

Še dve slabosti sicer odlično organizirane konference je na zaključku – ko se je v imenu udeleženk in udeležencev s šopkom zahvalil glavni organizatorici in predsednici društva **Jagodi Granić** – previdno omenil znameniti hrvaški jezikoslovec **Damir Kalogjera**: velika časovna nedisciplina pretežnega dela nastopajočih in hudo pomanjkanje časa za diskusijo. Pri tolikšnem številu nastopajočih in samo največ dveh vzporednih sekcijah bo drugo slabost težko preseči – glede prve pa bi se gotovo dalo pomagati že zdaj. Čeprav drži, da agresivnejše predsedovanje sekcijam nekako ne sodi v prijazno splitsko okolje, brez tega vendarle ne bo šlo, saj na samodisciplino predavateljev po izkušnjah sodeč nikakor ne kaže računati.

Kljub izrečenim kritičnim besedam je srečanje hrvaškega društva za uporabno

jezikoslovje nedvomno dobrodošla in dragocena priložnost za izmenjavo védenja in mnenj med tistimi jezikoslovci in jezikoslovkami, ki jezik poudarjeno raziskujejo v družbenem, psihološkem in besedilnem kontekstu. Množična slovenska udeležba pomeni željo po predstavitvi in sodelovanju slovenskega jezikoslovja v širšem južnoslovanskem prostoru, po drugi strani pa vsaj malo razkriva tudi pereče pomanjkanje dejavnosti slovenskega društva za uporabno jezikoslovje. V devetdesetih letih je prirejalo kakovostne in odmevne bienalne kongrese, trenutno pa miruje. Lahko kako pomagamo?

Marko Stabej
Univerza v Ljubljani, Filozofska fakulteta
marko.stabej@ff.uni-lj.si

LEXICOM Europe 2006 – Annual Workshop in Lexicography and Lexical Computing, Opatija, 4.–9. junij 2006

Letos se je v Opatiji odvijala že šesta leksikografska delavnica LEXICOM, ki jo od leta 2001 na različnih krajih po Evropi in v sodelovanju z različnimi (akademskimi) institucijami prireja angleško podjetje Lexicography MasterClass Ltd. Letošnji gostitelji, Pomorski fakultet, Sveučilište u Rijeci, so za en delovni teden v mediteransko letovišče privabili približno trideset udeležencev iz desetih evropskih držav, Združenih držav Amerike in Arube. Poleg številčne hrvaške zasedbe je po številu, a tudi delavni vnemi izstopala slovenska sedemčlanska delegacija iz univerzitetnega in založniškega sveta, ki je svoje bivanje v primorju izkoristila tudi za spoznavanje kolegov iz tujine in domovine. Vreme jim je bilo pri tem le v pomoč, saj so Angleži poleg znanja s seboj prinesli hladne in deževne dni, ki so bili kot nalašč za intenzivno delo in živahne debate ob kavi.

Lexicography MasterClass Ltd je podjetje, ki v različnih oblikah ponuja svoje leksikografsko in z leksikografijo povezano računalniško znanje. Poleg raznovrstnih izobraževanj, namenjenih široki paleti strokovnih profilov (študentom, leksikografom, prevajalcem, jezikovnim tehnologom, urednikom ...), se ukvarja tudi s svetovanjem in sodelovanjem pri korpusnih projektih, vprašanih rabe obstoječih ali izdelave novih računalniških leksikografskih orodij ter pri načrtovanju in izpeljavi konkretnih slovarskih projektov. Ker pa so podjetje pravzaprav ljudje, še nekaj besed o njih: **Sue Atkins**, **Michael Rundell** in **Adam Kilgarriff** so vrhunski evropski strokovnjaki s področja teoretične in praktične, tako dvojezične kot enojezične leksikografije ter računalniškega jezikoslovja, ki so sooblikovali že mnogo (velikih) slovarskih in jezikovnotehnoloških projektov, na primer *Collins-Robert English-French Dictionary*,

Oxford-Hachette English-French Dictionary, Macmillan English Dictionary for Advanced Learners, FrameNet, BNC ter Word Sketch Engine in BootCat, če jih naštejemo res le nekaj, eden najnovejših velikih projektov podjetja pa je bila zasnova novega angleško-irskega slovarja, ki je vključevala tudi jezikovnotehnološko podporo, med drugim izdelavo 30-milijonskega korpusa irščine in 25-milijonskega korpusa irske angleščine (več o projektu na <http://www.focloir.ie>, več o podjetju pa na <http://www.lexmasterclass.com>).

LEXICOM je intenzivno petdnevno izobraževanje, ki združuje kratka, zgoščena predavanja o teoretičnih izhodiščih leksikografije, računalniške leksikografije in korpusnega jezikoslovja ter samostojno ali skupinsko praktično delo, deloma organizirano v paralelnih sekcijah, slovaropisni in jezikovnotehnološki. Predavatelji so živahno, zanimivo in s precejšnjo mero humorja posredovali vsebine, povezane z zasnovo in izdelavo korpusnih slovarjev ter gradnjo in učinkovito izrabo spletnih korpusov v leksikografiji. Atkins in Rundell sta udeležence seznanila predvsem z osnovnimi in bistvenimi leksikografskimi problemi in možnimi rešitvami, ki segajo od zasnove slovarja do izdelave konkretnih geselskih člankov. Pri zasnovi slovarja je bila posebej izpostavljena pomembnost oblikovanja natančnih navodil za leksikografe, usklajenih z uredniško politiko, ta pa mora med drugim temeljiti na upoštevanju ciljnega uporabnika, saj je od tega odvisna marsikatera odločitev o vsebini in obliki slovarja. S tem v zvezi gre omeniti tudi predavanje o koristnosti izdelave predlog, tj. vzorčnih gesel z obveznimi in priporočljivimi elementi za skupine besed, ki pripadajo istemu razredu (npr. živali, dnevi v tednu, barve), da se izognemo nekonsistentnosti in leksikografom olajšamo delo. Načeta so bila tudi vprašanja črpanja podatkov o jeziku, kjer so bili ocenjevalno sopostavljeni različni viri: introspekcija, obstoječi jezikovni viri (npr. slovarji, slovnice), informacije rojenih govorcev, citatni lističi in korpusi. Ker se tehtnica s

plusi močno nagiba v smeri proti korpusom, smo bili seznanjeni z osnovami korpusnega jezikoslovja, tako z gradnjo kot izrabo korpusov v leksikografske namene. Poudarek je bil predvsem na vprašanju, kako izbrati relevantne korpusne informacije za pisanje konkretnega geselskega članka. Ta vprašanja so seveda tesno povezana s procesom izdelave slovarske podatkovne zbirke, ki mora vsebovati vse glavne podatke o pomenskih in slovničnih lastnostih leksikalnih enot, kolokacijah, frekvenci, tipični rabi in nekaj tipičnih (realnih) zgledov rabe. Na kratko je bila predstavljena teorija semantičnih shem (*frame semantics*), njena uporabnost v leksikografiji ter osnove dvojezične leksikografije. Ukvarjali pa smo se tudi z zelo pomembnimi in vse prej kot trivialnimi vprašanji določanja pomenov večpomenskih leksemov, slovarske obravnave prenesenih pomenov in oblikovanja definicij v enojezičnih slovarjih. Prav v tem delu smo bili podrobneje seznanjeni z različnimi teorijami, zgodovinskimi in sodobnimi rešitvami v angleških slovarjih ter s primeri dobre in slabe prakse. Ti se nanašajo na združevanje in ločevanje pomenov, razvrščanje pomenov v slovarskem geselskem članku ter problematiko različnih pristopov k oblikovanju definicij in odločitev o njihovi obliki, vsebini in funkciji. Na tem mestu nismo dobili enoznačnih odgovorov na primer o tem, kaj je dobra definicija, ker je vprašanje večplastno in taki odgovori niti niso mogoči, so pa bili konkretni primeri dovolj pomenljivi za posredovanje navodil, na kaj vse mora leksikograf paziti, ko se loteva teh zahtevnih nalog.

Predstavljene jezikovnotehnološke vsebine so obravnavale načela zasnove in gradnje korpusov, gradnjo korpusov iz gradiva, dostopnega na spletu, ter prednosti uporabe specializiranih računalniških orodij za sestavo slovarjev. Med predavanji, ki so zahtevala nenehno sodelovanje udeležencev, smo se seznanili s številnimi funkcijami orodja Word Sketch Engine (<http://www.sketchengine.co.uk>) in z uporabo regularnih izrazov za naprednejše iskanje po označenih

korpusih (*Corpus Query Language*) ter pisanje gramatik. Pobljize smo spoznali tudi BootCat (<http://corpora.fi.muni.cz/bootcat>), orodje za gradnjo spletnih korpusov, in enega izmed zmogljivejših orodij za sestavo slovarjev, TshwaneLex, ki ga je na povabilo predavateljev predstavil **Gilles-Maurice de Schryver** iz podjetja TshwaneDJe HLT (<http://tshwanedje.com/tshwanelex>).

Predavanja so bila večinoma izvedena v prvih dneh delavnice, saj je bil četrtkov dan namenjen samostojnemu praktičnemu delu, petkov pa oceni naših izdelkov in premisleku o prihodnosti. Kljub precej omejenemu času so predavanja zaobjela zelo veliko vsebin, kar se seveda pozna pri manj intenzivnem poglobljanju v problematiko in manjši pluralnosti prikazanih teorij, pristopov in konkretnih rešitev. Vendar pa od tovrstnih intenzivnih tečajev kaj takega tudi ne gre pričakovati, sploh če je ciljna publika strokovna in naj bi torej bila po področju razgledana. Se pa je ravno zaradi tega zastavljalo vprašanje relevantnosti odpredavanjih vsebin, saj se vsaj na prvi pogled zdi, da bi vsakdo, ki se kakorkoli ukvarja z leksikografijo in jezikovnimi tehnologijami, te vsebine vsaj deloma moral poznati že prej. Temu je najverjetneje botrovala velika heterogenost skupine udeležencev, ki smo izhajali iz zelo različnih okolij, bili različno stari, različnih strokovnih profilov in z različnimi izkušnjami. S stališča podiplomskih študentk, ki sva med študijem pridobili teoretične osnove v obravnavanih temah in imava z njimi že nekaj konkretnih izkušenj, pa lahko rečeva, da je bila osvežitev znanja dobrodošla, še posebej zato, ker je bila dopolnjena z nadgradnjo, spoznavanjem stanovskih kolegov in skoraj nenehno pripravljenostjo predavateljev odgovarjati na naša še kako raznovrstna vprašanja.

Praktično leksikografsko delo je bilo povezano z izdelavo geselskih člankov dveh angleških besed, in sicer na podlagi podatkov iz korpusa *BNC* (*British National Corpus*), ki smo ga pregledovali s pomočjo orodja Word Sketch Engine. Že prvi dan

delavnice smo se morali spopasti z besedo *honeymoon* (*medeni tedni*), ki je bila zaradi nizke frekvence in ne preveč zapletene semantične vrednosti leksikografu zelo prijazna, kljub temu pa po skupinah nismo oblikovali identičnih geselskih člankov in je bila tako vaja dober uvod v nadaljnja predavanja o problemih, s katerimi se leksikografi spopadajo pri vsakdanjem delu. Po »prijetnih medenih tednih« pa je nastopilo nekoliko bolj napeto delovno vzdušje, ko smo se v četrtek ukvarjali z izrazom *tense* (*napet, napeti, čas*). Naša naloga je bila izdelati geselski članek za enojezični slovar za tujce in dvojezični angleško-x slovar. Pri skupinskem delu se je pokazalo, da imamo zelo različne poglede na to, kako se lotiti takšne naloge, pa tudi pri pogajanjih o končnih odločitvah so posamezniki odločno zagovarjali svoje rešitve. Kakorkoli že, večinoma nam je v skupinah uspelo sestaviti vsaj en geselski članek, petkova debata o naših zadregah in rešitvah pa je bila zelo plodovita in poučna.

Praktični del jezikovnotehnoške sekcije je letos potekal v nekoliko okrnjeni obliki, saj so bili interesi večine udeležencev izrazito jezikoslovno obarvani, dinamiko delavnice pa je še dodatno upočasnili tehnološko nespretno rešen dostop do interneta. A kljub temu nam je uspelo podrobno preizkusiti vso zmogljivost orodja Word Sketch Engine, s pomočjo BootCata pa vsakemu sestaviti svoj spletni korpus. Tega smo nato dodali v Word Sketch Engine in pravkar naučene tehnike iskanja po angleškem korpusu preizkusili tudi v svojem jeziku. BootCat se je izkazal kot priročno orodje za hitro gradnjo tematsko specializiranih korpusov iz gradiva, dostopnega na spletu, za angleščino, nemščino, francoščino, italijanščino in španščino pa omogoča tudi besednovrstno označevanje in lematizacijo. Tako pripravljen korpus je nato mogoče preiskovati z orodjem Word Sketch Engine, ki poleg konkordančnika ponuja možnosti iskanja kolokacij, slovničnih relacij med besedami, semantično sorodnih besed in primerjave podobnosti in razlik med dvema izrazoma.

Za konec. Po napornem tednu so organizatorji že sicer odlično izveden seminar popestrili še s skupno večerjo, ki smo jo vsi prisotni z veseljem izkoristili za intenzivno druženje in izmenjavo izkušenj. Poleg dobre organizacije pa gre izpostaviti še izčrpno seminarsko gradivo z zapiski s predavanj in vajami, ki so ga za udeležence pripravili predavatelji. Seminar smo zaključili zazrti v prihodnost, ki naj bi prinesla celovitejše rešitve za računalniško podprto leksikografijo, še večjo integracijo korpusov in slovarjev

ter inteligentnejšo ponudbo slovarskih informacij za specifičnega uporabnika. Zaradi velikega števila udeležencev iz Slovenije, ki priča o potrebnosti tovrstnih izobraževanj v našem prostoru, pa si za konec zaželimo še, da bi podobne leksikografske delavnice prav kmalu zaživele tudi pri nas.

Darja Fišer in Tadeja Rozman
Univerza v Ljubljani, Filozofska fakulteta
darja.fiser1@guest.arnes.si
tadeja.rozman@ff.uni-lj.si