

Historie matematické lingvistiky

Blanka Sedlačková

Úvod

In: Blanka Sedlačková (author): Historie matematické lingvistiky. (Czech). Brno: Akademické nakladatelství CERM v Brně, 2012. pp. 1–[4].

Persistent URL: <http://dml.cz/dmlcz/402311>

Terms of use:

© Blanka Sedlačková

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>

Úvod

Matematická lingvistika je vědecká disciplína, která v současné době znamená značný rozvoj. Její aplikace se staly běžnou součástí našeho života (zmiňme například jazykové databáze, elektronické slovníky, překlady textů pomocí počítačů, opravy překlepů v textových editorech aj.). Matematickou lingvistikou rozumíme pomezí disciplínu, která stojí na hranici mezi matematikou a jazykovědou (nově i umělou inteligencí a informatikou). Svými výsledky se do historie matematické lingvistiky zapsali ovšem nejen lingvisté či matematici, ale i fyzikové, literární vědci a další. Interdisciplinární charakter matematické lingvistiky poměrně komplikuje snahu shromáždit informace týkající se její historie – je třeba pátrat v celé řadě oborů (jazykověda, matematika, informatika, literární věda aj.). Podrobněji se historii matematické lingvistiky u nás věnuje pouze práce [6], další informace (zpravidla jen kusé) pak nacházíme roztroušeny v různých článcích, popularizačních příručkách, monografiích apod. (viz [49], [58], [66], [77]).

Cílem této publikace, která vznikla původně jako dizertační práce doktorského studijního programu Obecné otázky matematiky a informatiky na Přírodovědecké fakultě Masarykovy univerzity v Brně, bylo jednak shromáždit tyto izolované poznatky do jedné práce, jednak doplnit neúplné nebo stručné informace o nové skutečnosti. Ačkoliv je matematická lingvistika oborem poměrně mladým a její počátky klademe až do padesátých let 20. století, lze využívání matematických metod v lingvistice vysledovat hluboko do minulosti. První aplikace matematiky v lingvistice nacházíme dokonce již ve starověku a středověku. V tomto období se formovaly například počátky těsnopisu, významnou roli hrála kabala se svým kombinatorickým přístupem k jazyku apod. Proto stranou našeho zájmu nezůstaly ani tyto první pokusy, neboť mohou přinést nový pohled na problematiku matematické lingvistiky i jiných vědních oborů s ní souvisejících (např. kombinatorika, textologie aj.). Nemalá pozornost je konečně věnována těm výsledkům, kterými se do historie (či prehistorie) matematické lingvistiky zapsali čeští původci.

Celá práce je rozčleněna na dvě kapitoly. Nejprve je v první kapitole blíže představen obor matematická lingvistika – jeho vznik, cíle, metody, aplikace. Podrobněji je čtenář rovněž seznámen s jednotlivými částmi matematické lingvistiky, na něž bývá toto odvětví tradičně děleno, a to s lingvistikou kvantitativní (tou částí matematické lingvistiky, která využívá kvantitativní matematické metody), algebraickou lingvistikou (pracující s nekvantitativními mate-

matickými metodami, jako jsou například teorie grafů, kombinatorika, algebra, logika či teorie množin) a lingvistikou počítačovou (praktickou aplikací obou předchozích částí), i když hranice mezi jednotlivými částmi matematické lingvistiky nejsou samozřejmě vždy zcela zřetelné a dochází k jejich vzájemnému propojování. Například značný rozvoj počítačové lingvistiky v dnešní době by byl jen sotva možný bez využívání výsledků zbývajících dvou teoretických částí matematické lingvistiky, tedy lingvistiky kvantitativní a lingvistiky algebraické.

Těžiště práce představuje druhá kapitola, v níž jsou podle výše uvedených kritérií podrobně představeny některé partie z historie matematické lingvistiky. Nejprve je pozornost zaměřena na starověk a středověk, v nichž se většina aplikací matematiky na jazyk pohybovala víceméně na úrovni hry či mystiky. Přesto již zde můžeme vysledovat počátky takových matematických disciplín, jako je teorie kódování či kombinatorika. Blíže je ukázáno pojetí kombinatoriky v kabale a u významného středověkého myslitele Raimunda Lulla, jehož myšlenky ovlivnily například G. W. Leibnize. Důležitým mezníkem ve vývoji matematické (přesněji kvantitativní) lingvistiky je využívání pojmu frekvence (četnost) jazykového jevu (hlásky, písmena, slova, fonémy atd.). Ta hraje důležitou roli při sestavování frekvenčních a konkordančních slovníků, z nichž některé významné jsou v práci představeny blíže. Dále je ukázána řada dalších (především praktických) aplikací pojmu frekvence, a to například její využití při tvorbě těsnopisných systémů, optimalizaci klávesnice psacího stroje, šifrování, kódování (Morseova abeceda) apod. V následujících třech částech druhé kapitoly jsou představeni tři průkopníci využívání kvantitativních matematických metod při studiu jazyků. Ruský matematik V. J. Bunjakovský již v roce 1847 vůbec jako první upozornil na možnost využívání matematických metod v lingvistice. Další ruský matematik A. A. Markov svou teorií markovských řetězců nejen významně zasáhl do historie matematické lingvistiky, ale především do samotné matematiky. A konečně si představíme práci amerického lingvisty německého původu G. K. Zipfa, který se zabýval vztahy mezi frekvencí slov a jejich rankem a po němž bývá kvantitativní lingvistika někdy nazývána lingvistikou Zipfovou. V následujícím oddíle věnovaném textologii jsou podrobně rozebrány tři důležité textologické problémy, při jejichž řešení matematika sehrála důležitou roli. Konkrétně se jedná o tzv. stylometrii (neboli měření „stylu autora“), problematiku atribuce textu a sestavování stemmatu (čili schématu zachycujícího průběh textového procesu). Následuje oddíl zaměřený na teorii informace, v němž jsou vysvětleny některé základní pojmy tohoto odvětví, jako je například entropie a redundance, jež našly v matematické lingvistice rovněž své uplatnění. Dále je ukázána metoda nazývaná glottochronologie (též lexi-kostatistika), zajímavá aplikace matematiky, která se používá k určování stáří jazyků a která byla inspirována tzv. rozpadovým zákonem. Poslední tři oddíly druhé kapitoly se pak zabývají přínosem Čechů k „prehistorii“ využívání kvantitativních metod v lingvistice. Nejprve se zmíníme o Janu Amosi Komen-ském a jeho práci *Janua linguarum reserata*. Dále si ukážeme jednu z prvních hláskových statistik vůbec, která byla uveřejněna roku 1831 v časopise *Krok*, prvním českém vědeckém časopise. A na závěr se zastavíme u matematika, fyzika a astronoma Augustina Seydlera a jeho práce [57], v níž se pomocí po-

čtu pravděpodobnosti snažil dokázat, že *Rukopis královédvorský a zelenohorský* jsou padělky.

Text této práce vychází z dizertační práce, kterou jsem obhájila v prosinci roku 2010. Na prvním místě bych proto chtěla vyjádřit poděkování svému školiteli, doc. RNDr. Eduardu Fuchsovi, CSc., který byl nejen mým rádcem a kritikem, ale který dokázal v pravý čas nalézt slova podpory a povzbuzení. Velký dík patří rovněž oponentům, doc. RNDr. Zdeňku Karpíškovi, CSc. a doc. RNDr. Zdeňku Pospíšilovi, Dr., za pečlivé přečtení této práce a za jejich podnětné připomínky. Dále bych ráda poděkovala Nadaci Universitas, která umožnila vydání této knihy. Největší dík ovšem patří mému muži Markovi a našim třem dětem, kteří po celou dobu stáli po mém boku a byli mou největší motivací. Protože si mé poděkování zaslouží ještě celá řada dalších lidí a nerada bych některé z nich opomenula, nebudu jmenovat raději nikoho. Přesto věřím, že všichni, jichž se to týká, vědí, že velké díky patří i jim.

Blanka Sedlačíková

