

JESSE EGBERT, PAUL BAKER (EDS.): USING CORPUS METHODS TO TRIANGULATE LINGUISTIC ANALYSIS

New York and London: Routledge, 2020, 286 stran

ISBN: 978-1-138-08254-0



Triangulace — tedy zkoumání téhož jevu různými metodami, různými výzkumníky, s využitím různých zdrojů dat nebo teoretických modelů — se uplatňuje zejména ve společenských vědách, obvykle jako prostředek, který ověřuje validitu určité výzkumné metody.

Také v souvislosti s korpusovou lingvistikou se v poslední době čím dál více mluví o nutnosti jejího propojení s jinými odvětvími lingvistiky nebo s jinými vědními disciplínami. Metodologická triangulace zde může vést na jedné straně k rigoróznějším závěrům korpusově založeného výzkumu a na druhé straně ukazuje jeho využitelnost v různých oblastech, včetně těch nelingvistických.

Kniha navazuje na publikaci týchž autorů *Triangulating Methodological Approaches in Corpus Linguistics*, která se ovšem soustředila jen na metodologickou pluralitu v rámci korpusové lingvistiky. V současné publikaci shromáždili editoři empirické studie, které překračují hranice korpusové lingvistiky a testují možnosti triangulace metod a dat z oblasti psycholingvistiky, aplikované lingvistiky nebo analýzy diskurzu s metodami a daty korpusově-lingvistickými. Metodologická triangulace je zde definována široce jako využití dvou nebo více nezávislých metod v téže studii k hledání odpovědi na tytéž výzkumné otázky (s. 6). Triangulace v lingvistickém výzkumu neslouží jen k ověření validity výsledků. Výsledkem triangulace metod je zde komplexní pohled na zkoumané otázky.

Vztah mezi použitými metodami lze charakterizovat buď jako vztah konvergence, nebo korelace (,convergent', nebo ,correlational methodological triangulation', kapitola 11). Příkladem prvního přístupu, kdy se k zodpovězení jedné otázky používají dvě různé metody s nadějí, že jejich výsledky poskytnou komplementární konvergentní výsledky, může být kapitola 3. Odpověď na otázku, kterou si zde McEnery, Bakerová a Dayrellová kladou — Kdy nastala v Británii v 19. století období extrémního sucha a jaké byly jejich dopady? — je důležitá pro dlouhodobé meteorologické a hydrologické modelování. Autoři kombinovali korpusově-lingvistické metody (frekvence zmínek o suchu v novinových článcích jako indikátor jeho délky a intenzity), detailní čtení vybraných textů, geoparsing (přiřazování geografických identifikátorů ke sloům a frázím v textu) a analýzu historických hydrologických záznamů. Konvergence výsledků analýzy korpusových dat a hydrologických dat potvrdila důvěryhodnost obou metod a autoři popsali také dosud neznámá období sucha, která objevili korpusovými metodami. Zajímavé je v této studii také ,cyklické' uspořádání metod, kdy se autoři k jednotlivým přístupům vrací během výzkumu několikrát a postupně kladou další otázky nebo ty původní zpřesňují.

Podobně používají triangulace dvou nezávislých metod, které se snaží odpovědět na tutéž otázku, také Schnurová a Csomayová (kapitola 2) nebo Baker (kapitola 4), kteří kombinují korpusově-lingvistické metody s metodami analýzy diskurzu. Schnurová a Csomayová hledají jazykové indikátory, s jejichž pomocí můžeme segmentovat texty (zde přepisy univerzitních přednášek). Na rozdíl od předchozí studie



jsou zde dvě metody (manuální a počítačová) použity na sobě nezávisle, na jejich pořadí nezáleží. Díky kombinaci metod vidíme text ze dvou komplementárních pohledů — funkčního (distribuce funkčních kategorií v textu) a formálního (identifikace jazykových prostředků charakteristických pro jednotlivé segmenty textu). Bakerova kapitola zkoumá, jak britské noviny píšící o problému obezity. Propojují se zde metody korpusové lingvistiky (kolokace, analýza konkordančních řádků, klíčová slova) s kritickou analýzou diskurzu (detailní kritické čtení vybraných sad článků). Stejnou měrou jako reprezentace obezity v tisku autora zajímají metodologické otázky výběru dat, způsobu jejich porovnávání a vlivu pořadí, v němž na sebe kroky analýzy navazují, na interpretaci výsledků.

Na testování konvergence výsledků získaných dvěma nezávislými metodami je založena také kapitola 7, ve které se Egbert a Davies zaměřili na významové vztahy mezi dvěma substantivy, z nichž první je modifikátorem druhého, např. *baby oil*. Použité metody označují autoři jako ‚use-based‘ (identifikace více než 1500 kompozit v diachronním korpusu americké angličtiny COHA a popis diachronních změn v jejich frekvenci) a ‚user-based‘ (sémantickou klasifikaci kompozit prováděli manuálně ‚kódovači‘, zaměstnanci Mechanical Turk). Na základě statistické analýzy takto získaných dat popisují autoři vývojové tendence v zastoupení různých sémantických vztahů mezi substantivy ve zkoumaných kompozitech. Ukazují například, jak se významové vztahy vyvíjejí od konkrétních vztahů lokace a složení (např. *kitchen table*, *stone wall*) k abstraktnějším vztahům a procesům (např. *birth control*).

Také druhý typ metodologické triangulace — korelaci metod — můžeme ilustrovat na příkladu jedné z kapitol. Na rozdíl od konvergence, kde různé metody odpovídají na tutéž otázku, zde jde o korelaci souborů dat získaných různými metodami ve snaze zjistit, jaký je mezi těmito daty vztah. Tento typ triangulace se uplatňuje zejména ve studiích, které kombinují korpusově-lingvistické a psycholingvistické metody (kapitoly 8, 9 a 10). Tyto studie spojuje zájem o to, zda existuje korelace mezi tím, co vnímáme jako frekventované nebo podstatné a co tak můžeme popsat statisticky s oporou ve velkých jazykových korpusech. Experimenty popsané v těchto kapitolách pracují s teorií označovanou jako ‚priming‘ — vytváření kognitivních spojení na základě opakované expozice určitým vzorcům nebo spojením slov. Hughesová a Hardie (kapitola 8) se věnují tématu, které se objevuje v korpusově založených studiích často — kolokacím. S využitím elektroencefalografie ukazují, jak metody užívané v korpusové lingvistice k identifikaci kolokací korespondují s neurologickými odezvami při rozpoznávání kolokátů. Z jiného úhlu zkoumá ‚priming‘ Gries (kapitola 9). Koreluje experimentální data (výsledky experimentu, v němž mají studenti angličtiny doplňovat neúplné věty) s modelem vycházejícím z korpusu rodilých mluvčích angličtiny a ukazuje, jak strukturní ‚priming‘ ovlivňuje, jaké konstrukce používají nerodilí mluvčí angličtiny. V kapitole 10 popisuje Ellis experiment, kde se pomocí toho, jak dlouho trvá mluvčímu přečíst slovo (‚voice onset time‘) zjišťuje, jaký vliv na rozpoznání slova (v tomto případě předložky následující po slovese) má frekvence jeho výskytu po určitém slově (zjišťovaná v referenčním korpusu), délka, ustálenost a sémantická prototypičnost spojení daných slov.

Korelace metod se objevuje i ve studiích, kde se korpusových metod užívá spolu s metodami aplikované lingvistiky (kapitoly 5 a 6). Pátá kapitola (LaFlair, Staplesová

a Yan) ukazuje, jak kombinace multidimenzionální analýzy a testovacích schémat vede ke zjemnění a zpřesnění kritérií hodnocení studentských prací. Gablasová v kapitole 6 koreluje frekvenční data z referenčního korpusu s asociačními testy. Díky triangulaci metod prokázala, že faktory, které fungují jako spouštěče lexikálních asociací, zahrnují jak slova nadřazená, tak slova vysoce frekventovaná.

Knihu uzavírá kapitola, kde se Baker a Egbert vracejí k postupům uplatněným v předchozích kapitolách a hodnotí, do jaké míry a jakým způsobem přispěla triangulace k výsledku popsaných experimentů, a obecněji, jak může triangulace lingvistických metod zahrnující metody korpusové lingvistiky přispět k pochopení různých procesů a variace v jazyce. Shrnují zde také, co autoři jednotlivých studií popisují jako výhody a nedostatky zvolených přístupů. Jak autoři zdůrazňují, bude kniha užitečným východiskem pro korpusové lingvisty, protože „korpusová lingvistika nemůže plně rozvinout svůj potenciál, pokud se korpusový výzkum bude provádět izolovaně, nezávisle na jiných výzkumných metodách“ (s. 280). Zároveň může být zajímavá pro badatele z oblasti psycholingvistiky, aplikované lingvistiky, sociolingvistiky nebo analýzy diskurzu, protože tam všude přináší triangulace s korpusově-lingvistickými metodami nové, originální úhly pohledu na problémy, které zkoumají. Kniha tak potvrzuje trend ke konvergenci korpusové lingvistiky a ostatních oblastí lingvistiky, který — jak autoři připomínají — popsali v r. 2012 McEnery a Hardie (s. 226): „Logickým vyústěním tohoto vývoje bude zánik korpusové lingvistiky jako samostatného oboru — situace, kdy korpusové metody budou (kde je to vhodné) používat všichni lingvisté.“

LITERATURA

BAKER, P. — J. EGBERT (eds.) (2016):
*Triangulating methodological approaches in
corpus-linguistic research*. London: Routledge.

MCENERY, T. — A. HARDIE (2012): *Corpus
linguistics: Method, theory and practice*.
Cambridge: Cambridge University Press.

Markéta Malá | Katedra anglického jazyka a literatury,
Pedagogická fakulta Univerzity Karlovy | Celetná 13, 116 39 Praha 1
ORCID ID: 0000-0003-3611-8433
Marketa.Mala@ff.cuni.cz

