

Ondřej Tichý, Digitization of Old and Middle English Dictionaries.

Diplomová práce, Ústav anglistiky a amerikanistiky FF UK, Praha, květen 2007, 100 s.

(Posudek vedoucího diplomové práce)

Diplomová práce Ondřeje Tichého podrobně mapuje jednotlivé kroky v rámci projektu převodu největšího existujícího „klasického“ staroanglického slovníku – *An Anglo-Saxon Dictionary* autorů Bosworthe a Tollera (Oxford University Press 1838-1972) – do elektronické podoby. Tento projekt probíhal pod záštitou ÚAA a za finanční pomoci Vzdělávací nadace Jana Husa v uplynulých dvou letech a Ondřej Tichý byl jeho iniciátorem a hlavním technologem i koordinátorem. Na předložené diplomové práci je třeba ocenit především její praktickou využitelnost (jakkoli se to s ohledem na druh slovníku může zdát paradoxní), netradičně mezioborový charakter (diplomand uplatnil nejen své vzdělání lingvistické, ale využil i svých znalostí počítačového odborníka), promyšlenost i pracnost (výběru a digitalizaci uvedeného slovníku předcházela skutečná *spadework*, soustavná analýza *všech* podstatných lexikografických zdrojů věnovaných staré, resp. střední angličtině).

Praktická využitelnost výsledku Tichého diplomového projektu je dána již problematickým charakterem existujících lexikografických zdrojů. Při práci totiž běžný uživatel naráží zejména na mezerovitost (u slovníků zaměřených na část období, jež by podle svého označení a záměru měly mapovat v relativní úplnosti – to je případ např. Stratmannova středoanglického slovníku, orientovaného především na ranou fázi tohoto období), na zastaralost (jde vesměs o produkty druhé poloviny předminulého a první poloviny minulého století), na nepřístupnost (v důsledku vysoké ceny, a tedy také malého počtu vydání), na malou praktickou použitelnost (případ „klasických“ slovníků, které slouží zejména k lineárnímu překládání), a konečně na neúplnost či nedokončenost (velký Kurathův a Kuhnův středoanglický slovník, dílo desítek specialistů, byl po mnoha letech dokončen teprve nedávno a elektrizován, resp. zpřístupněn na webu v době bezprostředně minulé; obdobně rozsáhlý „Torontský“ staroanglický slovník je teprve u písmene F).

Tichého elektronický slovník přichází ve vhodnou chvíli i proto, že na rozdíl od doby, kdy vznikla jeho klasická podoba, slouží dnes daleko většímu okruhu potřeb, než je pouhá práce filologická. Zejména po provedení dalších kroků, o nichž diplomová práce pojednává v nárysu budoucího výhledu, Tichého aplikaci využijí odborníci i studenti zájmů diachronních i synchronních, lingvisté, filologové, literáti i překladatelé. Praktičnost tohoto projektu umocňuje dále skutečnost, že staroanglické období je v porovnání s těmi, jež v dějinách angličtiny následovala, přirozeně hůře podchycené i podchytitelné. A konečně je třeba ocenit fakt, že takto bude v blízké budoucnosti Bosworthův a Tollerův slovník coby jedinečný zdroj informací o nejstarším historickém období angličtiny možno doplnit a opravit.

Promyšlenost a pracnost předkládaného diplomového projektu spolu bezprostředně souvisejí. Tichý ve svém textu postupně podává přehled o existujících lexikografických zdrojích, ve vztahu k nim definuje kritéria žádoucího elektronického slovníku (zejména úplnost a přehlednost dat) a z hlediska makrostruktury a mikrostruktury analyzuje nejvýznamnější slovníky, jež by pro elektrizaci přicházely v úvahu. Zajímavě mj. konstatuje, že rozdíly mezi moderními, elektronickými verzemi a slovníky klasické podoby nejsou zásadní, přestože moderní slovníky vznikají díky součinnosti početných týmů specialistů a za využití nejmodernější techniky. Volbu Bosworthova a Tollerova slovníku Tichý posléze zdůvodňuje a podrobně analyticky přibližuje jednotlivé kroky projektu elektrizace od skenování až po vlastní vývoj slovníkové aplikace. Zde je dále namístě zdůraznit, že Tichý svůj projekt rozvinul na základě počáteční spolupráce

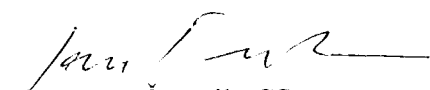
s týmem Seana Christa na Swarthmore College (*Germanic Lexicon Project*; USA) – využil jeho podkladů ve fázi skenování automatického rozpoznávání znaků z digitálních obrazů a poté, co zorganizoval závěrečné fáze práce na ručních opravách textu, Christovy návrhy ohledně sémantického značkování a dalších aspektů modifikoval a rozvinul tak, aby získanou elektronickou verzi bylo možno dále zdokonalovat. V přípravné, metodické fázi diplomand analyzoval více než 50 zdrojů dat (s vyloučením řady dalších – především těch úzce vymezených, malého rozsahu a hloubky, resp. u glosářů položek obsahujících méně než 3000 slov), a navíc konzultoval recenze a zprávy o nich.

Závěrečné kapitoly diplomové práce, věnované nejbližším kroků do budoucna, načrtávají záměr podrobnějšího a soustavnějšího sémantického označování slovníkových dat včetně nutnosti identifikovat pravidla, jež by umožnila rozlišení části sémantických značek automatickou cestou, resp. nastiňují základní obrysy záměru zkonstruovat morfologického analyzátoru staré angličtiny, prvního nástroje svého druhu, jenž by měl zásadní význam pro další morfologickou, slootovornou a lexikálně-sémantickou analýzu tohoto flektivního jazyka, a to především z hlediska provázání různých lexikografických zdrojů a korpusů staré angličtiny. Budoucí problémy, které se zde předjímají, vyplývají z typologických vlastností staré angličtiny (výrazná intreflexe, tj. četné morfonologické variace), z variantnosti pravopisu a z „neúplnosti“ její dochované gramatiky (jež ovšem alespoň částečně souvisí s lexikem, srov. četná hapax legomena; lze se proto nadít, že oproti očekáváním budou tyto další kroky muset výrazně sázet na neautomatizované postupy, tj. snadná odvoditelnost lemmat ze slovních tvarů je pouhým předpokladem).

Má vlastně jediná kritická poznámka se týká některých nedůsledností v jazykové podobě práce (nečetné potíže ve členech a v interpunkci, drobné formulační neobratnosti). Námítky, které by bylo možno vznést ohledně komentáře k vlastní aplikaci (např. požadavek delší, podrobnější a snad i názornější podoby kapitoly „Application Development“ či požadavek jakési předběžné příručky k samotné aplikaci), by po mém soudu dostatečně nezohlednily fakt, že vývoj aplikace byl do jisté míry projektem paralelním k projektu diplomovému, jakkoli spolu tyto projekty bezprostředně souvisejí. Nedávno dokončenou aplikaci je jistě třeba dále testovat; práce na textu diplomové práce byla jistě o to náročnější, že Ondřej Tichý zároveň dokázal organizovat, řídit a kontrolovat práci dosti početného, heterogenního a v některých fázích i poněkud amorfního týmu oprávců mnoha stovek rozpoznaných stránek slovníku.

Závěr: Diplomová práce Ondřeje Tichého, výjimečná hned v několika ohledech najednou – jako sonda do teoretické a praktické lexikografie staré angličtiny, jako studie z počítačové lexikografie, jako projekt funkční elektronické lexikografické aplikace – výrazně překračuje nároky na diplomové práce běžně kladené a zasluhuje předběžné hodnocení výborně.

Praha, 18. května 2007


doc. PhDr. Jan Čermák, CSc.
vedoucí diplomové práce