

## POLETNA ŠOLA KORPUSNEGA JEZIKOSLOVJA V LANCASTRU

Jaka ČIBEJ

Filozofska fakulteta Univerze v Ljubljani

Dafne MARKO

Ljubljana

*Čibej, J., Marko, D. (2016): Poletna šola korpusnega jezikoslovja v Lancastru. Slovenščina 2.0, 4 (1): 10–14.*

DOI: <http://dx.doi.org/10.4312/slo2.0.2016.1.10-14>.

### 1 UVOD

Med 12. in 15. julijem je na Univerzi v Lancastru potekala poletna šola korpusnega jezikoslovja Lancaster Summer Schools in Corpus Linguistics and Other Digital Methods. Poletno šolo so organizirali UCREL (University Centre for Computer Corpus Research on Language), ERC (Evropski svet za raziskave – European Research Council), CASS (ESRC Centre for Corpus Approaches to Social Science) in ESRC (Economic and Social Research Council), razdeljena pa je bila na šest programov, prilagojenih različnim področjem: Korpusno jezikoslovje za proučevanje jezikov (Corpus Linguistics for Language Studies), Korpusno jezikoslovje za družbene vede (Corpus Linguistics for Social Science), Korpusno jezikoslovje za humanistiko (Corpus Linguistics for Humanities), Statistika za korpusno jezikoslovje (Statistics for Corpus Linguistics), Geografski informacijski sistemi za digitalno humanistiko (Geographical Information Systems for the Digital Humanities) in Korpusno podprta obdelava naravnih jezikov (Corpus-based Natural Language Processing).

Poletna šola, ki je tokrat potekala že šestič, vsako leto sprejme udeležence iz širokega nabora področij, npr. jezikoslovja, sociologije, politologije, socialne antropologije, izobraževanja, ekonomije, zgodovine, literarnih ved, računalništva in informatike, namenjena pa je predvsem magistrskim in

doktorskim študentom ter raziskovalcem, ki jih zanima obdelava jezikovnih podatkov s kvantitativnimi korpusnimi metodami. Iz Slovenije so se letošnje poletne šole (smeri Statistika za korpusno jezikoslovje ter Korpusno jezikoslovje za proučevanje jezikov) udeležili Jaka Čibej (Filozofska fakulteta Univerze v Ljubljani), Dafne Marko (študentka Filozofske fakultete Univerze v Ljubljani) in Darja Fišer (Filozofska fakulteta Univerze v Ljubljani in Institut »Jožef Stefan«).

## **2 PROGRAM**

Poletna šola je zajemala dveurna predavanja in delavnice, dve enourni plenarni predavanji ter serijo treh 20-minutnih predavanj na temo preteklosti, sedanjosti in prihodnosti korpusnega jezikoslovja.

Na smeri Statistika za korpusno jezikoslovje je bil glavni predavatelj Vaclav Brezina (Univerza v Lancastru), pri delavnicah in praktičnih vajah pa sta pomagala tudi Matt Timperley in Irene Marin Cervantes (oba prav tako z Univerze v Lancastru). Poleg nekaterih ključnih statističnih pojmov in vidikov učinkovite vizualizacije podatkov so predstavili tudi korpusne vire in orodja, ki so jih razvili na Univerzi v Lancastru za izdelavo kolokacijskih mrež, izračun statističnih mer in obdelavo korpusnih podatkov. Predavala sta tudi Andrew Hardie (Univerza v Lancastru), ki se je osredotočil predvsem na statistične mere za prepoznavanje in luščenje kolokacij iz korpusov (npr. logaritemska verjetnost in vzajemna informativnost), ter Stefan Th. Gries (Univerza v Santa Barbari, ZDA), ki je predstavil nekaj praktičnih primerov uporabe inferenčne statistike za preverjanje in posploševanje hipotez na podlagi korpusnih podatkov.

Na smeri Korpusno jezikoslovje za proučevanje jezikov so se nekatera predavanja prekrivala s statistično smerjo, poleg že omenjenih strokovnjakov pa so predavanja in delavnice vodili še štirje predavatelji z Univerze v Lancastru in trije predavatelji z drugih ustanov. Michael Barlow (Univerza v Aucklandu) je predstavil pomen in uporabo orodij za gradnjo ter analizo vzporednih

korpusov, John Flowerdew (Mestna univerza v Hong Kongu) se je osredotočil predvsem na analizo akademskega diskurza s pomočjo korpusnih metod, Yukio Tono (Tokijska univerza za tuje jezike) pa je predstavil pristop, metode in tehnike označevanja napak v korpusu usvajanja tujega jezika. Claire Hardaker (Univerza v Lancastru) je v svojem predavanju predstavila osnove forenzičnega jezikoslovja ter na podlagi dveh preiskav izpostavila prednosti in slabosti korpusnih metod pri ugotavljanju avtorstva in prepoznavanju stilističnih posebnosti različnih piscev. Poleg korpusnih pristopov k analizi diskurza so udeleženci poletne šole skozi serijo praktičnih vaj spoznali tudi naprednejše iskalne pogoje, jezik CQP in regularne izraze, ki se uporabljajo za učinkovito korpusno analizo.

Prvo plenarno predavanje (z naslovom *Contextualising patterns found in literary texts: linking the fictional and the real world*) je pripravila Michaela Mahlberg (Univerza v Birminghamu). Na predavanju je predstavila nekaj pristopov k opazovanju diskurza v literarnih besedilih s pomočjo korpusov. Drugo plenarno predavanje (z naslovom *Corpus-assisted discourse studies: virtues and challenges – but not a new paradigm*) pa je vodil Alan Partington, ki je na primeru iskanja kolokatorjev za poimenovanja določenih verskih skupnosti prikazal možnosti opazovanja diskurza in družbenega odnosa s pomočjo korpusnih virov.

Ob zaključni večerji, ki je bila posvečena CASS (Centre for Corpus Approaches to Social Science), dobitniku nagrade *The Queen's Anniversary Prize for Higher and Further Education*, so bila izvedena še tri 20-minutna predavanja, ki so bila zasnovana kot strnjen pregled zgodovine korpusnega jezikoslovja, sedanjega položaja discipline ter pričakovanega razvoja v prihodnosti. Predavali so Charlotte Taylor (Univerza v Sussexu), Vaclav Brezina (Univerza v Lancastru) in Jesse Egbert (Univerza Brigham Young).

### **3 POMEN POLETNE ŠOLE**

Poletna šola korpusnega jezikoslovja v Lancasteru ponuja dober pregled korpusnih metod in je zato primerna tudi za študente in raziskovalce s področja jezikoslovja, ki še nimajo veliko izkušenj z analizo korpusnih podatkov. Tako pri smeri Korpusno jezikoslovje za proučevanje jezikov kot pri smeri Statistika za korpusno jezikoslovje je treba pohvaliti nezapleten, razumljiv in praktičen pristop k razlagi koristnih statističnih pojmov, ki je primeren tudi za udeležence, ki še nimajo trdne podlage v statistiki, a si želijo več znanja o statističnih testih, regularnih izrazih in načelih, ki se jih v sodobnem korpusnem jezikoslovju uporablja za izdelavo seznamov ključnih besed, kolokacij in kolokatorjev. Poletna šola ponuja strnjen teoretični pregled uporabnih metod in tehnik v korpusnem jezikoslovju, poleg tega pa se udeleženci v sklopu delavnic ter praktičnih vaj поблиže spoznajo s sodobnimi orodji in korpusnimi viri, ki so raziskovalcem s tega področja trenutno na voljo. Udeleženci imajo obenem tudi možnost, da strokovnemu odboru zastavijo vprašanja, povezana s teoretičnimi ali metodološkimi težavami, na katere so naleteli pri lastnem delu oz. raziskavi, ob zaključku poletne šole pa tovrstne težave prediskutirajo z vodilnimi strokovnjaki s področja korpusnega jezikoslovja.

Podrobnejši opisi vseh smeri in druge informacije o poletni šoli so na voljo na uradni spletni strani.

## **LANCASTER SUMMER SCHOOL IN CORPUS LINGUISTICS**

To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-  
Deljenje pod enakimi pogoji 4.0 Mednarodna.

This work is licensed under the Creative Commons Attribution ShareAlike 4.0  
License International.

<http://creativecommons.org/licenses/by/4.0/>

