

perimentů může být projekt studia morfologického primingu, v němž se sleduje, zda prezentace slova v určitém gramatickém tvaru, např. akuzativu, usnadňuje zpracování jiného slova ve stejném tvaru. Experimenty zkoumají rovněž to, zda se efekty liší v závislosti na způsobu realizace daných kategorií např. u různých deklinačních tříd. Experimenty zde využívají úlohu detekce slova (*lexical decision*), při níž účastníci sledují zobrazená slova a stisknutím tlačítka mají co nejrychleji rozhodnout, zda se jedná o skutečné slovo, nebo nikoliv. Tímto způsobem lze zajistit, že účastník si musí vybatvit a do značné míry zpracovat zobrazené slovo, jinak by nemohl spolehlivě odpovědět. Jednoduchá odpověď přitom dovoluje přesné a spolehlivé měření reakčního času.

Další informace např. o výzkumech realizovaných v laboratoři, o možnostech podávání projektů a o prioritních tématech jsou průběžně aktualizovány na webových stránkách laboratoře <<http://labels.ff.cuni.cz>>, které získají v roce 2013 i anglickou verzi. Zájemci o využití laboratoře pro vlastní výzkum nebo o konzultace s týmem laboratoře najdou na těchto webových stránkách všechny potřebné kontaktní informace.

Filip Smolík a Eva Lehečková

WORKSHOP CORPUS LINGUISTICS OR HOW TO DO STATISTICS WITH LANGUAGE

Ve dnech 9.–11. července 2012 se v kampusu Univerzity v Řezně konal workshop s názvem *Corpus Linguistics or How to Do Statistics with Language*. Hlavní náplní workshopu bylo studenty a badatele v oboru slovanských jazyků, kteří mají zájem při svém korpusovém výzkumu používat statistické metody, obeznámit s programem R. Informace o konání workshopu byla avizována dvěma způsoby. Organizátorka workshopu Sandra Birzerová jednak uveřejnila prostřednictvím Deutscher Slavistenverband oznámení o uskutečnění pracovního semináře a jednak na seminář osobně pozvala kolegy z oboru a jejich studenty. Samotný workshop pak vedla Dagmar Divjaková, jejíž pozvání bylo financováno organizací Vielberth-Stiftung — nadací, která podporuje výzkumnou činnost řezenské univerzity.

Dagmar Divjaková v současnosti působí na Univerzitě v Sheffieldu. Na Katedře ruských a slovanských studií (Department of Russian and Slavonic Studies) zastává roli ředitelky ruských a polských studií, dále spoluřídí Centrum pro jazykový výzkum (Centre for Linguistic Research) a je koordinátorkou etiky výzkumu v rámci Ústavu jazyků a kultur (School of Languages and Cultures). Ve své vědecké práci primárně vychází z tzv. usage-based kognitivní lingvistiky. Její zájem spočívá v porozumění tomu, jak se naše kognitivní schopnosti podílejí na vzniku jazykových struktur a vzorců, a ve zkoumání toho, co musí jazyk poskytnout při připsování významu slovům a konstrukcím těm, kteří se jej učí. Dokladem její kognitivnělingvistické orientace je i fakt, že plní funkci viceprezidentky / zvolené prezidentky SCLA (Slavic Cognitive Linguistic Association — Slovanské asociace kognitivní lingvistiky) a že se nachází v redakční radě časopisu *Cognitive Linguistics*. Vědecký profil Dagmar Divjakové doplňme o informaci, že je členkou výboru BASEES (Britské asociace pro slovanská a východoevropská studia).

V Řezně tato kognitivní lingvistka představila zhruba třiceti účastníkům workshopu základy práce s programem R. Tento matematický program se těší stále větší oblibě mezi těmi vědci a pracovníky, kteří potřebují ke své práci účinný software pro statistické analýzy a tvorbu grafů. Lingvistika je právě jednou z oblastí, v níž se program R začíná dostávat silněji do obecného povědomí. K jeho velké popularitě přispívá zejména to, že je volně dostupný (jedná se o tzv. open source software), je kompatibilní s více operačními systémy (Windows, Linux, Macintosh) a je průběžně aktualizován v závislosti na vývoji metod ve statistice. Na druhou stranu může někoho, kdo nemá s programováním mnoho zkušeností, překvapit neintuitivním uživatelským prostředím. Pro výpočet jednotlivých funkcí jsou používány příkazové řádky a výsledné statistické analýzy dostáváme ve formě výstupních řádků. Příkazy v programu R se přitom řídí vnitřní logikou, kterou je nutno si osvojit. K jejich zdárnému sestavení si též musíme zapamatovat podobu řady operátorů, funkcí a jejich argumentů. Každý, kdo se chystá využívat program R, je tedy odkázán k soustředěnému studiu materiálu objasňujícího základní principy práce v tomto programu. Poměrně velké množství takového materiálu, zahrnujícího návody pro jednotlivé statistické výpočty, se nachází na internetu. Alternativu — a vzhledem k její interaktivní povaze ji lze považovat snad i za vhodnější — k studiu příslušného materiálu představují přednášky a workshopy, na nichž se zacházení se softwarem přímo demonstruje. Do této kategorie spadal i workshop Dagmar Divjakové.

Statistické analýzy lze provádět na různých typech dat. Dagmar Divjaková se zaměřila ve svém výkladu na analýzu dat korpusových a celý workshop v podstatě pojala jako uplatnění programu R v korpusově orientované lingvistice.

První den workshopového cyklu byl zasvěcen samotným základům korpusové lingvistiky a statistiky. Dagmar Divjaková zasadila korpusovou lingvistiku do širokého lingvistického kontextu, v němž ji vymezila vůči tzv. pohovkové (armchair) lingvistice, zamýšlela se nad jejím charakterem (kladla si otázku, zda je nástrojem, modelem, nebo teorií) a zmínila hlavní pojmy, které jsou s ní spojeny. Následně účastníky uvedla do základů statistiky. Účastník workshopu tedy nemusel disponovat téměř žádnými vstupními znalostmi z těchto odvětví. Vše mu bylo srozumitelně vyloženo od základů.

Po probrání hlavních termínů statistiky (např. modus, medián, korelace) se v druhé půli dne účastníci workshopu seznamovali s programem R a učili se zadávat první příkazy (jak načítat soubory, jak zobrazovat názvy proměnných, jak zjistit medián daného vektoru apod.). Účastníci si mohli do programu R nahrát svůj vlastní soubor dat a zkusit příkazy na něm. V možnosti zpracovat si v programu svá data shledáváme velkou přednost tohoto workshopu. V případě, že vlastní soubor účastníci neměli, jim Dagmar Divjaková k procvičování příkazů poskytla soubory svoje.

Také další dva dny workshopového cyklu se skládaly ze dvou částí. První půle dne byla spíše výkladová a praktická cvičení prostupující výkladem probíhala na datech, která připravila vedoucí workshopu. V druhé půli dne měli účastníci možnost nabyté poznatky aplikovat na datech vlastních.

Náročnost statistických funkcí, o nichž Dagmar Divjaková hovořila a jež účastníci vzápětí zkusili provádět na svých datech, se každým dnem zvyšovala. Probrány byly například výpočet t-testu, chí-kvadrát testu či lineárních regresních modelů.

Workshop nabídl velmi dobrou základní orientaci práce v programu R. Účastníci workshopu se seznámili s počítáním výše uvedených statistických funkcí. V případě jakýchkoli potíží při práci s programem se mohli obrátit na Dagmar Divjakovou, která jim daný problém pomohla vyřešit. Vzhledem ke zmiňovanému neintuitivnímu uživatelskému prostředí a ke složitosti zápisu příkazů se tato možnost ukázala jako velmi vítaná.

Problém workshopu je možné spatřovat v teoretickém výkladu, respektive v náročnosti probíraných témat. Ve svém výkladu Dagmar Divjaková objasňovala některé základní informace ohledně sbírání dat pro statistickou analýzu a některé elementární statistické pojmy a metody (viz výše). Mimo to se však věnovala také poměrně pokročilým výpočtům, jako jsou právě např. lineární regresní modely. Tato jistá disproporce v náročnosti probíraných okruhů mohla především těm, kteří ve statistice nejsou dostatečně zblhlí, činit jisté potíže. Na druhou stranu je vzhledem k délce workshopu takovýto rychlejší postup v probíraných tématech pochopitelný.

Celkově lze workshop hodnotit pozitivně. Dagmar Divjaková uvedla jeho účastníky do práce s programem R velmi dobře. Účastníci workshopu se budou muset s programem samozřejmě seznamovat ještě delší čas, než se v něm budou schopni bez problémů orientovat a než v něm budou schopni používat pokročilé funkce. Nicméně poté, co si práci v programu R řádně osvojí (k čemuž by mělo přispět i prostudování literatury, kterou účastníkům Dagmar Divjaková doporučila), bude statistické zpracování dat v něm nejen časově úsporné, ale i elegantní. Řezenský workshop jim měl pouze usnadnit první kroky na cestě k tomuto cíli. A tento úkol zdárně naplnil.

Adam Kríž a Štěpán Matějka