

PROČ A JAK ZKOUMAT SOUKROMÉ DOPISY¹

Anna Řehořková

HLADKÁ, Zdeňka et al. (2013): *Soukromá korespondence jako lingvistický pramen*. Brno: Masarykova univerzita, 188 s.

HLADKÁ, Zdeňka — MARTINCOVÁ, Olga (2012): *Slova v soukromých dopisech: Lexikografická sonda*. Brno: Masarykova univerzita, 139 s.

Soukromé dopisy se stávají předmětem vědeckého zájmu především jako historické dokumenty. I v lingvistice jsou soubory těchto textů zřejmě nejčastěji využívány ke zjišťování staršího vývojového stavu jazyka, k historické sociolingvistické analýze apod., srov. např. Nevalainenová a Raumolin-Brunbergová (1996) a další studie z korpusové řady *Corpora of Early English Correspondence*, Elspass (2012) a literaturu tam uvedenou nebo výzkum korespondence Boženy Němcové ad. Ojedinelým počinem je v tomto ohledu projekt Zdeňky Hladké et al., Korpus soukromé korespondence (KSK), zachycující 2000 ručně psaných soukromých dopisů naší současnosti (1990–2005) a cílený na výzkum dnešní češtiny.² Náměty k tomuto výzkumu a jeho dosavadní výsledky prezentují již tři knižní publikace: po brožuře z roku 2005, která doprovázela kompletní vydání KSK na CD a vedle technických informací shrnovala i prvotní studie vzniklé při přípravě korpusu, vyšly v letech 2012–2013 monografie *Soukromá korespondence jako lingvistický pramen* a slovník *Slova v soukromých dopisech: Lexikografická sonda*.

Monografie byla avizována již v roce 2005 jako chystané systematictější a detailnější pojednání o možnostech, jak ze shromážděných dopisů čerpat různé lingvistické informace (lexikologické, stylistické aj.). Výsledné kolektivní dílo si však klade trochu jiný cíl: představit na jednom místě různě koncipované výzkumy soukromé korespondence vůbec (nejen současné) a ukázat především jejich metodologii. Podává přehled o třech výrazných českých projektech z nedávné doby zabývajících se epistolárními texty: zpracování a vytěžování KSK (1.–4. kapitola) a o dvou výzkumech korespondence známých osobností: Bedřicha Smetany a Boženy Němcové (5. a 6. kapitola). Krátce shrnuje také tradiční stylistickou reflexi dopisu (7. kapitola). Oproti pracím publikovaným v rámci těchto projektů jinde přidávají kapitoly vždy určité novum: ať jde o další téma z vyšší jazykové roviny (Marek Nekula — Lucie Rychnovská: Jazyková biografie Bedřicha Smetany a jazyk jeho korespondence se zřetelem

1 Text vznikl v rámci projektu GA ČR č. P406/12/0691 Korespondence Karla Havlíčka.

2 Dopisová část korpusu KSK (zahrnujícího v úplném vydání také e-maily a SMS) o zmíněných 2000 textech je od roku 2006 specializovanou, dále neměněnou složkou Českého národního korpusu (ksk-dopisy, zde dále KSK). Soukromé dopisy jsou tak v rámci ČNK zpřístupněny v nebyvalém množství: např. Britský národní korpus obsahuje 249 dokumentů z běžnědorozumívací komunikační oblasti (vedle dopisů i vzkazy, zápisky apod.), Národní korpus ruského jazyka 1405 soukromých písemností včetně deníků aj., o samostatném subkorpusu ani nemluvě. Sběr a zpracovávání soukromých dopisů pod vedením Zdeňky Hladké nadále pokračuje, a to směrem do minulosti (srov. Hladká et al., 2013a).

ke gramatickým jevům), rozšířený materiál a podrobněji probranou metodologickou stránku (Zdeňka Hladká: Teritoriálně a sociálně podmíněné diference v jazyce současné korespondence), nebo o vývojový pohled na témata a výsledky dlouhodobého výzkumu (Robert Adam et al.: Dopisy jako výzva (Jazykovědný výzkum korespondence Boženy Němcové)).

Lze jen zalitovat, že více prostoru není dáno právě studiím založeným na KSK. Dvě úvodní kapitoly (Zdeňka Hladká: Korpusové zpracování soukromé české korespondence v ÚJČ FF MU v Brně; Dana Hlaváčková: Korpusové zpracování korespondenčních textů: morfologické značkování) se zaměřují „pouze“ na koncepci a technické pozadí tvorby korpusu a do jisté míry reformulují text starší publikace.³ Přitom je to právě práce s korpusem, od níž by se dal očekávat nově koncipovaný výzkum (např. s využitím korpusů řady SYN jako srovnávacího pozadí).

Kvantitativní analýze materiálu KSK je věnována třetí a čtvrtá kapitola (Jana Machalová — Klára Osolsobě: Hypokoristika z rodných jmen v Korpusu soukromé korespondence; Zdeňka Hladká: Teritoriálně a sociálně podmíněné diference v jazyce současné korespondence). Autorky Jana Machalová a Klára Osolsobě hledají v rozmanitých způsobech tvoření domáckých podob rodných jmen jisté zákonitosti v podobě preferovaných slovtvorných postupů. Na základě seznamu hypokoristik již dříve excerpovaných z KSK vytvářejí (v návaznosti na práce Miloslavy Knappové a Miloše Dokulila) retrográdní morfematický slovník (přiložen ke kapitole). Podle kombinace typů sufixů s určitými typy slovtvorných základů pak rozdělují slovtvorné modely uplatňované při vytváření hypokoristik do tří skupin — derivace typicky deminutivními sufixy, zkracování a derivace typicky hypokoristickými, parasystémovými sufixy (např. *-as*) a následně kvantifikují poměr jejich zastoupení. Mimoto podávají např. podrobné srovnání předností a nevýhod korpusových dat s dotazníkovými a krátce pojednávají o vybraných subtématech (frekvenci parasystémových sufixů podle charakteristického konsonantu, posunům v gramatickém rodě hypokoristik aj.).

Početní výsledek, který by měl ukázat na potenciální slovtvorné preference, nemá být, jak autorky deklarují, vlastním přínosem kapitoly („frekvenční analýzy je [...] třeba brát s jistou rezervou“; Hladká et al., 2013, s. 37), záměrem je představit samotný výzkumný proces a jeho problémy jako informaci pro další badatele. Čtenář se tak dozví, že při případné slovtvorné analýze hypokoristik v KSK bude muset kvůli nedostatku mimojazykového kontextu řešit problém nejednoznačných slovtvorných základů (*Anda: Andrea vs. Anna*, tedy *And-a vs. An-da*, tedy typ zkracování vs. typ odvozování typicky hypokoristickým sufixem) a v přiloženém retrográdním slovníku zjistí, že se autorky rozhodly pro jednu z variant (*And-a*) a tu započítaly do frekvenčního přehledu. Má k dispozici i údaj o tom, že tyto nejednoznačné pří-

3 Pro získání praktických informací k užívání korpusu má smysl číst především zmiňovanou starší příručku (Hladká et al., 2005), novější kapitoly se zaměřují spíše na předání obecných zkušeností se zpracováváním textů pro korpus. Mnohé z těchto poznatků ovšem nejsou z praktického hlediska bez zajímavosti (např. doporučení upustit od morfologického značkování malých korpusů s nestandardní podobou jazyka pro přílišnou časovou náročnost nebo myšlenka ponechávat u obtížně interpretovatelných slovních tvarů více morfologických značek).

padu tvoří necelých 20 % hypokoristik, a proto vyvstává otázka, zda nemohly autorky uvést také alternativní výsledky bez započítaných sporných případů (nemusely by pak svou analýzu omlouvat). Navíc i když nejsou výsledky definitivní, měly by být dostatečně okomentovány a interpretovány. Jestliže autorky napíšou, že se „derivate typicky deminutivními sufixy *-k(a)*, *-(e)k* a *-ík* [...] uplatňuje u 43 % doložených formací“ (Hladká et al., 2013, s. 43), týká se tento výsledek jen hypokoristik, v nichž byl sufix jediným odvozovacím formantem, nebo i těch, v nichž se kombinoval s jinými (např. *-ás-ek*)? A pokud je typ odvozování deminutivními sufixy zastoupen ze 43 %, typ zkracování z 34 % a typ odvozování typicky hypokoristickými formanty z 23 %, znamená to, že žádný z nich není významně více nebo méně preferován než ostatní, nebo že jsou rozdíly podstatné?

Jiný kvantitativně orientovaný výzkum založený na KSK představila Zdeňka Hladká (4. kapitola). Autorka se podobně jako v některých předchozích studiích (např. Hladká — Šindlerová, 2004) zaměřila (1) na šíření vybraných prvků obecné češtiny za hranice českých nářečí v užším smyslu a (2) na teritoriální diference v míře užití spisovné češtiny v korespondenci. V rámci prvního tématu sledovala především relativní frekvence (procento reálných výskytů ze všech potenciálních) hláskových jevů $\acute{y} > ej$, $\acute{e} > í$ a protetického *v-* a některých morfologických typů (*dobrý kluci*; *prosej*, *sázej*, *dělaj*; *ved*, *nes*). Při zpracování druhého tématu vychází z části ze zjištěných frekvencí (případů neprovedení přehlásky $\acute{a} > \acute{e}$, $\acute{u} > i$, dále tzv. moravského kráčení, slovesných koncovek *-ijou*, *-ajou* a několika nářečních lexémů), z části pouze z prototypických ukázek (uvádí např. doklady naznačující souvislost mezi stylovým příznakem lexika a nespisovnou podobou: *těžký prachy* vs. *logické a přirozené systémy*).

Cílem kapitoly je přispět novými, čísly podloženými argumenty do neutuchající diskuse o užívání obecné a spisovné češtiny. V otázce šíření obecněčeských jevů dospívá autorka k závěru, že expanduje pouze diftong *ej*, a to jen do oblasti střední Moravy. Důležité (a v závěru neuvedené) jsou však podrobnosti vyplývající z předchozích částí kapitoly: úžení $\acute{e} > í$ je na střední Moravě nářečně náležité, a proto nelze na tomto území expanzi ani předpokládat; protetické *v-* proniká do dopisů velmi zřídka i v Čechách, a tudíž v tomto případě nehraje roli jen „omezující působení nářečního podloží“ (Hladká et al., 2013, s. 81), ale i faktor psanosti, který autorka zohledňuje až ve druhé části studie.

Problematická je na několika místech kapitoly i interpretace kvantitativních výsledků, neboť není testována významnost zjištěných rozdílů. Např. výskyt diftongu *ej* v koncové pozici je v českých a středomoravských dopisech zjevně vyrovnaný (Č 30 %, SM 32 %), ale pokud v koncovce následuje ještě konsonant (*-ejch*), jsou výsledky méně jasné a vyžadují ověření (Č 9 %, SM 5 %). Z podobného důvodu je nejistý výklad vlivu věku, vzdělání a pohlaví na výskyt diftongu *ej*, uvedený v Doplňku. Autorka srovnává hodnoty získané z korespondence z Prahy a Brna s hodnotami zjištěnými Françoisem Esvanem v Pražském a Brněnském mluveném korpusu (Čermák — Blatná, 2006) a tvrdí, že faktor pohlaví působí na výskyt diftongu v dopisech opačně než v mluvené řeči (v korespondenci je *ej* častější u žen než u mužů). Autorčiny výsledky byly sice normalizovány výpočtem tzv. korigované frekvence, navrženým právě Esvanem, nebyla však už ověřena signifikance rozdílů mezi muži a ženami, mladšími a staršími atd. (Esvan pracuje s většími daty a používá test chí-kvadrát).

V otázce míry užití spisovné češtiny v korespondenci autorka konstatuje, že v dopisech z Moravy a Slezska je tíhnutí psané formy ke spisovnosti silnější než v Čechách. Tento závěr opírá o fakt, že výskyt sledovaných celomoravských jevů (např. neprovedení přehlásky 'u > i) je výrazně omezen na konkrétní lexémy a to snižuje celkovou frekvenci daných jevů. V případě středomoravské oblasti však autorka nebere v potaz předchozí poznatek o průniku $\acute{y} > ej$ a $\acute{e} > í$ v míře shodné s českými dopisy (východomoravská a slezská oblast se v tomto od středomoravské výrazně liší). Za pozornost stojí také poznámka, že škála dalších moravských interdialektických i místních jevů v dopisech je velmi široká (to ovšem pravděpodobně platí i pro českou oblast). Zjištění o obecněčeských jevech v dopisech z Čech měla autorka možnost porovnat s výzkumem mluvené řeči Concetty Maglione (2003) a vyjádřit, o kolik procent frekvence nespisovných jevů v psaných textech klesla (otázku srovnatelnosti dat nekomentuje). V případě moravských a slezských dopisů údaje pro srovnání k dispozici nebyly, ale Zdeňka Hladká přesto tvrdí, že pokles je větší (Hladká et al., 2013, s. 85–86). Cennější než obecné závěry jsou tak na kapitole konkrétní fakta, např. výrazně nízká frekvence všech nespisovných jevů (např. i tvaru *bysme*) v dopisech ze Slezska.

Další tři kapitoly představují kvalitativně orientované výzkumy soukromých dopisů, a to buď korespondence známých osobností (kromě dvou výše zmíněných kapitol částečně i Jana Hoffmannová: Soukromý dopis: mezi psaným textem a mluveným dialogem — v oddíle věnovaném dopisům Jana Nerudy a Karla Čapka ženám), nebo textů anonymních pisatelů v KSK (Jana Hoffmannová v další samostatné části).⁴ Srovnání prvních dvou kapitol dobře ukazuje nosnost rozlišování různých úhlů pohledu na dopis při výzkumu korespondence Boženy Němcové (dobový dokument vs. textová projekce komunikace, v níž se projevuje např. stylizace produktora do různých rolí), na něž odkazuje i sociolingvisticky, nikoli literárněvědně inspirovaný výzkum korespondence Bedřicha Smetany (v rámci rekonstrukce jazykové biografie⁵ je rozlišena životní, subjektivní a textová rovina metajazykových výpovědí). Zaměření na průměrného uživatele jazyka své doby, jakým Bedřich Smetana byl, navíc umožňuje nejen identifikovat jeho subjektivní strategie (např. ideologické přehodnocení vlastního původně neutrálně vnímaného německého vzdělání na násilnou germanizaci), ale také získat z jazykové biografie (zahrnující např. i tehdejší jazyková nařízení a další informace z různých zdrojů) údaje obecněji platné pro dobovou jazykovou situaci (např. o přítomnosti češtiny v německém školství).

Zobecnitelnost poznatků kvalitativního rázu na základě typičnosti zkoumaného subjektu a konfrontace s dalšími dostupnými informacemi, deklarovaná v kapitole Marka Nekuly a Lucie Rychnovské, jistě není pro podobně orientované lingvisty novinkou, na studii je však třeba ocenit reflexi vlastních noetických východisek a systematické užívání pojmových nástrojů (např. teorie jazykového managementu). Kvali-

4 Toto rozčlenění sleduje vyváženost materiálového východiska (KSK vs. jiné) výzkumných kapitol v publikaci.

5 Metoda vychází původně z narativních interview s mluvčími o jejich jazykových zkušenostech v různých komunikačních situacích a nyní je aplikována na historickou korespondenci (deníky ad.) jako na zdroj podobných metajazykových dat.

tativní přístup, který bez kvantitativního doplnění naopak mnoho nového nepřináší, reprezentuje poslední oddíl kapitoly Jany Hoffmannové, v němž jsou známé obecné rysy mluvenostní syntaxe pouze doloženy na novém materiálu (tato část ovšem nabízí zajímavé podněty k dalším výzkumům, např. tázacích dovětků v češtině nebo významově vyprázdněného mezivýpovědního konektoru *ale*).

Jako samostatnou kapitolu monografie lze s nadsázkou chápat i publikaci *Slova v soukromých dopisech: Lexikografická sonda* autorek Zdeňky Hladké a Olgy Martincové (2012). (Obě práce vznikaly ve stejné době a navzájem na sebe odkazují.) Diferenčně pojatý slovník využívá materiálu KSK (konkrétně tematicky homogenního souboru 100 dopisů od pisatelů mladších 30 let z celé ČR) jako zdroje v různém smyslu příznakové slovní zásoby současné češtiny (nářeční, expresivní, knižní aj.). Více než 800 takto získaných hesel obsažených v heslové části doplňuje mj. samostatně zpracovaný registr frazémů, seznam dopisů se stručnými obsahovými charakteristikami a údaji o pisatelích či fotokopie vybraných dopisů.

Cílem lexikografické sondy je, jak se zdá, především nabídnout spolehlivě doložený lexikální materiál tvůrcům nového výkladového slovníku češtiny. Tomu odpovídá jak široké zaměření na příznakové lexikum (nikoli na slova příznačná právě pro dopisy), tak i způsob zpracování hesel, podrobně komentovaný v úvodních statích. Pokud je slovo zachyceno v některém z existujících slovníků (zejm. v SSJČ a SSČ; v slovnících neologismů ad.), opírá se definice významu o výklad tam uvedený a akcentuje ten významový rys, který je aktivován v kontextu určitého dopisu (např. *fandit nedok. (komu) hovor. s velkým (SSČ: náruživým) zájmem stranit, přát vítězství (SSČ: stranit určitému sportu, mužstvu ap.): [...] Škoda, že nevyhrál nejlepší film Samotáři. Těm jsem fandila.*). Zásada přizpůsobovat definici aktuálnímu užití slova přináší užitečné a materiálově podložené informace o stálosti či posunech jeho významu.⁶

Málo výhodné je naopak používání tradičních tzv. stylových příznaků (srov. Homoláč — Mrázková, 2014). Autorky shledávají samo vymezení příznakovosti jako dosud otevřenou otázku a řídí se de facto možnostmi přiřadit slovo k některé ze zavedených skupin.⁷ Na systematický postup při klasifikaci lexémů zřejmě rezignují: konstatují pouze, že jsou si vědomy neurčitosti spojené s podáváním příznaků pomocí kvalifikátorů, a uvádějí jediný příklad svého postupu, hesla *čígo* a *dějzl*, 'dějepis' se značením slang. (zdůvodněném v prvním případě formou, ve druhém navíc skupinou uživatelů). Ačkoli autorky věnují úsilí tomu, aby aktuálnost uváděných kvalifikátorů ověřily v dalších zdrojích (KSK, ČNK, internet), nezbyvá čtenáři, kterému chybí bližší informace o postupu značení, než hádat, proč je *debil* expr., ale *střevo* vulg. (obojí jako

6 Pokud však jde o definování dosud nezachycených slov mladých lidí, nemusí dopisy poskytovat dostatečný kontext k přesnému vymezení jejich významu: např. *hulit*, (silně) kouřit může v citátu znamenat i ,omámit se drogou, zdrogovat se, zfetovat se' („jsem zkouřen hašíšem“) je pojat příliš široce — ne všechny drogy, např. LSD, je obvyklé kouřit.

7 Výrazy obecněčeské, oblastní (specifikované — slez. — i nespecifikované), nářeční, slangové, hovorové, knižní; expresivní výrazy (s kladným citovým zabarvením, hanlivá, zhrubělá, vulgární); zastarávající a zastaralé; dobové (tj. příznačné pro současný životní styl, např. *CD*) a frekventované (v mluvě mládeže).

nadávkou), *zanesený* (o lidském vztahu) poněk. kniž., ale *zplodit* (napsat) pouze expr., *písemka* (hovor.), ale *seminárka* (slang. (kvůli odlišné frekvenci v SYN?) ap.

Navzdory některým zjištěným problémům je závěrem třeba zdůraznit, že slovník i výše probíraná monografie poukázaly na mnoho zajímavých aspektů soukromé korespondence a představily široký záběr výzkumu v této oblasti, na který lze dále navazovat. Dopisy zůstávají díky své vnitřní dynamice otevřenou výzvou jak pro statisticky poučenou kvantitativní analýzu, tak pro konceptuální nástroje kvalitativně orientovaných přístupů.

LITERATURA:

- ČERMÁK, František — BLATNÁ, Renata (eds.) (2006): *Korpusová lingvistika — stav a modelové přístupy*. Praha: Nakladatelství Lidové noviny.
- ELSPASS, Stephan (2012): The use of private letters and diaries in sociolinguistic investigation. In: Juan Manuel Hernández-Campoy — Juan Camilo Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, s. 156–169.
- HLADKÁ, Zdeňka et al. (2005): *Čeština v současné soukromé korespondenci: Dopisy, e-maily, SMS*. Brno: Masarykova univerzita.
- HLADKÁ, Zdeňka et al. (2013a): *111 let českého dopisu v korpusovém zpracování*. Brno: Masarykova univerzita.
- HLADKÁ, Zdeňka — ŠINDLEROVÁ, Hana (2004): Jakou češtinou si dopisujeme na Moravě. In: Jiří Fiala (ed.), *AUPO, Fac. Phil., Moravica*, 1. Olomouc: Univerzita Palackého v Olomouci, s. 105–114.
- HOMOLÁČ, Jiří — MRÁZKOVÁ, Kamila (2014): K stylistickému hodnocení jazykových prostředků, zvláště lexikálních. *Slovo a slovesnost*, 75 (1), s. 3–38.
- NEVALAINEN, Terttu — RAUMOLIN-BRUNBERG, Helena (eds.) (1996): *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*. Amsterdam — Atlanta, GA: Rodopi.

Anna Řehořková | Ústav Českého národního korpusu FF UK v Praze
 <anna.rehorkova@ff.cuni.cz>