

Historie matematické lingvistiky

Závěr

In: Blanka Sedlačíková (author): Historie matematické lingvistiky. (Czech). Brno: Akademické nakladatelství CERM v Brně, 2012. pp. 143–146.

Persistent URL: <http://dml.cz/dmlcz/402328>

Terms of use:

© Blanka Sedlačíková

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

Závěr

Cílem předkládané publikace bylo seznámit čtenáře s historií matematické lingvistiky, disciplíny stojící na pomezí mezi matematikou a jazykovědou (v současnosti ale i informatikou a umělou inteligencí). Počátky matematické lingvistiky jako samostatné disciplíny klademe do padesátých let 20. století, zpravidla za tento mezník bývá považován rok 1957, kdy se konal VIII. mezinárodní lingvistický kongres v Oslo. Ačkoliv se tedy jedná o obor poměrně mladý, není v silách jediného člověka popsat tak široké spektrum problémů, které jsou její součástí. Matematická lingvistika se tradičně rozděluje na tři části podle toho, kterých matematických metod při studiu jazyka využívá:

- 1) kvantitativní lingvistika využívá kvantitativní matematické metody, jako je například statistika, teorie pravděpodobnosti aj.;
- 2) lingvistika algebraická pracuje s nekvantitativními matematickými metodami (teorie množin, algebra, teorie grafů, logika apod.) a
- 3) lingvistika počítačová využívá ke studiu přirozených jazyků počítače (v podstatě jde o praktickou aplikaci dvou předchozích částí matematické lingvistiky).

I v případě, že bychom se zaměřili pouze na jednu z těchto tří oblastí matematické lingvistiky, nebylo by možné sestavit ucelený přehled této oblasti pouze jedním člověkem. Proto bylo nutné stanovit jistá kritéria výběru témat pro tuto práci. Protože o historii matematické lingvistiky nacházíme v dostupných pramenech (viz např. [6], [49], [58], [66], [77]) jen ojedinělé a často strohé zmínky, bylo hlavním záměrem práce tyto informace nejprve shromáždit, ale především doplnit o skutečnosti zcela nové, původní. Pokud to jen bylo možné, pracovala jsem v těchto případech s primárními prameny. Dalším kritériem, podle něhož jsem volila témata příhodná ke zpracování, byla důležitost dané problematiky v rámci matematické lingvistiky. V práci se proto objevují témata, která hrála zásadní roli ve vývoji matematické lingvistiky (teorie markovských řetězců, tzv. Zipfovy zákony, metoda glottochronologie aj.) A konečně jsem chtěla představit přínos českých vědců tomuto oboru. Kromě toho, že bylo třeba shromáždit a zpracovat poměrně velké množství rozmanitého materiálu od časopiseckých článků až po příručky encyklopedického charakteru, komplikoval situaci ten fakt, že autory těchto prací byly vědci z různých, na první pohled neslučitelných oborů (např. matematika, fyzika, lingvistika, informatika, textologie aj.).

Výsledkem této snahy je předkládaná práce, která je vedle Úvodu a Závěru tvořena dvěma hlavními kapitolami. Kapitola první je spíše obecného charak-

teru a jejím cílem je čtenářům představit ve stručnosti celý obor matematická lingvistika. Dozvídáme se tu o okolnostech vzniku matematické lingvistiky jako samostatné disciplíny, o jejích cílech, metodách či aplikacích. Rovněž jsou zde blíže představeny jednotlivé obory matematické lingvistiky, na něž bývá dnes běžně dělena. Konkrétně se jedná o lingvistiku kvantitativní, algebraickou a počítačovou, i když hranice mezi nimi se stále více a více stírají. Nejzásadnější je ovšem kapitola druhá, v níž je čtenář podle výše uvedených kritérií seznámen s některými výsledky z historie a „prehistorie“ matematické lingvistiky. Dočítáme se tu o poměrně jednoduchých aplikacích matematiky na jazyk (zpravidla na úrovni mystiky nebo hříčky), které byly využívány ve starověku a středověku. Ačkoliv v této době ještě zdaleka nemůžeme mluvit o matematické lingvistice, je zajímavé pozorovat, jakým způsobem se v této souvislosti formují počátky některých matematických disciplín, např. kombinatoriky aj. Navíc poznatky shromážděné v této kapitole se v tomto rozsahu nevyskytují v žádné z u nás běžně dostupných knih pojednávajících o matematické lingvistice. Skutečná matematická lingvistika (přesněji lingvistika kvantitativní) začíná až s vymezením pojmu frekvence jazykové jednotky (foném, písmeno, hláska, slovo, slabika aj.), proto je nemalá část práce věnována využití frekvence při tvorbě těsnopisných systémů, optimalizaci klávesnice psacího stroje, šifrování, kódování apod. Pozornost je zaměřena i na frekvenční a konkordanční slovníky, z nichž některé jsou blíže představeny. Samostatné oddíly jsou věnovány třem průkopníkům matematické lingvistiky, a to V. J. Bunjakovskému, A. A. Markovovi a americkému lingvistovi G. K. Zipfovi. Zejména kapitola pojednávající o ruském matematikovi A. A. Markovovi je poměrně rozsáhlá, neboť se snaží podrobně představit jeho zásadní práci [39], která byla vůbec první důslednou aplikací matematiky na jazyk a která navíc podnítila zájem matematiků o jazykovědnou problematiku. Rovněž sehrála významnou roli v matematické teorii markovských řetězců. Při řešení některých textologických problémů (atribuce textu, stylometrie, sestavování stemmatu) se významnou měrou také uplatňovala matematika, proto jsou tyto textologické problémy, které se zasloužily o rozvoj matematické lingvistiky, představeny podrobněji. Nelze opominout samozřejmě ani Shannonovu a Weaverovu teorii informace a některé její pojmy, zejména entropii a redundanci. Zajímavou aplikací matematiky je metoda nazývaná glottochronologie (nebo také lexikostatistika), která na základě statistiky zjišťuje stáří jazyka či jazyků. Její autor Morris Swadesh se inspiroval tzv. rozpadovým zákonem, jenž na základě rozpadu radioaktivních jader uhlíku určuje stáří organických látek. Poslední tři oddíly představují přínos českých vědců matematické lingvistice. Nejprve je pozornost věnována spisu *Janua linguarum reserata*, jehož autorem je Jan Amos Komenský. Dále je pozornost upřena na dvě statistiky uveřejněné roku 1831 v časopise *Krok*, prvním českém vědeckém časopise. První z nich je statistika hlásková, jejímž cílem je vyvrátit tvrzení, že čeština je, podobně jako němčina, jazyk těžký a tvrdý na výslovnost. Ukazuje se, že se jedná o jednu z prvních českých hláskových statistik, či spíše o jednu z prvních hláskových statistik v rámci slovanské jazykovědy vůbec. Druhá statistika zase porovnává počet slov v určitých jazycích a snaží se dokázat, že český jazyk je srovnatelný s ostatními jazyky co se týče slovní zásoby. Při pří-

právě této podkapitoly jsem náhodou narazila na zmínku o existenci statistiky, kterou si pro potřeby studia problematiky libozvučnosti sestavoval Jan Kollár. Nejednalo se sice o statistiku hláskovou či lexikální, ale nejspíše o statistiku grafémů, přesto ji lze podle dostupných informací časově vymezit rokem 1817 – byla by tedy o celých 14 let starší, než uvedené statistiky z časopisu *Krok*. Zdá se, že tato oblast matematické lingvistiky není zdaleka dostatečně prozkoumaná a že by bylo jistě zajímavé hledat další takové příklady užití matematických metod v lingvistice. A konečně poslední oddíl je věnován matematikovi, fyzikovi a astronomovi Augustinu Seydlerovi, který se svým článkem [57], v němž se pomocí počtu pravděpodobnosti snažil dokázat, že *Rukopisy královédvorský a zelenohorský* jsou novodobé podvrhy, zapojil v 80. letech 19. století do tzv. rukopisných sporů, které hýbaly celou tehdejší společností.

Doufám, že díky předložené práci mohou čtenáři zajímaví se o otázky týkající se historie matematické lingvistiky získat dobrý přehled o této problematice. Hlubší zájemci pak zde mohou nalézt (v čtených poznámkách pod čarou) množství odkazů na další literaturu (primární i sekundární). Kniha je psána takovým způsobem, aby byla srozumitelná a užitečná nejen matematikům (zejména z řad středoškolských učitelů a vysokoškolských studentů, kteří by zde mohli nalézt inspiraci pro svou práci), ale i zájemcům o jazyk či literaturu.

