (37) Nekoupil jsem si *auto*.
- (38) Musím si koupit *auto*.
- (39) Slovenská vláda vyzvala ministerstvo zahraničních věcí SR, aby [...] začalo připravovat *mezinárodní smlouvu*, která by upravila tento problém (PDT).

Za výrazy s nespecifickou individuální referencí s rysem [-reálnost] pokládáme také výrazy v generických propozicích, kterým v sémantickém zápisu odpovídá proměnná nacházející se v restriktoru operátoru GEN, jež je zde (tj. v restriktoru) omezena doplňující podmínkou. Příkladem může být věta (40), jejíž sémantický zápis je: GEN[*x*,*s* ;](IS (*x*, lev) & *x* ucítí potravu v *s*; *x* zařve v *s*).

- (40) Když lev ucítí potravu, zařve (GenBook; přeložil JP).

Proměnná *x*, reprezentující výraz *lev*, se nachází v restriktoru, avšak je zde doplněna ještě podmínkou *x* ucítí potravu v *s*, která vyjadřuje omezení platnosti zobecnění pouze na situace *s*, ve kterých lev ucítí potravu (z hlediska klasifikace vět podle GenBook se tedy jedná o generickou propozici, v níž dochází k zobecnění zároveň nad objekty a nad situacemi). Vzhledem k přítomnosti této podmínky výraz *lev* nemá

³⁶ V popsaných kontextech se samozřejmě mohou vyskytovat i jmenné fráze se specifickou referencí, např. pokud výraz *auto* v příkladových větách zaměníme za výraz *to auto*, bude se jednat o specifickou individuální referenci, neboť referentem bude konkrétní auto tvořící součást světa textu.

specifickou generickou referenci, nýbrž nespécifickou individuální referenci, neboť daná propozice nevypovídá o celém druhu „lev“, nýbrž pouze o těch exemplářích tohoto druhu, kteří se nacházejí v (hypotetických, tj. nereálných) situacích, kdy ucítí potravu.

Zaměříme se nyní na druhou skupinu výrazů s nespécifickou individuální referencí, tj. na výrazy s rysem [-vymezenost]. Jedná se o výrazy, jejichž referentem je otevřená množina objektů. „Otevřenost“ této množiny, tj. nemožnost charakterizovat množinu úplným výčtem jejích prvků, je umožněna tím, že se příslušný referenční výraz nachází v generické propozici. Jak již bylo zmíněno, generická propozice implikuje generalizaci (zobecnění), která s sebou nese generickou kvantifikaci příslušné eventuality. Generická kvantifikace je přitom vždy neurčitá (jak již bylo řečeno, je třeba ji odlišovat od kvantifikace univerzální), přičemž právě tato neurčitost způsobuje nemožnost charakterizovat příslušnou množinu výčtem všech jejích prvků (neboť počet prvků není určitý). Příkladem je výraz *auta* ve větě (19), který referuje k autům, která jsou vlastněna lidmi — počet takových aut není pevně dán, jedná se proto o otevřenou množinu a reference tohoto výrazu je tedy nespécifická.

Od výrazů se specifickou generickou referencí (které se, jak bylo zmíněno, rovněž vyskytují v generických propozicích) se výrazy s nespécifickou individuální referencí s rysem [-vymezenost] odlišují tím, že referent těchto výrazů není v těchto propozicích předmětem zobecnění. Tak např. ve větě (19) jsou předmětem zobecnění lidé, nikoli auta, propozice tedy nevypovídá o celém druhu „auto“, nýbrž pouze o některých autech (a nikoli o jiných, např. nevypovídá o těch autech, která nejsou vlastněna lidmi, nýbrž firmami). V sémantickém zápisu příslušných propozic se tato skutečnost odráží tím, že proměnná reprezentující výraz s rysem [-vymezenost] zde stojí mimo restriktor.³⁷

Výrazy s nespécifickou individuální referencí s rysem [-vymezenost] se mohou vyskytovat také v adnominálních pozicích, jako je tomu v případě výrazu *bot* ve spojení *tento výrobce bot*. Tyto případy pokládáme za nutné odlišit od výše popsaných případů typu *milovník čokolády*, neboť na rozdíl od nich tyto případy nelze popsat pomocí operátoru GEN tak, aby byla příslušná proměnná v restriktoru. Tento rozdíl je způsoben různou sémantikou řídicího substantiva a lze jej interpretovat tak, že zatímco milovník čokolády miluje všechno to, co je čokoláda (s možnými výjimkami), konkrétní výrobce bot rozhodně nevyrábí všechno to, co je bota (tj. drtivou většinu bot vyrobí někdo jiný než právě on).³⁸

V souvislosti s nespécifickou individuální referencí s rysem [-vymezenost] chceme ještě upozornit na skutečnost, že v některých případech má na referenci jednotlivých výrazů vliv aktuální členění věty, což může být chápáno jako jeden z argu-

37 Sémantický zápis věty (19) konstruujeme takto: GEN[x; y](IS(x, lidé); y je auto, x má y). Výraz *auto* je v tomto zápisu reprezentován proměnnou y, která stojí mimo restriktor.

38 Jediným případem, kdy by bylo možné uvažovat o generické referenci výrazu závisícím na substantivu typu *výrobce*, by tak byla situace, kdy by referent řídicího substantiva byl jediným existujícím výrobcem reprezentantů příslušného druhu. Taková informace se ovšem pochopitelně ve struktuře jazyka neodráží a interpretace podobných případů tedy závisí pouze na znalostech komunikantů o světě.

mentů pro situování aktuálního členění do roviny významové stavby věty. Příkladem je dvojice vět (41) a (42):

(41) Ševci prodávají boty.

(42) Boty prodávají ševci.

Při nepříznačkové intonaci obsahuje věta (41) zobecnění nad množinou ševců, věta (42) pak zobecnění nad množinou bot. Proto ve větě (41) bude v restriktoru sémantického zápisu (pouze) proměnná reprezentující výraz *ševci*, zatímco ve větě (42) (pouze) proměnná reprezentující výraz *boty*. Důsledkem je, že ve větě (41) má výraz *ševci* specifickou generickou referenci a výraz *boty* nespécifickou individuální referenci, zatímco ve větě (42) je tomu naopak. Tento rozdíl v typech reference odpovídá i intuitivnímu chápání významu obou vět.

Pro úplnost zmiňujeme ještě existenci výrazů, které mají jak rys [-reálnost], tak rys [-vymezenost]. Jsou pro ně typické výše uvedené charakteristiky obou zmíněných rysů, příkladem je již uvedená věta (20).

Závěrem této části chceme upozornit na skutečnost, že pomocí typu nespécifické individuální reference lze podle našeho názoru adekvátně popisovat všechny případy, které jsme v části 3. označili za neřešitelné či obtížně řešitelné v rámci modelu PDT. Ve větě (7) předpokládá náš model u výrazu *dceru* i u výrazu *jí* nespécifickou individuální referenci, přičemž mezi oběma výrazy je vztah koreference. Podobně je tomu i u vět (8) a (9). Toto pojetí odstraňuje všechny problémy modelu PDT, na které jsme v souvislosti s těmito případy výše upozornili.

Dále ve větách (10) a (11) budou mít výrazy *oči* nespécifickou individuální referenci, přičemž mezi nimi nebude vztah koreference (neboť se jedná o navzájem různé množiny prvků). Ve větě (12) bude mít výraz *oči obratlovců* specifickou generickou referenci (neboť se nachází v restriktoru operátoru GEN) a nebude v koreferenčním vztahu s žádným z předchozích výrazů (mezi ním a výrazem *oči* v (10) lze však uvažovat o jiném typu referenčního vztahu, totiž o vztahu exemplář druhu — druh). Otázkou je, jak chápat referenci výrazu *tyto oči*. Ve své bakalářské práci (viz Pergler, 2010, s. 50) jsme navrhovali chápat tento výraz jako jeden z předmětů zobecnění příslušné generické věty, tedy jako výraz se specifickou generickou referencí. V takovém případě by (podobně jako v případě výrazu *oči obratlovců*) tento výraz nevstupoval do koreferenčního vztahu, a jednalo by se zde tedy o nekoreferenční anaforu, jejímž antecedentem by byl výraz *oči* ve větě (11).³⁹ Domníváme se však, že lze uvažovat i o jiném řešení — je možné předpokládat, že v případě výrazu *tyto oči* nedochází k zobecnění ve výše popsaném smyslu a že tento výraz přímo referuje k množině prvků, která je dána kontextem (tj. antecedentem anaforického vztahu). V tomto případě by se zde jednalo o nespécifickou individuální referenci a výraz by byl koreferenční s výrazem *oči* ve větě (11).

³⁹ Toto řešení může na první pohled působit neobvykle, avšak nekoreferenční anafora, u níž antecedent má individuální referenci a anafor (tj. odkazující výraz) generickou referenci, je poměrně běžná. V GenBook (s. 353–354) je tento jev vysvětlován takto: „[A]ny occurrence of a CN [= common noun; JP] in discourse would simultaneously introduce the kind in the main discourse representation structure, thus [...] making it always accessible for anaphora.“

Konečně ve větách (13) a (15) budou mít výrazy *stůl* a *něj* nespécifickou individuální referenci, přičemž obě zájmena budou ve vztahu koreference se svými antecedenty (v rámci příslušné věty). Referenční výrazy typu *stůl, který je velký* ve větě (14) nadále pokládáme za generické.

4.4 NESPECIFICKÁ GENERICKÁ REFERENCE

Čtvrtým typem reference, který logicky vzniká kombinací aplikace kritéria generičnosti a kritéria specifičnosti, je nespécifická generická reference. Uvědomujeme si, že v běžných textech bude tento typ pravděpodobně zastoupen výrazně méně než ostatní tři typy, přesto však pokládáme za užitečné s ním při analýze vztahů reference pracovat.

Přítomností rysu generičnosti je tento typ reference podobný specifičné generické referenci. Liší se však od ní tím, že reference je nespécifická, tedy příslušný výraz musí mít aspoň jeden z rysů [-reálnost], [-vymezenost]. Tyto rysy se však na rozdíl od nespécifické individuální reference netýkají individuálních objektů, nýbrž druhů (ve výše popsaném pojetí). Nespécifickou generickou referenci mají tedy ty výrazy, jež referují k druhům, které nejsou reálnou součástí světa textu nebo které netvoří uzavřenou množinu (množinu druhů).

Podmínky, za nichž můžeme rozhodnout o generičnosti a o nespécifčnosti reference, byly podrobně popsány výše, zde proto uvedeme jen několik příkladů výrazů s tímto typem reference.

(43) Chtěl bych objevit *nějaké nové zvíře*.

(44) Lidé obvykle mají *své oblíbené jídlo*.

(45) Lidé by chtěli objevit *nějakou novou rostlinu*.

Výraz *nějaké nové zvíře* ve větě (43) je příkladem nespécifické generické reference s rysem [-reálnost] — jeho reference je generická, neboť výraz je použit ve významu „nový druh zvířete“, a zároveň nereálná vzhledem k modálnímu slovesu *chtít*. Výraz *své oblíbené jídlo* ve větě (44) má rys [-vymezenost], a to vzhledem ke generičnosti příslušné propozice, kde jsou předmětem zobecnění lidé (nikoli jídla). Konečně výraz *nějakou novou rostlinu* ve větě (45) má jak rys [-reálnost], tak rys [-vymezenost] (na rozdíl od věty (43) zde jde o generickou propozici, kde druhy rostlin nejsou předmětem zobecnění).

5. KOREFERENCE

Na základě popsané klasifikace reference je možné provést také klasifikaci vztahů koreference. Na rozdíl od přístupu PDT, kde je typ koreferenčního vztahu definován pouze na základě typu reference druhého z členů koreferenčního páru, se však domníváme, že typ koreferenčního vztahu je třeba určovat na základě typu reference obou členů tohoto páru. Otázkou tedy je, jaké dvojice výrazů (z hlediska typu reference) se mohou účastnit vztahu koreference.

Za samozřejmou pokládáme skutečnost, že se na koreferenčním vztahu mohou podílet dvojice výrazů, které mají navzájem stejný typ reference, těmto případům

se tedy nebudeme podrobněji věnovat. Za pozornost naopak stojí ty případy koreference, u kterých se příslušné výrazy co do typu reference liší. Ukážeme nyní, že se tyto výrazy mohou lišit jak co do specifičnosti reference, tak co do její generičnosti.

Pokud jde o specifičnost reference, domníváme se, že koreferenční vztah nemůže existovat mezi dvěma výrazy, z nichž jeden má referenci specifickou a jeden referenci nespécifickou s rysem [-vymezenost]. Důvodem je, že jeden z výrazů referuje k uzavřené množině a druhý k množině otevřené, přičemž uzavřená a otevřená množina nemohou být totožné.

V našem pojetí je však možný koreferenční vztah mezi výrazem se specifickou referencí a výrazem s nespécifickou referencí s rysem [-reálnost] (a zároveň [+vymezenost]). Koreferenci předpokládáme tehdy, pokud by při případné realizaci hypotetické eventuality, jejíž hypotetičnost zapříčiňuje rys [-reálnost] u jednoho z výrazů, došlo u obou výrazů k identitě referentů. Příkladem mohou být výrazy *projekt* a *projektu* ve větách (46)–(48):

- (46) Museli vyhotovit architektonickou studii stavby a *projekt* na její využití (PDT).
- (47) To *squatteři* již jako nadace Ladronka učinili (PDT).
- (48) Stavební část *projektu* prý pro ně vypracovali studenti 3. ročníku Fakulty architektury ČVUT (PDT).

Výraz *projekt* ve větě (46) má rys [-reálnost], což je umožněno skutečností, že eventualita pojmenovaná příslušným predikátem je chápána pouze jako možnost. Z dalších vět však vyplývá, že tato eventualita již byla ve světě textu realizována, a proto má výraz *projektu* ve větě (48) specifickou individuální referenci. Oba výrazy jsou navzájem koreferenční, neboť v případě, že by byl mutační děj vyjádřený predikátem ve větě (46) ve světě textu realizován, tj. k vyhotovení projektu by skutečně došlo, byl by referent výrazu *projekt* identický s referentem výrazu *projektu*.

Vztah koreference je možný také mezi výrazy lišícími se co do generičnosti reference. Je to možné z toho důvodu, že některé výrazy mohou referovat k druhu nikoli kvůli typu reference, nýbrž díky svému lexikálnímu významu. Jedná se např. o výrazy *druh*, *typ* apod. V GenBook (s. 6) se předpokládá, že reference těchto výrazů je vždy generická. Toto řešení však pokládáme za problematické, neboť podle našeho názoru se i u těchto výrazů může opozice individuálnost — generičnost realizovat, jak dokládají věty (49) a (50):

- (49) *Tento druh lva* žije v Africe (specifická individuální reference).
- (50) Zoologie studuje *druhy zvířat* (specifická generická reference).

Předpokládáme proto, že v případech podobných větě (49) mají tyto výrazy individuální referenci. Je tedy zřejmé, kdy se může realizovat vztah koreference mezi výrazy lišícími se typem reference (z hlediska kritéria generičnosti): je tomu tak tehdy, pokud výraz s individuální referencí referuje k druhu díky svému lexikálnímu významu, přičemž tento druh je shodný s druhem, k němuž referuje výraz s referencí generickou. Příkladem jsou koreferenční výrazy *snovačka jedovatá* a *tento druh* ve větách (51) a (52):

(51) *Snovačka jedovatá* je velmi nebezpečné zvíře.

(52) *Tento druh* ale u nás nežije.

Lze tedy shrnout, že koreferenční vztah může existovat jak mezi výrazy se stejným typem reference, tak mezi výrazy, které se v typu reference liší (v kritériu specifičnosti i v kritériu generičnosti). Jeden z recenzentů v této souvislosti vyslovuje pochybnost, zda je námi navrhovaná klasifikace referenčních typů pro popis koreference užitečná, když možnosti případné koreference nijak neomezuje. Domníváme se, že užitečná je — abychom mohli popisovat koreferenci (ve výše vymezeném smyslu referenční identity), musíme si nejprve ujasnit, k čemu jednotlivé výrazy v textu referují, přičemž k tomu, abychom to mohli zjistit, musíme vědět, o jaký typ reference se v tom kterém případě jedná. Pokud si typ reference a referenty jednotlivých výrazů v textu ujasníme, můžeme tak odhalit nejen případy koreference, nýbrž také to, že v některých případech, např. v příkladu (53), který uvádí zmíněný recenzent, se o koreferenci ve výše popsaném smyslu ve skutečnosti nejedná.⁴⁰

(53) *Pět blbounů nejapných* sedělo na skále. Přemýšleli o svém vyhynutí.

6. ZÁVĚR

V článku jsme provedli základní vymezení pojmu koreference a nejdůležitějších vlastností této relace. Zároveň jsme ukázali, že koreferenci je nutné důsledně odlišovat od vztahů vnitrotextového odkazování v textu (endofory), neboť tyto jevy se v textech mohou vyskytovat nezávisle na sobě. Podali jsme stručný přehled modelu reference a koreference užívaného v Pražském závislostním korpusu a na základě rozboru případů, které tento model nedokáže adekvátně popisovat, jsme dospěli k potřebě vymezit podrobnější klasifikaci referenčních vztahů. Takovou klasifikaci jsme dále navrhli: vycházeli jsme ze dvou nezávislých kritérií (specifičnosti a generičnosti) a dospěli jsme tak ke čtyřem typům reference (specifické individuální, specifické generické, nespecifické individuální a nespecifické generické).⁴¹ Nakonec jsme ukázali, jaké dvojice výrazů (z hlediska této klasifikace) se mohou účastnit koreferenčních vztahů v textu.

Jak již bylo zmíněno, naše klasifikace reference není klasifikací vyčerpávající, neboť bere v úvahu pouze ty charakteristiky reference, které jsou relevantní pro popis koreference (pojímané jako referenční identita). Z příkladů, které jsme uváděli, je dále zřejmé, že jsme se soustředili především na referenci výrazů se substantivní platností a konkrétním významem. Předpokládáme, že navrženou klasifikaci je možno uplatnit i na substantivní výrazy s abstraktním významem a významem

⁴⁰ Výraz *pět blbounů nejapných* má specifickou individuální referenci a referuje k pěti exemplářům, anaforické zájmeno *svůj* má specifickou generickou referenci a odkazuje k celému druhu. Vzhledem k tomu, že oba výrazy nemají stejný referent, zde koreference není, a jedná se tak o nekoreferenční anaforu (srov. výše oddíl 1.).

⁴¹ Uplatnění této klasifikace při empirickém rozboru několika konkrétních textů jsme ukázali ve své bakalářské práci (viz Pergler, 2010, s. 51–57).

dějovým. Reference takových výrazů však musí být předmětem podrobnějšího výzkumu. Kromě toho je třeba věnovat pozornost také výrazům s jinou slovnědruhovou charakteristikou, a to především adjektivům a slovesům. Jak jsme již uvedli, v modelu PDT jsou tyto výrazy za jistých podmínek také chápány jako součást koreferenčních řetězců, tyto podmínky však nejsou dostatečně explicitně popsány. Třetím tématem, na které chceme v závěru upozornit, jsou další referenční vztahy v textu. Vymezili jsme koreferenci jako vztah mezi výrazy, mezi jejichž referenty existuje identita. Bylo by však možné pracovat i s dalšími vztahy, které by byly založeny na vzájemném poměru referentů.⁴² V souvislosti s navrženou klasifikací referenčních typů by bylo vhodné uvažovat např. o relaci reprezentace, tedy vztahu mezi výrazy, z nichž jeden označuje druh a druhý exemplář tohoto druhu (v libovolném pořadí).⁴³ Za zvláště zajímavé pak v této souvislosti pokládáme téma nekoreferenční anafory, kde by bylo vhodné zkoumat, za jakých podmínek⁴⁴ se nekoreferenční anafora může vyskytovat.

LITERATURA:

- ADAMEC, Přemysl (1980): K vyjadřování referenční určenosti v češtině a v ruštině. *Slovo a slovesnost*, 41(4), s. 257–264.
- GenBook: CARLSON, Gregory N. — PELLETIER, Francis Jeffrey (eds.) (1995): *The Generic Book*. Chicago, IL — London: University of Chicago Press.
- HAJIČOVÁ, Eva — OLIVA, Karel — SGALL, Petr (1987): Odkazování v gramatice a v textu. *Slovo a slovesnost*, 48(3), s. 199–212.
- HLAVSA, Zdeněk (1975): *Denotace objektu a její prostředky v současné češtině*. Praha: Academia.
- MIKULOVÁ, Marie — BÉMOVÁ, Alevtina — HAJIČ, Jan — HAJIČOVÁ, Eva — HAVELKA, Jiří — KOLÁŘOVÁ, Veronika — KUČOVÁ, Lucie — LOPATKOVÁ, Markéta — PAJAS, Petr — PANEVOVÁ, Jarmila — ŠEVČÍKOVÁ, Magda — SGALL, Petr — ŠTĚPÁNEK, Jan — UREŠOVÁ, Zdeňka — VESELÁ, Kateřina — ŽABOKRTSKÝ, Zdeněk (2005): *Anotace na tektogramatické rovině Pražského závislostního korpusu: anotátorská příručka* [Technická zpráva ÚFAL MMF UK v Praze, 31]. Praha: MFF UK v Praze.
- NĚDOLUŽKO, Anna (2011): *Rozšířená textová koreference a asociální anafora: koncepce anotace českých dat v Pražském závislostním korpusu*. Praha: ÚFAL MMF UK v Praze.
- PALEK, Bohumil (1988): *Referenční výstavba textu*. Praha: UK v Praze.
- PERGLER, Jiří (2010): *Koreferenční řetězce s nespécifickou a generickou referencí v češtině* [nepublikovaná bakalářská práce]. Praha: ÚČJTK FF UK v Praze.
- RACHILINA, Jekaterina Vladimirovna — KREJDLIN, Grigorij Jefimovič (1981): Denotativnyj status otgлагоl'nych iměn. *Naučno-techničeskaja informacija*, 12, s. 17–22.

Jiří Pergler | Ústav českého jazyka a teorie komunikace FF UK v Praze

<ji.perg@seznam.cz>

42 O některých takových vztazích pojednává např. Palek (1988). Zmiňuje v této souvislosti vztahy inkluze, členství (v množině) a rozdílu.

43 Srov. příklad (53).

44 Tj. např. jaký referenční vztah (jiný než koreference) se může vyskytovat mezi anaforem a jeho antecedentem.