

Univerzita Karlova
Filozofická fakulta
Ústav Českého národního korpusu

Petr Plecháč

**ATRIBUCE AUTORSTVÍ
BÁSNICKÝCH TEXTŮ**
AUTHORSHIP ATTRIBUTION
OF POETIC TEXTS

Autoreferát dizertační práce

Praha 2019

vedoucí dizertační práce: doc. Mgr. Václav Cvrček, Ph.D.

studijní program: Filologie

studijní obor: Matematická lingvistika

Obsah

ÚVOD.....	5
1 KVANTITATIVNÍ METODY URČOVÁNÍ AUTORSTVÍ.....	6
2 VERSOLOGICKÉ RYSY.....	8
3 EXPERIMENTY.....	10
4 APLIKACE.....	12
ZÁVĚR.....	17
CITOVANÁ LITERATURA.....	18
PŘEHLED ODBORNÉ ČINNOSTI.....	27

Úvod

Současná stylometrie, konkrétně ta její část zabývající se atribucí autorství, zažívá nebývalý rozmach. V posledních letech byla navržena, testována a aplikována celá řada textových charakteristik, na jejichž základě lze s pomocí metod multidimenzionální statistiky a/nebo strojového učení s velmi vysokou úspěšností odlišit texty psané různými autory. Ve většině případů se ovšem – z pohledu statistiky – jedná o tzv. řídké jevy nebo o jevy s distribucí LNRE – *large number of rare events* (např. četnosti jednotlivých slovních tvarů, slovních *n*-gramů, znakových *n*-gramů). Pro úspěšnou analýzu jsou tak obvykle potřeba poměrně rozsáhlé vzorky (tisíce až desetitisíce slov). V oblasti uměleckých textů se proto většina studií omezuje na rozsáhlé prozaické texty. Básnické texty – přestože neznámé nebo zpochybněné autorství provází nejčastěji právě je – zůstávají stranou.

Na druhou stranu, v poezii se lze opřít také o celou řadu binárních proměnných nebo proměnných, které mohou nabývat jen relativně malý počet možných hodnot, z oblasti versologie. Určitá část versologických charakteristik textu sice spadá do sféry vědomě volených a tudíž pro rozlišení autorství nevhodných – např. typ básnického metra (jamb, trochej, daktyl...), délka strofy, nebo typ rýmového schématu – u velké části z nich lze ale vědomou kontrolu předpokládat jen stěží (např. tendence při konkrétním výběru možných rytmických realizací metrického schématu, četnosti výskytu jednotlivých hlásek v rýmech). Přestože tyto charakteristiky bývají tradičně považovány za specifický rys autorské, případně alespoň dobové poetiky, při atribuci autorství byly dosud využívány pouze sporadicky.

Cílem této dizertační práce je testovat využitelnost takových versologických rysů při atribuci autorství na materiálu česky, německy, španělsky a anglicky psané poezie a aplikovat takový přístup na konkrétní případy sporného autorství.

Kapitola 1 je rozšířenou verzí oddílů (2) *Motivations* a (3) *History and related works* článku *Versification and authorship attribution. Pilot study on Czech, German, Spanish, and English poetry* (Plecháč–Bobenhausen–Hammerich 2018).

Kapitola 3.2 vychází částečně z textu *A collocation-driven method of discovering rhymes (in Czech, English, and French poetry)* (Plecháč 2018).

Kapitola 4.2 je přepracovanou verzí oddílů (2) *Atribuce „Kříže pod Petřínem“ provedená Pavlem Vašákem*, (3) *Kvadratická kosinova Delta* a (4) *Bootstrapovaná kosinová Delta* článku *Problém Barák–Neruda z pohledu současné stylometrie* (Plecháč–Flaišman 2017).

Kopie prohlášení spoluautorů jsou součástí dizertační práce.

1 Kvantitativní metody určování autorství

Kapitola podává přehled vývoje stylometrie, přičemž se omezuje na oblast zabývající se atribucí autorství. Vzhledem k tomu, že historie těchto metod byla podrobně zmapována v řadě přehledových publikací (zejm. Koppel–Schler–Argamon 2009; Stamatatos 2009; Juola 2006; Grieve 2005; Holmes 1998; 1994), jsou zde pouze stručně popsány hlavní milníky. Vedle toho jsou podrobněji rozvedeny:

- metody užívané v této práci, konkrétně:
 - míra Delta zavedená Johnem F. Burrowsem (2002; 2003) a její modifikace (Argamon 2008; Smith-Aldridge 2011);
 - *Support Vector Machine* (jedna z nejčastěji užívaných metod strojového učení při rozpoznávání autorství);
- práce zabývající se versologickými rysy, které bývají obvykle v přehledových publikacích přehlíženy, zejména práce členů *New Shakspeare Society* (např. Ingram 1874a; Fleay 1874a, Furnivall 1874b) a tzv. ruské versologické školy (Tomaševskij 1923/2008; Lotman–Lotman 1986; Šapir 1997, 1998; Tarlinskaja 1987, 2014);
- metoda iniciálních grafémů zavedená G. U. Yulem (1944), u níž upozorňujeme na překvapivé souvislosti s moderními multidimenzionálními metodami;
- stylometrické práce v českém prostředí (Seydler 1886; Jakobson 1935; Těšitelová–Nebeská–Králík 1976; Vašák 1966, 1980)

Závěry výše uvedených bodů můžeme shrnout následovně: od svého vzniku v 19. století vychází stylometrie z obecné představy, že autora díla lze určit na základě *podobnosti* mezi určitou *numerickou reprezentací* daného textu a numerickými reprezentacemi textů produkovaných kandidáty na jeho autorství.

Zatímco v rané stylometrii 19. století byly využívány jednoduché reprezentace jako např. délka slova měřená počtem znaků (Mendenhall), ve 20. století postupně přechází stylometrie k mnohem komplexnějším charakteristikám. Ruku v ruce s tímto vývojem se proměňovala i definice *podobnosti*: od jednoduchého (většinou optického) porovnání blízkosti dvou naměřených hodnot, přes inferenční statistiku, multidimenzionální analýzy, až po klasifikaci pomocí strojového učení.

Zmíněných charakteristik využívá současná stylometrie celou řadu: frekvence nejčastějších slov, znakových a slovních *n*-gramů, interpunkčních znaků... Jedna podstatná součást literárního stylu (jedné podstatné literární formy) ovšem zůstává zcela mimo její pozornost. Přestože u versologických rysů se tradičně předpokládá závislost na osobě autora (např. v češtině existuje bezpočet studií na téma „čím je Máchův jamb specifický“),

s jejich testováním a využitím pro potřeby atribuce básnických děl se v moderní stylometrii prakticky neseťkáme. K tomu připojme:

- Většina rysů užívaných ve stylometrii (např. frekvence slov, znakových a slovních *n*-gramů) představuje z pohledu statistiky řídké jevy, resp. přesněji jevy s distribucí LNRE (large number of rare events; srov. Baayen 2001). Ty lze smysluplně analyzovat pouze v rozsáhlých vzorcích textů (minimálně tisíce až desetitisíce slov). U básnických textů se ovšem s takovou situací v praxi setkáme zcela výjimečně – neznámé nebo zpochybněné autorství obvykle provází jednu báseň, případně několik básní, ale zřídka celou básnickou sbírku. Versologické rysy na druhou stranu bývají obvykle binární (např. přízvuková/nepřízvuková slabika) nebo mohou nabývat jen poměrně malého počtu možných hodnot (např. výskyt hlásek v rýmu) a mohou tak být smysluplně analyzovány i v mnohem menších vzorcích.
- Versologické rysy lze považovat za kontextově méně závislé než obvyklé rysy užívané ve stylometrii (frekvence slov, znakových a slovních *n*-gramů). Lexikon se v rámci textů produkovaných jedním autorem bude napříč žánry a tématy nutně značně odlišovat, zatímco můžeme předpokládat, že versologické charakteristiky budou zůstávat více méně stabilní. (Dodejme, že zde není řeč o vědomě volených proměnných jako je typ metra (jamb, trochej, daktyl...) nebo strofické schéma, resp. určitá pevná forma (sonet, rondel, limerik...), ale o způsobu realizace těchto abstraktních vzorců jazykovým materiálem, který je jen těžko racionálně kontrolovatelný/napodobitelný.)
- Lexikon ve veršovaných textech závisí nejen na autorovi a žánrových/tematických charakteristikách, ale je ovlivňován i zvoleným básnickým metrem (např. Forstall a Scheirer (2010) prokázali v latinských časoměrných textech asociaci mezi metrem a frekvencemi některých znakových bigramů).
- Někteří stylometrici doporučují k dosažení spolehlivějších výsledků kombinovanou analýzu několika rysů zároveň, např. vektory určené četnostmi nejfrekvencovanějších slov, znakových *n*-gramů a slovních *n*-gramů (srov. Mikros–Perifanos 2013; Eder 2011). Tyto rysy spolu ale navzájem značně korelují. Na druhou stranu můžeme předpokládat, že versologické rysy jsou na výše jmenovaných prakticky nezávislé. Kombinovaná analýza lexikálních a versologických rysů by tak měla přinést lepší výsledky než analýza těchto rysů samostatně.

Následující kapitoly se proto pokoušejí s použitím moderních atribučních metod testovat využitelnost versologických rysů coby ukazatelů autorství, a to na česky, německy, španělsky a anglicky psaném materiálu. Takový přístup byl (dle mého nejlepšího vědomí a svědomí) zatím testován pouze v omezené míře na malých vzorcích latinské (Forstall–Jacobson–Scheirer 2011) a staroarabské poezie (Al-Falahi–Ramdani–Bellafkih 2017), v obou případech s nepřilíš uspokojivými výsledky. V současné době probíhají přípravné práce na projektech, které se z tohoto pohledu zaměřují na středověkou nizozemskou poezii (Kestemont–Haverals 2018) a portugalsky psanou poezii (Mittmann–Pergherdos Santos 2019). První výsledky našeho testování byly publikovány v Plecháč–Bobenhausen–Hammerich 2018.¹

1 Nedlouho před dokončením této práce byl v *Nature Human Behaviour* publikován článek využívající versologické rysy pro rozlišení autorství staroanglické poezie (Neidorf et al. 2019).

2 Versologické rysy

Kapitola se zabývá základními charakteristikami veršovaných textů – básnickým rytmem, rýmem a eufonií – a možnostmi jejich kvantitativního modelování.

U **rytmické stránky** textu jsou diskutovány tři způsoby:

- Rytmičtý profil (vektor určený relativními četnostmi přízvuků na jednotlivých metrických pozicích). Jedná se o tradiční metodu rytmičtý analýzy (např. Mukařovský 1934/2001), v níž se ovšem ztrácí veškerá informace o kontextu jednotlivých slabik (srov. např. Dobritsyn 2016). Zároveň je aplikovatelný pouze u striktně sylabotónických versifikací (nemožnost modelování tzv. extrametrických pozic, „headless lines“, nepoužitelný u tónických veršů).
- Četnosti rytmičtý typů (vektor určený relativními četnostmi binárních reprezentací přízvučných/nepřízvučných slabik jednotlivých veršů). Metoda produkuje poměrně řídká data, a je proto vhodná spíše pro modelování rozsáhlejších textů.
- Četnosti rytmičtý n -gramů (metoda navržená pro potřeby této práce). Jedná se *de facto* o střední cestu mezi oběma výše zmíněnými metodami. Veršovaný text je reprezentován četnostmi rytmičtý realizací podřetězců metrické osnovy o arbitrárně zvolené délce n (viz OBR. 2.1).

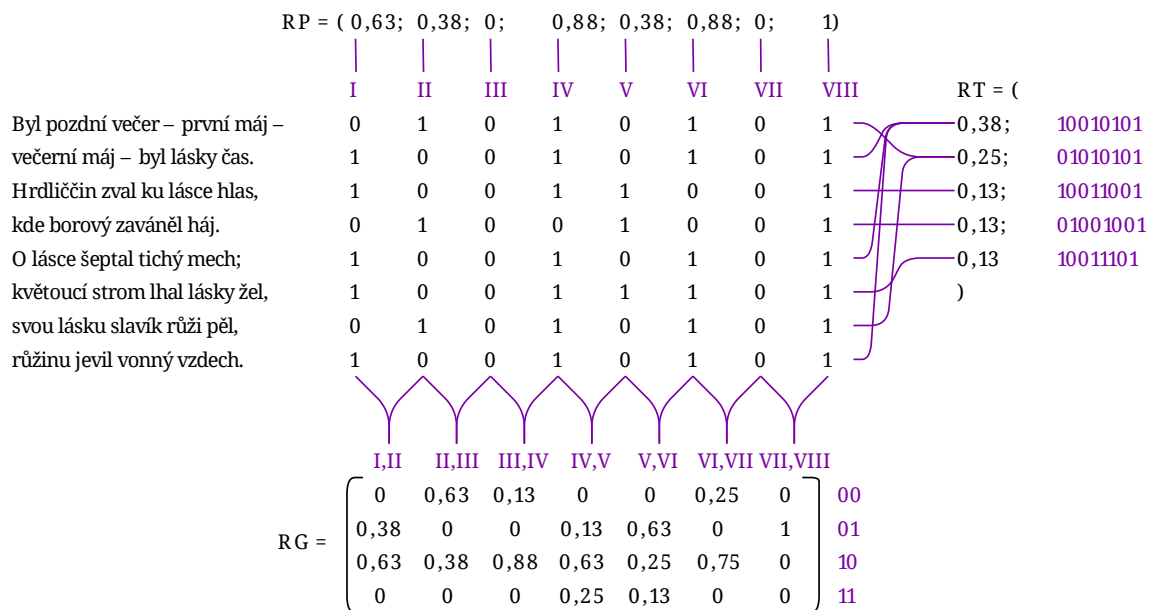
V dalších experimentech jsou sylabotónické (české, anglické) a sylabické (španělské) texty reprezentovány četnostmi rytmičtý 2-, 3- a 4-gramů, tónické texty (německé) četnostmi rytmičtý typů.

Rým je reprezentován jako neuspořádaná dvojice následujících charakteristik rýmových slov:

- morfologické rysy (v češtině slovní druh určený první pozicí morfologické značky, v ostatních jazycích celá značka produkovaná stochastickým taggerem TreeTagger),
- délka slova měřená počtem slabik,
- počet slabik za poslední přízvučnou slabikou,
- koda poslední slabiky,
- jádro poslední slabiky,
- pretura poslední slabiky + koda předposlední slabiky (pouze ženské rýmy),
- jádro předposlední slabiky (pouze ženské rýmy).

Textové vzorky reprezentujeme relativními četnostmi jednotlivých neuspořádaných dvojic v rámci dané kategorie.

Vzhledem k v práci popsaným výhradám nepracujeme se známou metodou měření „**eufoničnosti**“ textu navrženou Gabrielem Altmannem (1966a, 1966b), ale aproximujeme ji prostými četnostmi výskytu jednotlivých hlásek.



OBR. 2.1: Ilustrace metod rytmičké analýzy. (1) Rytmičký profil (RP), (2) četnosti rytmičkých typů (RT), (3) četnosti rytmičkých bigramů (RG).

3 Experimenty

Kapitola popisuje výsledky atribučních experimentů s 10 subkorpusy extrahovanými z původních korpusů sdružujícími autory narozené v určitém časovém rozmezí (CS1, CS2, CS3, DE1, DE2, DE3, ES1, ES2, EN1, EN2). Při stanovení těchto rozmezí jsme se řídili jednak požadavkem na dostatečný objem dat, jednak jsme se snažili rámcově vycházet z obvyklé literárněhistorické periodizace (např. CS1 obsahuje autory „obrozenecké“, CS2 obsahuje převážně autory spojované s „lumírovskou“, resp. „ruchovskou“ školou, EN1 obsahuje kanonické představitele anglického romantismu).

U každého subkorpusu byla provedena sada experimentů s modely založenými na:

- versologických rysech popsaných v kap. 2 (versologické modely),
- četnostech 150, 250 a 500 nejfrekventovanějších lemmat (lexikální modely),
- kombinaci obou sad rysů (kombinované modely).

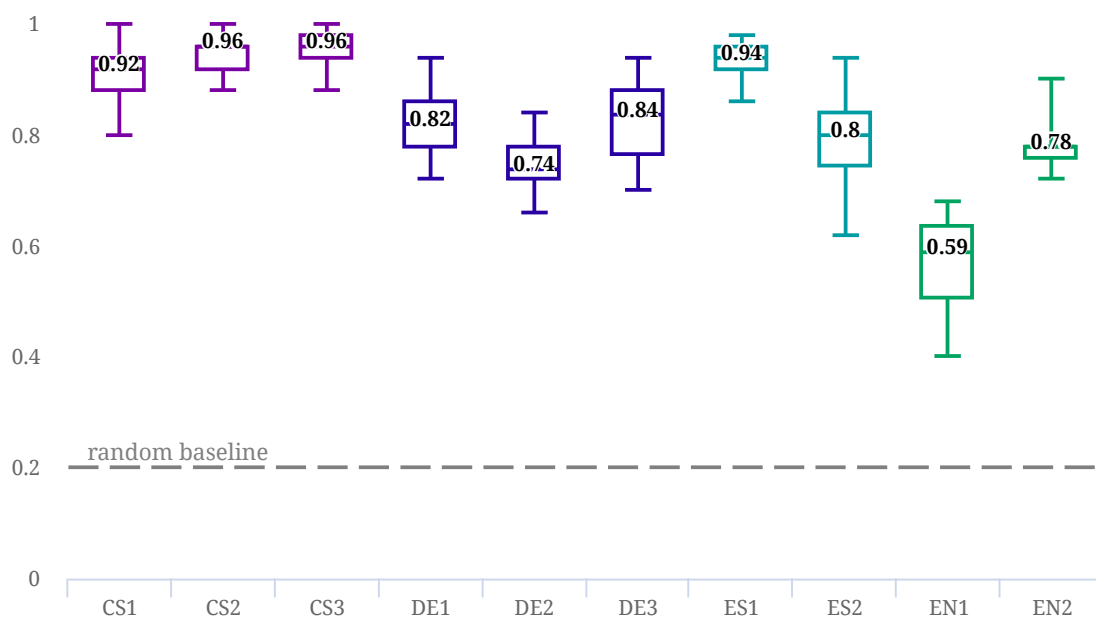
Průběh experimentu byl ve všech případech následující:

- subkorpus je segmentován na vzorky obsahující 100 veršů,
- náhodným výběrem je extrahováno 50 vzorků (po 10 vzorcích od 5 autorů),
- křížová validace modelu (jako klasifikátor zvolen *Support Vector Machine*),
- celý proces je 30krát zopakován vždy s novým náhodným výběrem vzorků (pro každý model je tak získáno 30 odhadů úspěšnosti).

Výsledky ukázaly, že versologické rysy lze považovat za smysluplný indikátor autorství:

- úspěšnost všech versologických modelů vysoce převyšuje hranici *random baseline* (= 0,2); viz OBR. 3.1,
- úspěšnost versologických modelů ojediněle převyšuje i úspěšnost lexikálních modelů,
- úspěšnost kombinovaných modelů je zpravidla vyšší než úspěšnost lexikálních modelů samotných,
- situace, kdy lexikální a versologický model predikuje stejného autora, znamená téměř vždy vyšší spolehlivost atribuce než situace, kdy je predikce provedena pouze lexikálním modelem.

Zároveň se ukázalo, že úspěšnost versologických a kombinovaných modelů je značně závislá na konkrétní versifikaci. Jako nejspolehlivější se versologické modely ukázaly u českého sylabotónického verše a (v o něco menší míře) španělského sylabického verše. Poměrně dobrých výsledků dosáhly versologické modely i v případě německého tónického verše.



OBR. 3.1: Výsledky křížových validací modelů založených na versologických rysech (30 iterací s náhodným výběrem vzorků).

Nejhorší výsledky vykázaly versologické modely u anglického sylabotónického verše. To jsme se pokusili vysvětlit jednak pevností anglického slovosledu, který omezuje možnosti individualizace rytmu, jednak málo rozvinutou flexí, což má za důsledek chudší rýmový slovník oproti ostatním jazykům. I přesto ale mohou versologické rysy najít v anglické versifikaci uplatnění. Jak ukazuje následující kapitola, přinejmenším v případě binární nebo ternární klasifikace a většího množství dat.

4 Aplikace

V poslední části práce jsou výše testované přístupy aplikovány na dva případy sporného autorství: (1) veršované drama *The Famous History of the Life of King Henry the Eighth* poprvé otištěné v Prvním foliu her Williama Shakespeara, u nějž se ovšem předpokládá i autorská účast Johna Fletchera, případně i dalších autorů, a (2) básně publikované pod jménem Josefa Baráka (1833–1883), u nichž existuje hypotéza, že jejich skutečným autorem je Jan Neruda.

William Shakespeare a John Fletcher: *Henry VIII*

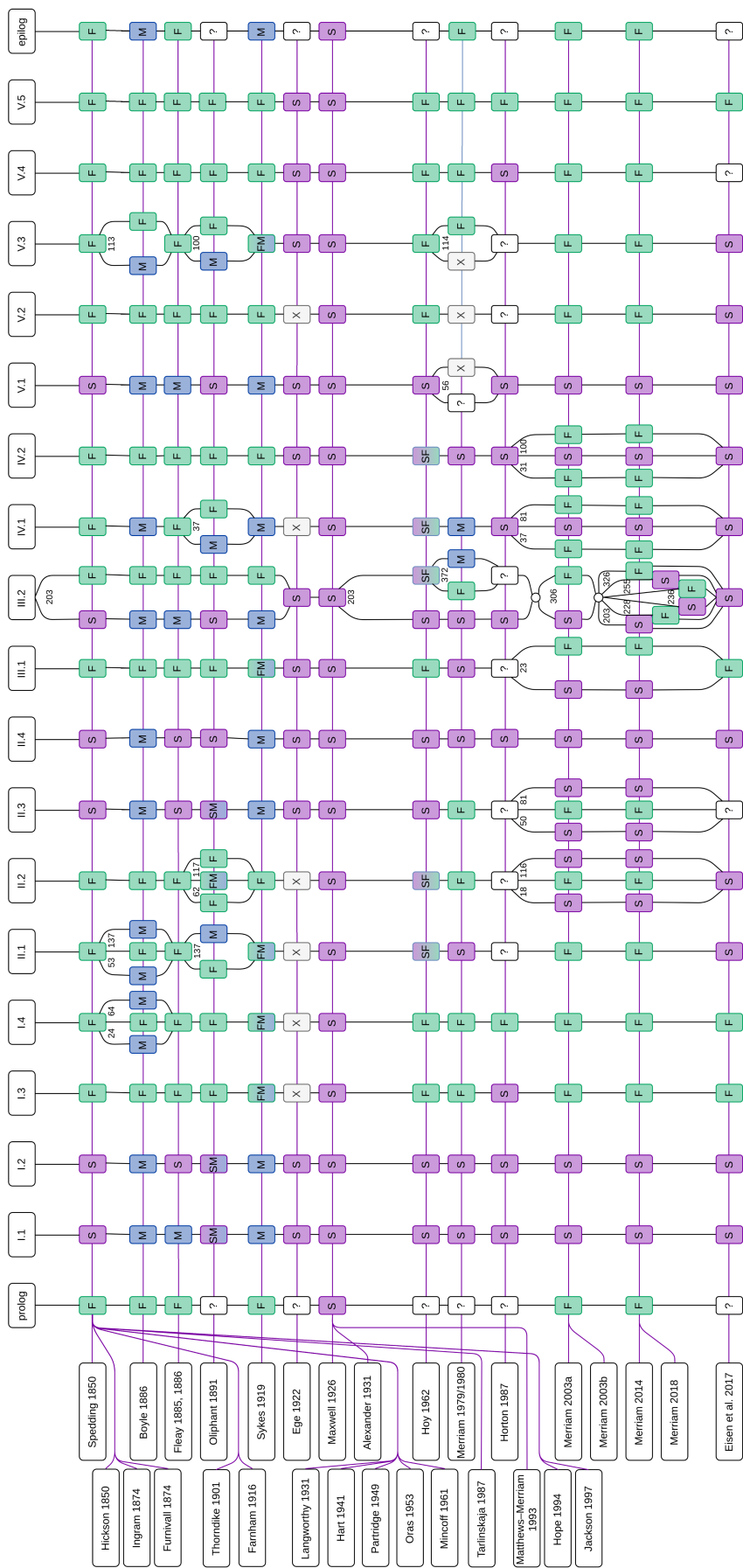
Veršované drama *The Famous History of the Life of King Henry the Eighth* (nadále *H8*) bylo až do poloviny 19. století považované za nesporné Shakespearovo dílo. V roce 1850 vyslovil James Spedding hypotézu, že části hry (konkrétně scény 1.3, 1.4, 2.1, 2.2, 3.1, 4.1, 4.2, 5.2–5.5 a druhou polovinu scény 3.2) napsal ve skutečnosti John Fletcher. Hlavním argumentem se stává nápadně vysoká variabilita poměru četností žensky zakončených veršů (např. „Till this time pomp was single, but now *married*“) a četností mužsky zakončených veršů (např. „The view of earthly glory: men might *say*“) napříč jednotlivými scénami. Od té doby byla publikována řada článků jak podporujících Speddingovu atribuci, tak obhajujících Shakespeara coby jediného autora, tak i formulujících nové hypotézy ohledně rozdělení scén mezi Shakespeara a Fletchera, případně atribuuující části díla dalším autorům: Philipu Massingerovi nebo dosud neznámému autorovi (viz OBR. 4.1).

Vlastní atribuci jsme provedli pomocí modelu natrénovaného na 4 Shakespearových, 4 Fletcherových a 3 Massingerových hrách (*Support Vector Machine*) založeného na četnostech 500 nejfrekventovanějších rytmických typů a četnostech 500 nejfrekventovanějších slovních tvarů. Křížová validace ukázala, že se jedná o model, který je s velmi vysokou úspěšností schopen rozpoznat autora.

Klasifikace jednotlivých scén *H8* ukázala:

- účast Philipa Massingera je velmi nepravděpodobná,
- je vysoce pravděpodobné, že text vzniknul ve spoluautorství Williama Shakespeara a Johna Fletchera,
- výsledky do značné míry korespondují se Speddingovou atribucí.

Vedle toho jsme se pokusili provést i atribuci, která nestojí na předpokladu, že rozdělení autorství mezi Williama Shakespeara a Johna Fletchera koresponduje s rozdělením scén. K tomuto účelu byla použita technika „klouzavé atribuce“ (*rolling attribution*; Eder 2016): text *H8* byl segmentován na 569 překrývajících se úseků (délka úseků: 100 veršů, posun okna: 5 veršů). Každý z těchto úseků byl klasifikován pomocí výše popsaného modelu a z výsledků klasifikace bylo pomocí Plattova škálování (*Platt's scaling*; Platt 1999) odhadnuto rozdělení pravděpodobnosti autorství mezi oba kandidáty.



OBR. 4.1: Přehled atribucí hry *Henry VIII*. S = William Shakespeare, F = John Fletcher, M = Philip Massinger, X = jiný, neznámý autor, ? = neurčeno. U scén se smíšeným autorstvím, kde atribuce neudává konkrétní rozdělení: FM = Shakespeare/Massinger, SF = Fletcher/Massinger, SM = Shakespeare/Fletcher. U scén se smíšeným autorstvím, kde atribuce udává konkrétní rozdělení, označují údaje mezi vertikálními spojnicemi číslo verše, od něhož začíná část připisovaná následujícímu autorovi (číslování veršů dle Pooler, C. K. (ed.) (1915). *The Famous History of the Life of King Henry VIII*. London: The Arden Shakespeare).

OBR. 4.2, 4.3, 4.4 zobrazují výsledky klouzavé atribuce na základě (1) 500 nejčtenějších rytmických typů, (2) 500 nejčtenějších slovních tvarů a (3) 1000 rozměrných vektorů kombinujících rysy (1) a (2). Každý sloupec odpovídá jednomu pětiverší a zobrazuje průměr pravděpodobností Shakespearova a Fletcherova autorství, které mu byly v rámci různých úseků přiděleny. Pro lepší přehlednost jsou hodnoty u Fletchera zobrazeny jako záporné (vzdálenost mezi vrcholy Shakespearova a Fletcherova sloupce je tak vždy rovna jedné).

I výsledky atribuce nezohledňující dělení na scény mluví do značné míry ve prospěch Speddingovy atribuce (včetně předpokládaného smíšeného autorství druhé scény třetího jednání). Hlavní rozdíl oproti Speddingově atribuci představují u našich modelů ambivalentní výsledky u obou scén čtvrtého jednání. Už sám Spedding ovšem vyslovil ohledně atribuce těchto scén Fletcherovi určité pochyby. Ostatní rozdíly jsou spíše marginální; buď svědčí ve prospěch drobných úprav, které vůči Speddingově atribuci navrhuje Thomas Merriam (2003a, 2003b, 2014, 2018), nebo pocházejí ze situací, kdy si výsledky jednotlivých modelů odporují.

Dodejme, že při aplikaci metody „klouzavé atribuce“ na 4 Shakespearovy a 4 Fletcherovy hry obsažené v trénovacím korpusu bylo 99,77 % úseků atribuováno svému skutečnému autorovi; výše popsanou analýzu *H8* tak můžeme pokládat za vysoce spolehlivou.

Autorství básní připisovaných Josefu Barákovi

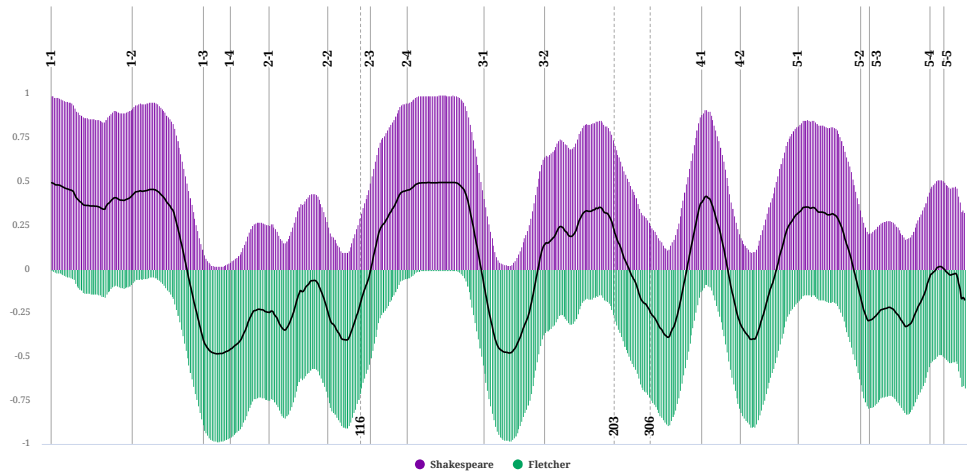
V 50. letech 20. století formuloval Oldřich Králík hypotézu, že autorem básní podepsaných Josefem Barákem (publikováno v letech 1858–1862) a prózy *Kříž pod Petřínem* byl ve skutečnosti jeho přítel Jan Neruda. Tato hypotéza se setkala se silnou kritikou např. Felixe Vodičky nebo Emanuela Macka, z našeho pohledu je ovšem nejpodstatnější reakce Pavla Vašáka opírající se o tehdejší stylometrické metody a vyvracející Nerudovo autorství *Kříže pod Petřínem*.

Vašákovu atribuci pokládáme za nespolehlivou z několika důvodů:

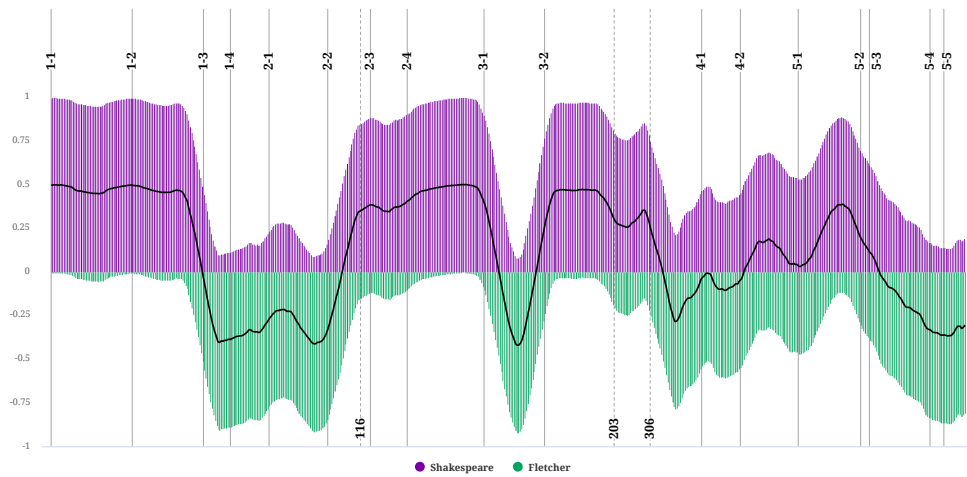
- nespolehlivé textové rysy; tři Vašákem analyzované charakteristiky (průměrná délka věty, průměrná délka slova a zastoupení slov různé délky) byly při rozsáhlém empirickém testování na anglicky psaném materiálu vyhodnoceny jako vůbec nejslabší, nejméně spolehlivé ukazatele autorství (Grieve 2007),
- nedostatečně velké vzorky (v jednom případě je text reprezentován průměrnou délkou pouhých 7 vět),
- vágní kritéria pro zamítnutí Nerudova autorství (stejným způsobem lze zamítnout Nerudovo autorství jím prokazatelně napsaných textů).

Otázku autorství textů připisovaných Josefu Barákovi jsme se proto rozhodli znovu otevřít, tentokrát analýzou objemnější veršované části díla.

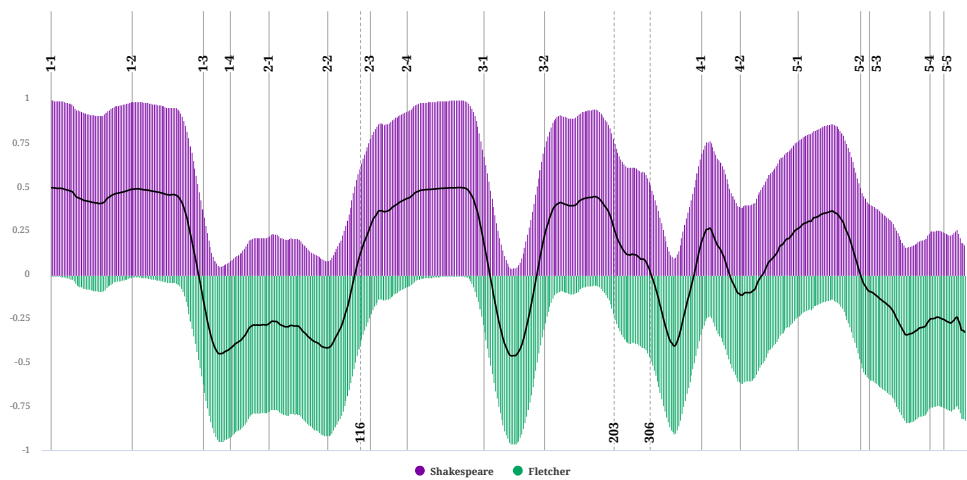
Nejvhodnějším klasifikátorem se ukázala být Smith-Aldridgova kosinová Delta (klasifikace metodou nejbližšího souseda měřeno kosinovou podobností) za použití n nejfrekventovanějších lemmat ($n \in \{50, 100, 150, \dots, 1000\}$) a rýmových charakteristik. Trénovací korpus byl sestaven ze vzorků básní Jana Nerudy a jeho současníků (Heyduk, Hálek, Jahn, Martinec, Šolc, Pflieger Moravský) pocházejících zhruba z období předpokládaného vzniku sporných básní.



OBR. 4.2: Klouzavá atribuce $H8$ na základě 500 nejčtetnějších rytmických typů.



OBR. 4.3: Klouzavá atribuce $H8$ na základě 500 nejčtetnějších slovních tvarů.



OBR. 4.4: Klouzavá atribuce $H8$ na základě 500 nejčtetnějších rytmických typů a 500 nejčtetnějších slovních tvarů.

Tento model, který lze považovat za poměrně spolehlivý (úspěšnost na trénovacím korpusu $\sim 0,9$), označil při různých hladinách n za nejbližšího souseda sporných básní vždy vzorek z Nerudových *Knih veršů*. Dodejme ale, že zvolená metoda sice vykazuje vysokou úspěšnost, ale pouze za předpokladu, že *texty skutečného autora jsou v trénovacím korpusu obsaženy*. Jinými slovy, nějaká položka z trénovacího korpusu bude vždy nejbližším sousedem sporných básní, ať už je skutečný autor v trénovacím korpusu přítomen, nebo ne. Výsledek „nelze určit, kdo je autorem sporných básní“ Delta nenabízí.

Toto omezení jsme se pokusili překonat aplikací tzv. bootstrapované Delty (Eder 2013), techniky, která byla navržena jako jedna z možných pojistek proti chybné atribuci v případě, kdy skutečný autor není v trénovacím korpusu přítomen. I tentokrát byl jako pravděpodobný autor vyhodnocen Jan Neruda.

Závěr

Cílem dizertační práce bylo testovat a aplikovat modely rozpoznávání autorství založené na versologických rysech. Shrňeme-li nejpodstatnější závěry, které testování přineslo, můžeme konstatovat:

- Úspěšnost rozpoznávání autorství na základě versologických rysů převyšuje v české, německé, španělské i anglické versifikaci několikanásobně hodnotu *random baseline*.
- Rozdíly mezi úspěšností rozpoznávání autorství v různých jazycích lze vysvětlovat zejm. jazykově-typologickými faktory.
- Úspěšnost versologických modelů ojediněle převyšuje úspěšnost obvyklých lexikálních modelů.
- Úspěšnost kombinovaných versologicko-lexikálních modelů je zpravidla vyšší než úspěšnost lexikálních modelů samotných.
- Situace, kdy lexikální a versologický model predikuje stejného autora, znamená téměř vždy vyšší spolehlivost atribuce než situace, kdy je predikce provedena pouze lexikálním modelem.

V další části práce byly pomocí versologických modelů, lexikálních modelů a kombinovaných versologicko-lexikálních modelů úspěšně analyzovány dva soubory textů se sporným autorstvím: (1) veršované drama *The Famous History of the Life of King Henry the Eighth* a (2) básně publikované pod jménem Josefa Baráka.

Má-li mít tato práce nějaký zcela jednoznačný závěr, pak ten, že versologické rysy představují relevantní stylometrický indikátor a (přestože jejich extrahování z textu je obvykle zdatelně komplikovanější než u tradičně využívaných rysů jako jsou četnosti slovních tvarů, lemmat, znakových či slovních *n*-gramů) mohou a měly by být při atribuci autorství básnických textů využívány.

Citovaná literatura

- Abney, S. (2007). *Semisupervised Learning for Computational Linguistics*. Boca Raton – London – New York: CRC.
- Al-Falahi, A. – Ramdani, M. – Bellafkih, M. (2017). Machine learning for authorship attribution in Arabic poetry. *International Journal of Future Computer and Communication* 6(2), 42–46.
- Alexander, P. (1931). *Conjectural History, or Shakespeare's Henry VIII. Essays and Studies* 16, 85–119.
- Altmann, G. (1966a). The measurement of Euphony. In J. Levý (ed.), *Teorie verše I. Sborník brněnské versologické konference, 13.–16. května 1964*. Brno: UJEP, 263–264.
- Altmann, G. (1966b). Binomial Index of Euphony for Indonesian Poetry. *Asian and African Studies* 2, 62–67.
- Antilla, A. – Heuser, R. (2016). Phonological and metrical variation across genres. *Proceedings of the Annual Meeting on Phonology 3*. Linguistic Society of America: Washington D.C.
- Argamon, S. (2008). Interpreting Burrows's Delta: Geometric and probabilistic foundations. *Literary and Linguistic Computing* 23(2), 131–147.
- Baayen, R. H. (2001). *Word Frequency Distributions*. Dordrecht: Kluwer Academic Publishers.
- Bělyj, A. (1910). *Symbolizm: Kniga Statej*. Moskva: Musaget.
- Bobenhausen, K. (2011). The Metricalizer. Automated metrical markup for German poetry. In C. Küper (ed.), *Current Trends in Metrical Analysis*. Frankfurt am Main et al.: Peter Lang, 119–131.
- Bobenhausen, K. – Hammerich, B. (2015). Métrique littéraire, métrique linguistique et métrique algorithmique de l'allemand mises en jeu dans le programme Metricalizer². *Langages* 199, 67–87.
- Boyle, R. (1886). *Henry VIII. An investigation into the origin and authorship of the play*. *Transactions of the New Shakspeare Society 1880–86*, 443–487.
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2, 121–167.
- Burrows, J. F. (1989). »An ocean where each kind«: Statistical analysis and some major determinants of literary style. *Computer and the humanities* 23(4–5), 309–321.
- Burrows, J. F. (2002) »Delta«: a measure of stylistic difference and a guide to likely authorship. *Literary and Linguistic Computing* 17(3), 267–287.
- Burrows, J. F. (2003). Questions of authorship: attribution and beyond. *Computers and the Humanities* 37(1), 5–32.

- Burrows, J. F. – Hassall, A. J. (1988). Anna Boleyn and the Authenticity of Fielding's Feminine Narrative. *Eighteenth Century Studies*, 21(4), 427–453.
- Craig, H. – Kinney, A. F. (2009). *Shakespeare, Computers and the Mystery of Authorship*. Cambridge: Cambridge University Press.
- Čech, R. – Popescu, I.-I. – Altmann, G. (2011). Euphony in Slovak lyric poetry. *Glottometrics* 22, 5–16.
- Čech, R. – Popescu, I.-I. – Altmann, G. (2014). *Metody kvantitativní analýzy (nejen) básnických textů*. Olomouc: Univerzita Palackého.
- Červenka, M. (1971). Osmislabičná řada ve verši a v próze. In *idem*, *Statistické obrazy verše*. Praha: ÚČSL AV ČR, 31 – 50.
- Červenka, M. (1998). Máchovo místo ve vývoji českého verše. *Česká literatura* 46(5), 485–489.
- Červenka, M. (2002). Hlásková instrumentace. In M. Červenka, M. Jankovič, M. Kubínová, M. Langerová (eds.), *Pohledy zblízka. Zvuk, význam, obraz*. Praha: Torst, 7–54.
- Červenka, M. (2006). *Kapitoly o českém verši*. Praha: Karolinum.
- Červenka, M. – Sgallová, K. (1978). Český verš. In Z. Kopczyńska, L. Pszczołowska (eds.), *Słowiańska metryka porównawcza I. Słownik rytmiczny i sposoby jego wykorzystania*. Wrocław et al.: Ossolineum, 45–94.
- Diederich, J. – Kindermann, J. – Leopold, E. – Paass, G. (2003). Authorship attribution with support vector machines, *Applied Intelligence* 19(1), 109–123.
- Dobritsyn, A. (2016). Rhythmic entropy as a measure of rhythmic diversity (The example of Russian iambic tetrameter). *Studia Metrica et Poetica* 3(1), 33–52.
- Eddy, H. (1887). The characteristic curve of composition. *Science* 9(216), 297.
- Eder, M. (2011). Style markers in authorship attribution. A cross-language study of the authorial fingerprint. *Studies in Polish Linguistics* 6, 99–114.
- Eder, M. (2013). Bootstrapping Delta: A safety net in open-set authorship attribution. *Digital Humanities 2013: Conference Abstracts*. Lincoln: University of Nebraska-Lincoln, 169–172.
- Eder, M. (2015). Does size matter? Authorship attribution, short samples, big problem. *Digital Scholarship in the Humanities* 30(2), 167–182.
- Eder, M. (2016). Rolling stylometry. *Digital Scholarship in the Humanities* 31(3), 457–469.
- Eder, M. (2017). Short samples in authorship attribution: A new approach. *Digital Humanities 2017: Conference abstracts*. Montreal: McGill University, 221–224.
- Ege, K. (1922). Shakespeare's Anteil an „Henry VIII“. *Shakespeare Jahrbuch* 58, 99–115.
- Eisen, M. – Riberio, A. – Segarra, S. – Egan, G. (2017). Stylometric analysis of early modern period English plays. *Digital Scholarship in the Humanities* 33(3), 500–528.
- Farnham, W. (1916). Colloquial contractions in Beaumont, Fletcher, Massinger and Shakespeare as a test of authorship. *Publications of the Modern Language Association of America* 31, 326–358.
- Fleay, F. G. (1874a). On the authorship of *The Taming of the Shrew*. *Transactions of the New Shakspeare Society* 1, 85–129.
- Fleay, F. G. (1874b). On the authorship of *Timon of Athens*. *Transactions of the New Shakspeare Society* 1, 130–194.

- Fleay, F. G. (1874c). A fresh confirmation of Mr. Spedding's division and date of the play of Henry VIII. *Transactions of the New Shakspeare Society* 1, appendix 23.
- Fleay, F. G. (1874d). Mr. Hickson's division of *The Two Noble Kinsmen*, confirmed by Metrical Tests. *Transactions of the New Shakspeare Society* 1, appendix 61–64.
- Fleay, F. G. (1876). *Shakespeare Manual*. London: Macmillan.
- Fleay, F. G. (1885). Mr. Boyle's theory as to „Henry VIII“. *Athenæum* 2994, 14. 3., 355.
- Fleay, F. G. (1886). *A Chronicle History of Life and Work of William Shakespeare*. London: John C. Nimmo.
- Forstall, C. W. – Jacobson, S. L. – Scheirer, W. J. (2011). Evidence of intertextuality: investigating Paul the Deacon's *Angustae Vitae*. *Literary and Linguistic Computing* 26(3), 285–296.
- Forstall, C. W. – Scheirer, W. J. (2010). A statistical stylistic study of Latin elegiac couplets. 2010 Chicago Colloquium on Digital Humanities and Computer Science Abstracts, Evanston: Northwestern University, 21–22. 11.
- Fucks, W. (1952). On mathematical analysis of style. *Biometrika* 39(1–2): 122–129.
- Furnivall, F. J. (1874a). The founder's prospectus of the New Shakspeare Society. *Transactions of the New Shakspeare Society* 1, appendix 6–7.
- Furnivall, F. J. (1874b). Another fresh confirmation of Mr. Spedding's division and date of the play of Henry VIII. *Transactions of the New Shakspeare Society* 1, appendix 24.
- Furnivall, F. J. (1874c). Mr. Hickson's division of *The Two Noble Kinsmen*, confirmed by the the stopt-line test. *Transactions of the New Shakspeare Society* 1, appendix 64–65.
- Gebauer, J. (1886). Potřeba dalších zkoušek rukopisu Královedvorského a Zelenohorského. *Athenaeum* 3, 152–168.
- Giesbrecht, E. – Evert, S. (2009). Is Part-of-Speech Tagging a Solved Task? An Evaluation of POS Taggers for the German Web as Corpus. In I. Alegria – I. Leturia – S. Sharoff (eds.), *Proceedings of the 5th Web as Corpus Workshop (WAC5)*. San Sebastian.
- Göhring, A. (2009). Spanish Expansion of a Parallel Treebank. Magisterská diplomová práce. Universität Zürich.
- Grieve, J. (2005). Quantitative authorship attribution: A history and an evaluation of techniques. Master thesis. Simon Fraser University.
- Grieve, J. (2007). Quantitative authorship attribution: an evaluation of techniques. *Literary and Linguistic Computing* 22(3), 251–270.
- Grzybek, P. (2014). The emergence of stylometry: prolegomena to the history of term and concept. In: Kroó, Katalin; Torop, Peeter (eds.), *Text within Text – Culture within Culture*. Budapest, Tartu: L'Harmattan, 58–75.
- Hart, A. (1941). Vocabularies of Shakespeare's Plays. *The Review of English Studies* 19(74), 128–140.
- Hearst, M. A. (1998). Support vector machines. *IEEE Intelligent Systems* 13(4), 18–28.
- Herdan, G. (1956). Chaucer's authorship of the *Equatorie of the Planetis*: The use of romance vocabulary as evidence. *Language* 32(2), 254–259.
- Hickson, S. (1850). Who wrote Shakespeare's Henry VIII. *Notes and Queries* 2, 198.
- Holmes, D. I. (1994). Authorship attribution. *Computers and the Humanities* 28(2), 87–106.

- Holmes, D. I. (1998). The evolution of stylometry in humanities scholarship. *Literary and Linguistic Computing* 13(3), 111–117.
- Hoover, D. L. (2003). Another perspective on vocabulary richness, *Computers and the Humanities* 37(2), 151–178.
- Hoover, D. L. (2004a). Testing Burrows's Delta. *Literary and Linguistic Computing* 19(4), 453–475.
- Hoover, D. L. (2004b). Delta Prime? *Literary and Linguistic Computing* 19(4), 477–495.
- Hope, J. (1994/2009). *The Authorship of Shakespeare's Plays*. Cambridge: Cambridge University Press.
- Horsmann, T. – Erbs, N. – Zesch, T. (2015). Fast or accurate? A comparative evaluation of PoS tagging models. In *Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology (GSCL-2015)*
- Horton, T. B. (1987). *The Effectiveness of the Stylometry of Function Words in Discriminating between Shakespeare and Fletcher*. Dizertační práce. University of Edinburgh.
- Hoy, C. (1957). The Shares of Fletcher and his collaborators in the Beaumont and Fletcher Canon II. *Studies in Bibliography* 9, 143–162.
- Hoy, C. (1962). The shares of Fletcher and his collaborators in the Beaumont and Fletcher Canon VII. *Studies in Bibliography* 15, 71–90.
- Ibrahim, R. – Plecháč, P. – Říha, J. (2013). *Úvod do teorie verše*. Praha: Akropolis
- Ingram, J. K. (1874). On the „weak endings“ of Shakspeare, with some account of the history of the verse tests in general. *Transactions of the New Shakspeare Society* 1, 442–456.
- Jackson, M. P. (1997). Phrase lengths in Henry VIII. *Shakespeare and Fletcher. Notes and Queries* 44(1), 75–80.
- Jackson, M. P. (2013). All Is True? Or Henry VIII: Authors and ideologies. *Notes and Queries* 60(3), 441–444.
- Jacobs, A. M. (2018). The Gutenberg English Poetry Corpus: Exemplary quantitative narrative analyses. *Frontiers in Digital Humanities* 5:5. doi: 10.3389/fdigh.2018.00005
- Jakobson, R. (1935). K časovým otázkám nauky o českém verši. *Slovo a slovesnost* 1(1), 46–53.
- Jakobson, R. (1938/1995). K popisu Máchova verše. In *idem: Poetická funkce*. Jinočany: H & H, 427–476
- Jannidis, F. – Pielström, S. – Schöch, C. – Vitt, T. (2015). Improving Burrows' Delta. An empirical evaluation of text distance measures. *Digital Humanities Conference 2015*, Sydney.
- Jirát, V. (1931–1932). Hudebnost Máchova rýmu. *Časopis pro moderní filologii* 18(1–2), 24–34 & 147–157.
- Juola, P. (2006). Authorship attribution. *Foundations and Trends in Informational Retrieval* 1(3), 233–334.

- Kestemont, M. (2014). Function words in authorship attribution. From black magic to theory? Proceedings of the 3rd Workshop on Computational Linguistics for Literature. Göteborg: ACL, 59–66.
- Kestemont, M. – Haverals, W. (2018). Metrical analyses of medieval Dutch poetry for the purpose of genre and authorship analysis [konferenční příspěvek]. Plotting Poetry II: Bringing Deep Learning to Computational Poetry Analysis. Berlin: Freie Universität Berlin, 12.–14. 9. 2018.
- Koppel, M. – Schler, J. (2004). Authorship verification as a one-class classification problem. Proceedings of the 21st International Conference on Machine Learning. New York: ACM, 62–68.
- Koppel, M. – Schler, J. – Argamon, S. (2009). Computational methods in authorship attribution. Journal of the Association for Information Science and Technology 60(1), 9–26.
- Králík, O. (1956). Svědectví Anny Holinové. Host do domu 3(3), 107–109.
- Králík, O. (1957). Neruda nebo Barák? Literární noviny 6(47), 6.
- Králík, O. (ed.) (1958). Z doby Májů. Olomouc: Krajské nakladatelství v Olomouci.
- Kubát, M. (2016). Kvantitativní analýza žánrů. Ostrava: Ostravská univerzita.
- Kuboň, V. – Lopatková, M. – Hercig, T. (2016). Searching for a Measure of Word Order Freedom. In B. Brejová (ed.), Proceedings of the 16th ITAT Conference Information Technologies – Applications and Theory. CreateSpace, 11–17.
- Langworthy, C. (1931). A verse-sentence analysis of Shakespeare's Plays. Publications of the Modern Language Association of America 46(3), 738–751.
- Larsen, W. A. – Rencher, A. C. – Layton, T. (1980). Who wrote the Book of Mormon? An analysis of wordprints. BYU Studies Quarterly 20(3), 225–251.
- Levý, J. (1962). Izochronie taktů a izosylabismus jako činitelé básnického rytmu. Slovo a slovesnost 23(1–2), 1–8 & 83–94.
- Lotman, J. M. – Lotman, M. (1986). Vokrug desjatoj glavy „Evgenija Onegina“. In N. N. Petrunina (ed.), Puškin: Issledovanija i materialy XII. Moskva/Leningrad: Nauka, 124–151.
- Macek, E. (1974). Králík kontra Barák. Literární archiv PNP VIII–IX, 495–580.
- Malone, Edmond. (1787/1803). A dissertation on parts one, two and three of Henry the Sixth tending to show that those plays were not written originally by Shakespeare. In Plays of William Shakespeare 14. London, 219–263.
- Matthews, R. A. J. – Merriam, T. V. N. (1993). Neural computation in stylometry I. An application to the works of Shakespeare and Fletcher. Literary and Linguistic Computing 8(4), 203–209.
- Mascol, C. (1888a). Curves of pauline and pseudo-pauline style I. Unitarian Review 30, 452–460.
- Mascol, C. (1888b). Curves of pauline and pseudo-pauline style II. Unitarian Review 30, 539–546.
- Matthews, R. – Merriam, T. (1993). Neural computation in stylometry I; An application to the works of Shakespeare and Fletcher. Literary and Linguistic Computing 8(4), 203–209.

- Maxwell, B. (1923). Fletcher and Henry the Eighth. *The Manly Anniversary Studies in Language and Literature*. Chicago: University of Chicago, 104-112.
- Maxwell, B. (1926). Sidelights on Shakespeare; Sidelights on Elizabethan Drama [recenze]. *Modern Philology* 23(3), 365–372.
- Mendenhall, T. C. (1887). The characteristic curves of composition. *Science* 9(214), 237–249.
- Mendenhall, T. C. (1901). A mechanical solution to a literary problem. *Popular Science Monthly* 9, 97–110.
- Merriam, T. (1979). What Shakespeare wrote in ‚Henry VIII‘: part one. *The Bard* 2(3), 81–94.
- Merriam, T. (1980). What Shakespeare wrote in ‚Henry VIII‘: part two. *The Bard* 2(4), 111–118.
- Merriam, T. (2000). Edward III. *Literary and Linguistic Computing* 15(2), 157–186.
- Merriam, T. (2003a). Taylor’s method applied to Shakespeare and Fletcher. *Notes and Queries* 50(4), 419–423.
- Merriam, T. (2003b). Though this be supplementarity, yet there is method in’t. *Notes and Queries* 50(4), 423–426.
- Merriam, T. (2005). *The Identity of Shakespeare in Henry VIII*. Tokyo: Renaissance Institute.
- Merriam, T. (2014). A reply to ‚All is True or Henry VIII: Authors and ideologies‘. *Notes and Queries* 61(2), 253–256.
- Merriam T. (2018). Henry VIII, All Is True?. *Notes and Queries* 65(1), 84–88.
- Mikros, G. K. – Perifanos, K. A. (2013). Authorship attribution in Greek tweets using author’s multilevel *n*-gram profile. In *Papers from the 2013 AAAI Spring Symposium. "Analyzing Microtext"*, 25–27 March 2013. Palo Alto: AAAI Press, 17–23.
- Mincoff, M. (1961). Henry VIII and Fletcher. *Shakespeare Quarterly* 12(3), 239–260.
- Mittmann, A. – Pergher, P. H. – dos Santos, A. L. 2019. What rhythmic signature says about poetic corpora. In P. Plecháč – B. Scherr – T. Skulacheva – H. Bermúdez-Sabel – R. Kolár (eds.), *Quantitative Approaches to Versification*. Praha: Ústav pro českou literaturu AV ČR, 153–172.
- de Morgan, A. (1851/1882). To Rev. W. Heald, Aug. 18, 1851 [dopis]. In S. E. de Morgan (ed.), *Memoir of Augustus de Morgan*. London: Longmans, Green, & co., 214–216.
- Mosteller, F. – Wallace D. L. (1964). *Inference and Disputed Authorship*. Reading: Addison-Wesley.
- Mukařovský, J. (1934/2001). *Obecné zásady a vývoj novočeského verše*. In idem *Studie II*, Brno: Host, 116–198.
- Navarro-Colorado, B. (2015). A computational linguistic approach to Spanish golden age sonnets: metrical and semantic aspects. *Computational Linguistics for Literature NAACL 2015*, Denver (Co).
- Navarro-Colorado, B. (2017). A metrical scansion system for fixed-metre Spanish poetry. *Digital Scholarship in the Humanities* 33(1), 112–127.
- Navarro-Colorado, B. – Ribes-Lafoz, M. – Sánchez, N. (2016). Metrical annotation of a large corpus of Spanish sonnets. Representation, scansion and evaluation. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož.

- Neidorf, L. – Krieger, M. S. – Yakubek, M. – Chaudhuri, P. – Dexter, J. P. (2019). Large-scale quantitative profiling of the Old English verse tradition. *Nature Human Behaviour* 8. 4. 2019, doi: 0.1038/s41562-019-0570-1.
- Oliphant, E. H. C. (1891). The Works of Beaumont and Fletcher. *Englische Studien* 15, 321–360.
- Oliphant, E. H. C. (1927). *The Plays of Beaumont and Fletcher: An Attempt to Determine their Respective Shares*. New Haven.
- Oras, A. (1953). »Extra monosyllables« in Henry VIII and the problem of authorship. *Journal of English and Germanic Philology* 52, 198–213.
- Partridge, A. C. (1949). *The Problem of Henry VIII Reopened. Some Linguistic Criteria for the Two Styles Apparent in the Play*. Cambridge: Bowes & Bowes.
- Pelillo, M. (2014). Alhazen and the nearest neighbor rule. *Pattern Recognition Letters* 38, 34–37.
- Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers* 10(3), 61–74.
- Plecháč, P. (2016). Czech verse processing system KVĚTA: Phonetic and metrical components. *Glottotheory* 7, 159–174.
- Plecháč, P. (2018). A collocation-driven method of discovering rhymes (in Czech, English, and French poetry). In M. Fidler – V. Cvrček (eds.), *Taming the Corpus. From Inflection and Lexis to Interpretation*. Cham: Springer, 79–95.
- Plecháč, P. – Bobenhausen, K. – Hammerich, B. (2018). Versification and authorship attribution. Pilot study on Czech, German, Spanish, and English Poetry. *Studia Metrica et Poetica* 5(2), 29–54
- Plecháč, P. – Flaišman, J. (2017). Problém Barák–Neruda z pohledu současné stylometrie. *Česká literatura* 65(5), 743–769.
- Plecháč, P. – Kolár, R. (2015). The Corpus of Czech Verse. *Studia Metrica et Poetica* 2(1), 107–118.
- Popescu, I.-I. – Altmann, G. – Grzybek, P. – Jayaram, B. D. – Köhler, R. – Krupa, V. – Mačutek, J. – Pustet, R. – Uhlířová, L. – Vidya, M. N. (2009). *Word Frequency Studies*. Berlin: De Gruyter.
- Popescu, I.-I. – Lupea, M. – Tatar, D. – Altmann, G. (2015). *Quantitative Analysis of Poetic Texts*. Berlin: De Gruyter.
- Rybickí, J. – Eder, M. (2011). Deeper Delta across genres and languages: do we really need the most frequent words? *Literary and Linguistic Computing* 26(3), 315–321.
- Rygl, J. (2014). Automatic adaptation of author's stylometric features to document types. In P. Sojka et al. (eds.), *Text, Speech, and Dialogue. 17th International Conference*. Berlin – Heidelberg: Springer, 59–61.
- Rygl, J. (2016). Building Corpora for Stylometric Research. In P. Sojka et al. (eds.), *Text, Speech, and Dialogue. 19th International Conference*. Berlin – Heidelberg: Springer, 20-27,
- Sedlačiková, B. (2012). *Historie matematické lingvistiky*. Brno: CERM.

- Segarra, S. – Eisen, M. – Riberio, A. (2013). Authorship attribution using function words adjacency networks. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5563–5567.
- Seydler, A. (1886). Počet pravděpodobnosti v přítomném sporu. *Athenaeum* 3, 299–308.
- Sherman, L. A. (1888). Some observations upon sentence-length in English prose. *The University of Nebraska Studies* 1(4), 337–366.
- Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of International Conference on New Methods in Language Processing*. Manchester.
- Skoumalová, H. (2011). Porovnání úspěšnosti tagování korpusu. In V. Petkevič – A. Rosen (eds.), *Korpusová lingvistika Praha 2011/3. Gramatika a značkování korpusů*. Praha: Nakladatelství Lidové Noviny, 199–207.
- Smith, P. W. H. – Aldridge, W. (2011). Improving authorship attribution: Optimizing Burrows' Delta method. *Journal of Quantitative Linguistics* 18(1), 63–88.
- Spedding, J. (1850). Who wrote Shakespeare's Henry VIII. *The Gentleman's Magazine*, 115–123.
- Spoustová, D. – Hajič, J. – Votrubec, J. – Krbec, P. – Květoň, P. (2007). The best of two worlds: Cooperation of statistical and rule-based taggers for Czech. *Proceedings of the Workshop on Balto-Slavonic Natural Language Processing*. ACL, 67–74.
- Stamatatos, E. (2009). A Survey of modern authorship attribution methods. *Journal of the Association for Information Science and Technology* 60(3), 538–556.
- Sykes, H. D. (1919). *Sidelights on Shakespeare*. Stratford-upon-Avon: The Shakespeare Head Press.
- Šapir, M. (1997). Fenomen Batenkova i problem mistifikacii (lingvistichovedčeskij aspekt 1–2). *Philologica* 4, 85–144.
- Šapir, M. (1998). Fenomen Batenkova i problem mistifikacii (lingvistichovedčeskij aspekt 3–4). *Philologica* 5, 49–132.
- Tabata, T. (2012). Approaching Dickens' style through random forests. *Digital Humanities 2012: Conference Abstracts*. Hamburg: Universität Hamburg, 388–391.
- Tarlinskaja, M. (1987). *Shakespeare's Verse: Iambic Pentameter and the Poet's Idiosyncrasies*. New York: Peter Lang.
- Tarlinskaja, M. (2014). *Shakespeare and the versification of English Drama, 1561–1642*. Farnham et al.: Ashgate.
- Těšitelová, M. – Nebeská, I. – Králík, J. (1976). On the quantitative characteristics of the Czech texts of disputed authorship. *Prague Studies in Mathematical Linguistics* 5, 119–147.
- Thorndike, A. H. (1901). *The Influence of Beaumont and Fletcher on Shakespeare*. Worcester: Oliver B. Wood.
- Tomaševskij, B. V. (1923/2008). Pjativopnyj jamb Puškina. In *Izbrannyje raboty o stiche*. Moskva & Sankt Peterburg: Akademia, 140–242.
- Vašák, P. (1966). Statistika a sporné autorství. *Slovo a slovesnost* 27(4), 364–370.
- Vašák, P. (1972). Metody ustanovenija spornogo avtorstva (problema Neruda — Barák). *Prague Studies in Mathematical Linguistics III (Praha)*, s. 143–162.
- Vašák, P. (1974). Barákovo autorství jako problém. *Česká literatura* 22(2), 145–155.

- Vašák, P. (1980). *Metody určování autorství*. Praha: Academia.
- Vodička, F. (1958). Ještě jednou: Neruda nebo Barák. *Literární noviny* VII, 25. 1., 6.
- Vickers, B. (2004). *Shakespeare, Co-Author. A Historical Study of Five Collaborative Plays*. Oxford: Oxford University Press.
- Weber, H. (1812). Observations on the participation of Shakespeare in *The Two Noble Kinsmen*. In idem (ed.), *The Works of Beaumont and Fletcher in Fourteen Volumes* 13. Edinburgh: J. Ballantyne & Co., 151–169.
- Williams, C. B. (1975). Mendenhall's studies of word-length distribution in the works of Shakespeare and Bacon. *Biometrika* 62(1), 207–212.
- Wimmer, G. – Altmann, G. – Hřebíček, L. – Ondrejovič, S. – Wimmerová, S. (2003). *Úvod do analýzy textov*. Bratislava: VEDA.
- Wolf, F. (1928). Použití počtu pravděpodobnosti k identifikaci textu. Inaugurace rektorů v Brně 1928/29 a 1929/30. Brno: Masarykova Universita, s. 99–105.
- Woodward, R. H. – Goldsmith, P. L. (1964). *Cumulative sum techniques*. Edinburgh: Oliver and Boyd.
- Yule, G. U. (1939). On sentence-length as a statistical characteristic of style in prose, with application to two cases of disputed authorship. *Biometrika* 30, 363–390.
- Yule, G. U. (1944). *The Statistical Study of Literary Vocabulary*. Cambridge: CUP.
- Zhao, Y. – Zobel, J. (2005). Effective and scalable authorship attribution using function words. In G. G. Lee et al. (eds.), *Information Retrieval Technology. AIRS 2005. Lecture Notes in Computer Science*. Berlin – Heidelberg: Springer, 174–189.
- Zipf, G. K. (1932). *Selected Studies on the Principle of Relative Frequency in Language*. Cambridge: HUP.

Přehled odborné činnosti

Publikace

Knižní publikace

- (1) Plecháč, P. – Scherr, B. P. – Skulacheva, T. – Bermúdez-Sabel, H. – Kolár, R. (eds.) (2019). *Quantitative Approaches to Versification*. Prague: ICL CAS.
- (2) Plecháč, P. – Kolár, R. (2017). *Kapitoly z korpusové versologie*. Praha: Akropolis.
- (3) Ibrahim, R. – Plecháč, P. – Říha, J. (2013). *Úvod do teorie verše*. Praha: Akropolis.

Články, kapitoly v knihách

- (4) Plecháč, P. – Birnbaum, D. J. (2019). Assessing the reliability of stress as a feature of authorship attribution in syllabic and accentual syllabic verse. In P. Plecháč, B. P. Scherr, T. Skulacheva, H. Bermúdez-Sabel, R. Kolár (eds.), *Quantitative Approaches to Versification*. Prague: ICL CAS, 201–210.
- (5) Místecký, M. – Čech, R. – Plecháč, P. (2019). In search of a euphony unit: A case study in the Czech 1830s poetry. *Slovo a slovesnost* 80(1), 27–38.
- (6) Plecháč, P. – Bobenhausen, K. – Hammerich, B. (2018). Versification and authorship attribution. Pilot study on Czech, German, Spanish, and English poetry. *Studia Metrica et Poetica* 5(2), 29–54.
- (7) Plecháč, P. (2018). A collocation-driven method of Discovering Rhymes (in Czech, English, and French Poetry). In M. Fidler, V. Cvrček (eds.), *Taming the Corpus. From Inflection and Lexis to Interpretation*. Cham: Springer, 79–95.
- (8) Plecháč, P. – Flaišman, J. (2018). Atribuce textu z pohledu současné stylometrie. In M. Kosák, J. Flaišman (eds.), *Editologie (od náčrtu ke knize)*. Praha: ÚČL, 147–178. [přetisk (13)]
- (9) Kolár, R. – Plecháč, P. (2018). Ještě k Seifertovu ranému volnému verši. *Česká literatura* 66(4), 553–562.
- (10) Kolár, R. – Plecháč, P. (2018). Tomanovy pohádky krve a Korpus českého verše. *Český jazyk a literatura* 68(4), 187–190.
- (11) Kolár, R. – Plecháč, P. (2018). Strofické kombinace v české poezii přelomu 19. a 20. století. *Litikon* 3(1), 95–102.
- (12) Plecháč, P. – Hlaváčová, J. – Merthová, K. – Kolár, R. (2017). Distribuce předpon v českém sylabotónickém trocheji. *Slovo a slovesnost* 78(4), 322–332.

- (13) Plecháč, P. – Flaišman, J. (2017). Problém Barák–Neruda z pohledu současné stylometrie. *Česká literatura* 65(5), 743–769.
- (14) Kolár, R. – Plecháč, P. (2017). Častotnost častej reči v češskoj poeziji. *Trudy Instituta ruskogo jazyka im. V. V. Vinogradova XIV: Slavjanskij stich*, 113–125.
- (15) Plecháč, P. – Kolár, R. (2017). Točnyje metody v češskom stichovedenii. *Trudy Instituta ruskogo jazyka im. V. V. Vinogradova XI: Slavjanskij stich*, 31–43.
- (16) Plecháč, P. (2016). Czech Verse Processing System KVĚTA: Phonetic and Metrical Components. *Glottology* 7(2), 159–174.
- (17) Kolár, R. – Plecháč, P. (2016). Rým a strofika v české poezii 19. století. *Český jazyk a literatura* 67(2), 59–66.
- (18) Kolár, R. – Plecháč, P. (2016). Hus v české poezii 19. století. In M. Jaluška, D. Soukup (eds.), *Jan Hus, husitství a husitské války a jejich dopad na českou kulturu*. Praha: Akropolis, 135–145.
- (19) Plecháč, P. – Kolár, R. (2015). The Corpus of Czech Verse. *Studia Metrica et Poetica* 2(1), 107–118.
- (20) Kolár, R. – Plecháč, P. (2015). Databáze českých meter a výzkum českého verše 19. století. *Česká literatura* 63(2), 236–246.
- (21) Plecháč, P. – Kolár, R. (2015). Korpus českého verše. In J. Hlaváčová (ed.), *WDH 2015 – Sborník Semináře o digitálních zdrojích a službách ve společenských a humanitních vědách*. Praha: Univerzita Karlova, 74–78.
- (22) Ibrahim, R. – Plecháč, P. (2014). La théorie du vers et le Cercle linguistique de Prague. *La Linguistique* 50(2), 101–114.
- (23) Plecháč, P. – Říha, J. (2014). Measuring Euphony. In G. V. Vekshin (ed.), *Methodology and Practices of Russian Formalism: Brikovskij sbornik II*. Moskva: Azbukovnik, 194–199.
- (24) Plecháč, P. – Ibrahim, R. (2014). Frequency of Parts of Speech in Adolf Heyduk's Poetry. In T. Kubíček, A. Lass (eds.), *Roman O. Jakobson; A Work in Progress*. Olomouc: VUP, 133–144.
- (25) Plecháč, P. – Ibrahim, R. (2013). Phonological and Morphological Means Compensating for Non-Metricality in 19th Century Czech Verse. *Prace Filologiczne* 59(3), 31–50.
- (26) Plecháč, P. – Ibrahim, R. (2013). Automatyczna analiza wiersza: punkt wyjścia. In W. Sadowski (ed.), *Potencjał wiersza*. Warszawa: IBL, 273–280.
- (27) Ibrahim, R. – Plecháč, P. (2012). Teorie verše ve stopovém množství. In A. Vermířovský (ed.), *Poetický Cikháj v Brně 2010*. Brno: Tribun EU, 155–163.
- (28) Plecháč, P. (2012). Miroslav Červenka a generativní model metrické normy českého sylabotónického verše. *Česká literatura* 60(3), 398–408.
- (29) Ibrahim, R. – Plecháč, P. (2011). Toward Automatic Analysis of Czech Verse. In B. P. Scherr, J. Bailey, E. V. Kazartsev (eds.), *Formal Methods in Poetics*. Lüdenscheid: RAM-Verlag, 295–305.
- (30) Plecháč, P. (2009). Přízvuková kontura veršů J. H. Krchovského jako intertextový znak. *Bohemica Olomucensia* 1, 262–265.
- (31) Plecháč, P. (2009). Metrická teorie Otto Jespersena. *Aluze* 12(2), 99–101.

- (32) Plecháč, P. (2008). Česká versifikace a generativní metrika. *Aluze* 11(3), 2008, 86–93.
- (33) Plecháč, P. (2008). Na rozhraní estetických norem: Poezie Bohdana Jelínka. In V. Polách (ed.), *Jazyková interakce a jazykové rozhraní a teorie „cutting-edge“*. Olomouc: VUP, 219–221.

Slovníková hesla

- (34) Kolár, R. – Plecháč, P. (2018). Rytmičtý impuls; Strukturální versologie; Verš; Zvukosled; Zvuková stránka díla. In O. Sládek et al., *Slovník literárněvědného strukturalismu*. Praha & Brno: ÚČL & Host.
- (35) Kolár, R. – Plecháč, P. (2017). Čeština a versifikační systémy. In P. Karlík, M. Nekula, J. Pleskalová (eds.), *Nový encyklopedický slovník češtiny*. Praha: NLN 2017.

Software

- (36) Plecháč, P. (2017). Sticho. Stylometric analysis of poetic texts based on versification, dostupné z <https://github.com/versotym/stichometry>.
- (37) Plecháč, P. (2017). RhymeTagger. Collocation-driven algorithm for discovering rhymes in a corpus of poetic texts, dostupné z <https://github.com/versotym/rhymeTagger>.
- (38) Plecháč, P. – Kolár, R. (2017). Gunstick 2.0. Databáze českých rýmů. Ústav pro českou literaturu AV ČR, dostupné z <http://www.versologie.cz>.
- (39) Plecháč, P. (2017). Hex 2.0. Klíčová slova v české poezii. Ústav pro českou literaturu AV ČR, dostupné z <http://www.versologie.cz>.
- (40) Plecháč, P. (2017). Eufonometr 2.0. Fonetická transkripce a analýza eufonie textu. Ústav pro českou literaturu AV ČR, dostupné z <http://www.versologie.cz>.
- (41) Plecháč, P. – Změlík, R. (2015). KORBEL – Manažer beletristických korpusů; Autorský slovník Jana Čepa, dostupné z <http://slovníkcep.g6.cz>.
- (42) Plecháč, P. (2015). Babel – SQL dotazy do Korpusu českého verše. Ústav pro českou literaturu AV ČR, dostupné z <http://www.versologie.cz>.
- (43) Plecháč, P. – Ibrahim, R. (2013). Databáze českých meter. Ústav pro českou literaturu AV ČR, dostupné z <http://www.versologie.cz>.

Recenze

- (44) Plecháč, P. (2018). Jak spočítat žánry. *Česká literatura* 66(1), 122–124. [rec. M. Kubát (2016). *Kvantitativní analýza žánrů*. Ostrava: Ostravská univerzita.]
- (45) Martínez-Cantón, C. – Plecháč, P. – Ruiz Fabo, P. – Seláf, L. (2017). Plotting poetry: On mechanically enhanced reading, 5th–7th October, Basel, Switzerland. *Studia metrica et poetica* 4(2), 126–137. [zpráva z konference]
- (46) Plecháč, P. (2015). Statistické metody v literární vědě. *Česká literatura* 63(6), 974–975. [rec. D. L. Hoover – J. Culpeper – K. O'Halloran (2014). *Digital Literary Studies: Corpus Approaches to Poetry, Prose, and Drama*. New York / London: Routledge.]

- (47) Plecháč, P. (2015). O poezii v číslech. *Česká literatura* 63(3), 452–455. [rec. R. Čech – I.-I. Popescu – G. Altmann (2014). *Metody kvantitativní analýzy (nejen) básnických textů*. Olomouc: Univerzita Palackého.]
- (48) Plecháč, P. (2012). *Metrika (nejen) srovnávací*. *Česká literatura* 60(3), 445–448. [rec. M. Lotman, M. K. Lotman (2011 eds.). *Frontiers in Comparative Prosody*. Bern et al.: P. Lang.]
- (49) Plecháč, P. (2010). Pavel Jiráček potřetí. *Aluze* 13(2), 80–82. [rec. P. Jiráček (2009). *Kognitivní interpretce českého verše v rytmických souvislostech metrické teorie* Miroslava Červenky. Brno: Host.]
- (50) Plecháč, P. (2010). Z odkazu Miroslava Červenky. *Aluze* 13(1), 75–76. [rec. M. Červenka (2009). *Záznamník*. Brno: Atlantis.]
- (51) Plecháč, P. (2009). *Kognitivní poetika?* *Aluze* 12(3), 84–86. [rec. P. Jiráček (2008). *Význam a subjektivita v lyrice; Kognitivní struktury v lyrické představě*. Brno: Host.]
- (52) Plecháč, P. (2008). O spojení zvuku a smyslu ve verši. *Aluze* 11(2), 102–103. [rec. P. Jiráček (2007). *Lyrický rytmus; O spojení zvuku a smyslu ve verši*. Brno: Host.]

Překlady

- (53) Jespersen, O. (2009). *Metrum*. *Aluze* 12(2), 102–120.
- (54) Halle, M. – Keyser, S. J. (2008). *Teorie metra*. *Aluze* 11(3), 58–85.

Popularizační publikace

- (55) Ibrahim, R. – Plecháč, P. (2015). *Vrchlický 2.0*. *Biblio – příloha Lidových novin a Literárních novin*. 12. 2. 2015, 20–21.
- (56) Ibrahim, R. – Plecháč, P. (2014). *Báseň a počítač*. Praha: Academia.

Konference, workshopy, zvané přednášky

- (57) Plecháč, P. – Birnbaum, D. J. (2019). Assessing the reliability of stress as a feature of authorship attribution in syllabic accentual syllabic verse. *Quantitative Approaches to Versification*. Praha 24.–26. 6. 2019.
- (58) Plecháč, P. (2019). *Korpus českého verše*. Univerzita Konštantína Filozofa v Nitre, 15. 3. **[zvaná přednáška]**
- (59) Plecháč, P. (2019). *Metody kvantitativní versologie*. Workshop. Univerzita Konštantína Filozofa v Nitre, 15. 3.
- (60) Plecháč, P. (2018). *Authorship of Henry VIII*. Association for Slavic, East European & Eurasian Studies convention. Boston, 6.–9. 12.
- (61) Plecháč, P. (2018). *Versification and authorship recognition*. *Data, Humanities & Language: Tools & Applications*. Tartu, 26.–28. 9. 2018. **[keynote speaker]**
- (62) Plecháč, P. (2018). *Versification and authorship recognition (workshop)*. *Data, Humanities & Language: Tools & Applications*. Tartu, 26.–28. 9. 2018.

- (63) Plecháč, P. – Bobenhausen, K. – Hammerich, B. (2018). Automatic author recognition. Plotting Poetry II: Bringing Deep Learning to Computational Poetry Analysis. Berlin, 12.–14. 9. 2018. **[zvaná přednáška]**
- (64) Plecháč, P. (2018). Versification and authorship recognition. Central European Literary Theory (Structuralism, Marxism, Phenomenology), Praha, 10.–12. 7. 2018.
- (65) Plecháč, P. (2018). Stylometry and Versification. DH Lunch. Instytut Języka Polskiego PAN, Kraków, 9. 2. 2018. **[zvaná přednáška]**
- (66) Plecháč, P. (2017). Authorship attribution based on frequency characteristics of verse. Association for Slavic, East European & Eurasian Studies convention. Chicago, 9.–12. 11.
- (67) Plecháč, P. (2017). Slavic digital poetry roundtable (chair). Association for Slavic, East European & Eurasian Studies convention. Chicago, 9.–12. 11.
- (68) Plecháč, P. (2017). Collocation-driven method of discovering rhymes in a corpus of Czech, English, and French poetic texts. *Machiner la poésie – Plotting Poetry*. Basel, 5.–7. 10. 2017.
- (69) Plecháč, P. – Bobenhausen, K. – Hammerich, B. (2017). Versification and authorship attribution. *Frontiers in Comparative Metrics III*. Tallinn, 29.–30. 9. 2017.
- (70) Škrabal, M. – Piorecký, K. – Plecháč, P. – Kolár, R. – Kuběnský, P. – Procházka, P. (2017). A corpus of contemporary Czech poetry. Poster. CL2017 – The 9th International Corpus Linguistics Conference. Birmingham, 24.–28. 7.
- (71) Plecháč, P. (2017). Stylometrical analysis of versified texts. *Text Analysis Across Disciplines*. Budapest, 2. 5. **[zvaná přednáška]**
- (72) Plecháč, P. (2016). Corpus of Czech Verse and Beyond. Association for Slavic, East European & Eurasian Studies convention. Washington DC, 17.–20. 11.
- (73) Kolár, R. – Plecháč, P. – Říha, J. (2016). Functional structural study of verse. *Expérience et avenir du structuralisme – Vergangenheit un Zukunft des Strukturalismus – Past and Prospects of Structuralism – Minulost a budoucnost strukturalismu*. Praha, 24.–26. 10.
- (74) Plecháč, P. (2016). Measuring euphony. Přednáška v Max-Planck-Institut für Empirische Ästhetik, Frankfurt am Main, 25. 5. **[zvaná přednáška]**
- (75) Plecháč, P. (2015). 19th Century Czech Verse. Association for Slavic, East European & Eurasian Studies convention. Philadelphia, 19.–22. 11.
- (76) Plecháč, P. – Kolár, R. (2015). Korpus českého verše. Poster. WDH 2015 – Seminář o digitálních zdrojích a službách ve společenských a humanitních vědách. Praha, 22. 9.
- (77) Kolár, R. – Plecháč, P. (2015). Hus v české poezii 19. století. V. kongres světové literárněvědné bohemistiky. Praha, 1. 7.
- (78) Plecháč, P. (2015). Authorship Attribution Based on Frequency Characteristics of Verse. *Computational Approaches to Poetry*. Helsinki, 5. 6.
- (79) deCastro-Arrazola, V. – Plecháč, P. (2015). XML Schema for Encoding Verse Corpora. *The Structure of Verse: Formal, Experimental and Computational Approaches*. Leiden, 19.–20. 3.

- (80) Plecháč, P. – Ibrahim, R. (2014). How to Compile and Use a Verse Corpus. Association for Slavic, East European & Eurasian Studies convention. San Antonio (Texas), 20.–23. 11.
- (81) Ibrahim, R. – Plecháč, P. (2014). Korpus českého verše a jeho nástroje. Literatura, technologie, média. Praha, 8. 10.
- (82) Plecháč, P. (2014) Korpus českého verše a možnosti jeho využití. Korpusová lingvistika 2014. Praha, 17.–19. 9.
- (83) Plecháč, P. – Ibrahim, R. (2014). Corpus of Czech Verse. Slavjanskij stich XI. Moskva, 20.–24. 6.
- (84) Ibrahim, R. – Plecháč, P. – Říha, J. (2014). O českém „Vvedeniji v teoriju sticha“. Slavjanskij stich XI. Moskva, 20.–24. 6.
- (85) Plecháč, P. – Ibrahim, R. (2014). Database of Czech Verse. Frontiers in Comparative Metrics 2: in memoriam Lucylla Pszczołowska. Tallinn, 19.–20. 4.
- (86) Plecháč, P. – Ibrahim, R. (2013). Ot metra k ritmu i obratno. 100 let ruskogo formalizma. Moskva, 25.–29. 8.
- (87) Plecháč, P. – Říha, J. (2013). Zvukovaja garmonija: točnyje metody. Gasparovskije čtenija. Moskva, 18.–20. 4.
- (88) Ibrahim, R. – Plecháč, P. (2013). Častotnosť častěj reči v češskoj poezii. Gasparovskije čtenija. Moskva, 18.–20. 4.
- (89) Plecháč, P. – Říha, J. (2013). Measuring the Euphony. 2. Brikovskije čtenija: Metodologija i praktika ruskogo formalizma. Moskva, 20.–24. 3.
- (90) Plecháč, P. – Říha, J. (2013). Kritika textu a automatická analýza verše. Od knihy ke knize. Praha, 14. 2.
- (91) Plecháč, P. (2013). Avtomatičeskij analiz češskogo sticha. Stichovedenije, lingvistika, tĕkstologija; nasledije M. Červenki i aktualnyje zadači nauki o stiche. Moskva, 8.–9. 2.
- (92) Plecháč, P. – Říha, J. (2013). Verse Theory and Textual Scholarship. Stichovedenije, lingvistika, tekstologija; nasledije M. Červenki i aktualnyje zadači nauki o stiche. Moskva, 8.–9. 2.
- (93) Plecháč, P. – Ibrahim, R. (2012). Frequency of Parts of Speech in Czech Poetry of 19th Century. Jakobson ve struktuře (české) vědy. Olomouc, 10.–11. 12.
- (94) Ibrahim, R. – Plecháč, P. (2012). Automatyczna analiza wiersza: punkt wyjścia. Wiersz: Rytm – Dźwięk – Obraz – Semantyka. Warszawa, 10.–11. 11.
- (95) Plecháč, P. – Ibrahim, R. (2012). Verš podivuhodného kouzelníka. Podivuhodný kouzelník Vítězslav Nezval. Brno, 20.–22. 6.
- (96) Ibrahim, R. – Plecháč, P. (2011). Jak poznat metrum? Jan Mukařovský dnes: Tradice a perspektiva. Praha, 14.–15. 11.
- (97) Ibrahim, R. – Plecháč, P. (2010). Teorie verše ve stopovém množství. Poetický Cikháj v Brně. Brno, 26.–27. 3.
- (98) Plecháč, P. (2009). K jednomu aspektu Březinova rýmu. NEpravidelnost a NEjednoznačnost v non/verbální komunikaci; 10. mezinárodní konference Setkání mladých lingvistů. Olomouc, 11.–13. 5.

- (99) Plecháč, P. (2008). Přízvuková kontura veršů J. H. Krchovského jako intertextový znak. *Interkulturní komunikace a lingvistika*; 9. mezinárodní konference Setkání mladých lingvistů. Olomouc, 12.–14. 5.
- (100) Plecháč, P. (2007). Na rozhraní estetických norem: Poezie Bohdana Jelínka. *Jazyková interakce a jazykové rozhraní a strategie „cutting-edge“*; 8. mezinárodní konference Setkání mladých lingvistů. Olomouc, 14.–16. 5.
- (101) Plecháč, P. (2006). Prostorové dimenze fikčního světa Větrů od Pólů. *Svět za slovy a jejich tvary, svět za spojením slov*; 7. mezinárodní konference Setkání mladých lingvistů. Olomouc, 15.–17. 5.

Ohlasy

Ocenění

- (102) Prémie Otto Wichterleho 2016. Akademie věd ČR.
- (103) The Visegrad Group Academies Young Researcher Award 2015 in Languages and Literature. Polska Akademia Nauk – Akademie věd ČR – Magyar Tudományos Akadémia – Slovenská akadémia vied.
- (104) SCIAP 2014 – soutěžní přehlídka popularizace vědy. Středisko společných činností Akademie věd ČR. (Korpus českého verše – 3. místo v kategorii Internet)

Citace

- (105) Barash, O. – Vekshin, G. (2019). Some characteristics of sound patterns in English. In P. Plecháč, B. P. Scherr, T. Skulacheva, H. Bermúdez-Sabel, R. Kolár (eds.), *Quantitative Approaches to Versification*. Prague: ICL CAS, 11–22. **Cit. (23)**.
- (106) Martínez-Cantón, C. – Ruiz-Fabo, P. – González-Blanco, E. (2018). ANJA, ¿dónde están los encabalgamientos? In *Digital Humanities 2018: Book of Abstracts*. Ciudad de México: RHD, 632–633. **Cit. (19)**.
- (107) Mitmann, A. – von Wangenheim, A. – Dos Santos, A. L. (2018). Aoidos: A system for the automatic scansion of poetry written in Portuguese. In A. Gelbukh (ed.) *Computational Linguistics and Intelligent Text Processing*. Cham: Springer, 611–628. **Cit. (29)**.
- (108) Navrátil, M. (2018). Aktuálne otázky slovenskej textológie: Textologický diskurz na Slovensku. *Slovenská literatúra* 65(4), 294–301. **Cit. (13)**.
- (109) Změlík, R. (2018). Kvantitativní a korpusový výzkum v literární vědě: možnosti a přístupy. *Slovo a slovesnost* 79(1), 47–65. **Cit. (20)**.
- (110) Curado Malta, M. – Centenera, P. – González-Blanco, E. (2017). Using reverse engineering to define a domain model: the case of the development of a metadata application profile for European poetry. In M. Curado Malta, A. A. Baptista, P. Walk (eds.), *Developing Metadata Application Profile*. Hershey (Pennsylvania): IGI Global, 146–180. **Cit. (43)**.

- (111) Curado Malta, M. – Gonzalez-Blanco, E. – Martínez Cantón, C. – del Rio, G. (2017). A common conceptual model for the study of poetry in the digital humanities. In R. Lewis et al. (eds.) *Digital Humanities 2017: conference abstracts*. McGill University & Université de Montréal, 210–212. **Cit. (43)**.
- (112) Tureček, D. (2017). Stereotypy ideowe w obrazowaniu Słowacji w Czeskiej poezji lat 1860–1939. *Porównania* 21(2), 151–162. **Cit. (3)**.
- (113) Říha, J. (2017). Od děrných štítků ke korpusu českého verše, *Česká literatura*, 108–117. **Cit. (19, 20, 22, 56)**.
- (114) Tureček, D. (2017). Karel Vinařický: Mezi klasicismem a romantikou. *Bohemica litteraria* 20(1), 54–92. **Cit. (3)**.
- (115) Říha, J. (2017). Metrika a strofika Nerudových Balad a romancí. In V. Faktorová, J. Pácalová – Z. Urválková (eds.), *Víno, ženy, zpěvy: V(d)ěčné téma literárních dějin*. Česká Budějovice: JUČB, 165–175. **Cit. (2)**.
- (116) von Wangenheim, A. – Mitmann, A. – Dos Santos, A. L. (2016). A multi-level visualization scheme for poetry. *Proceedings 20th International Conference Information Visualisation IV 2016*. Lisboa: CPS, 312–317. **Cit. (29)**.
- (117) Mastalski, A. S. (2016). Sakadometryczna analiza wiersza: Ogólne założenia teoretyczne do badań empirycznych. *Neurolingwistyka Praktyczna* 2, 108–125. **Cit. (3)**.
- (118) Bičan, A. (2016). Phonological properties of Czech verse and prose Compared. In E. G. Rubio, E. Kislova, E. Kubicka (eds.), *Beiträge zum 19. Arbeitstreffen der Europäischen Slavistischen Linguistik (Polyslav)*. Wiesbaden: Harrassowitz Verlag, 27–37. **Cit. (23)**.
- (119) Darebný, J. – Vázquez Tourino, D. (2016). *E-manual de Métrica española*. Brno: FF MU. **Cit. (3)**.
- (120) Změlík, R. (2016). K základním možnostem modelování literárního díla v kontextu české literární vědy. *Tradice a současnost. Studia Słowianoznawcze* 12, 407–431. **Cit. (20, 56)**.
- (121) Změlík, R. (2016). Od textu ke korpusu, od korpusu k textu. In O. Fedoszov, A. Janiec-Nyitrai, Z. Muchová (eds.), *Od počátku k současnosti (1955–2015). 60 let bohemistiky na ELTE*. Budapest: ELTE BTK, 401–409. **Cit. (56)**.
- (122) Čermochová, K. (2016). Úvaha o rytmu v řeči, poezii i tak vůbec. *Slovo a smysl* 26, 78–88. **Cit. (3)**.
- (123) Pilščikov, I. – Starostin, A. (2015). Reconnaissance automatique des mètres des vers russes: une approche statistique sur corpus. *Langages* 199, 89–106. **Cit. (29)**.
- (124) Bičan, A. (2015). Distribution of Vocalic Quantity in Czech, *Grazer Linguistische Studien* 83, 2015, 133–138. **Cit. (3)**.
- (125) Bičan, A. (2015). Corpus-based Analysis of the Czech Syllable. In E. Guetiérrez Rubio, *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)* 18. München – Berlin – Washington: Harrasowitz Verlag, 26–36. **Cit. (3)**.
- (126) Bičan, A. (2015). Fonologický lexikální korpus češtiny a jeho analýza, *Bohemica Olomucensia* 7/3–4, 45–59. **Cit. (3)**.
- (127) Darebný, J. (2015). El verso de las traducciones de la vida es sueño al checo: el caso de la octava. *Romanica Olomucensia* 27(1), 113–124. **Cit. (3)**.

- (128) Změlík, R. (2015). Kvantitativně-korpusová analýza a literární věda. Olomouc: Univerzita Palackého. **Cit. (41, 56).**
- (129) Machala, L. et al. (2015). Panorama české literatury po roce 1989 (2). Praha: Knižní klub. **Cit. (3).**
- (130) Rainsford, T. M. – Scriver, O. (2014). Metrical Annotation for a Verse Treebank. Proceedings of the Thirteenth International Workshop on Treebanks and Linguistic Theories (TLT13). Tübingen, Universität Tübingen, 149–159. **Cit. (29).**
- (131) Pořízka, P. (2014). Tvorba korpusů a vytěžování jazykových dat: metody, modely, nástroje. Olomouc: Univerzita Palackého. **Cit. (42, 43).**
- (132) Štochl, J. (2014). Norma českého jambu. Česká literatura 62(5), 707–727. **Cit. (28, 32).**
- (133) Malura, J. – Smolka, Z. (2014). Úvod, který předkládá problémy. Česká literatura 62(5), 824–826. **Cit. (3).**
- (134) Štochl, J. (2011). Svět díla Jiřího Ortena. Praha: Torst. **Cit. (32).**
- (135) Sgallová, K. (2011). O versologii diskutujme věcně. Aluze 14/1, 76–77. **Cit. (49).**
- (136) Jiráček, P. (2010). Odpověď Petru Plecháčovi a trošku i Robertu Ibrahimovi. Aluze 13/3, 89–92. **Cit. (49).**

Recenze

- (137) Navrátil, M. (2018). Editologie (Od náčrtu ke knize), Litikon 3(2), 396–400. **Rec. (13).**
- (138) Místecký, M. (2018). Červenková teorie verše prizmatem korpusového bádání, Naše řeč 101(3), 182–184. **Rec. (2).**
- (139) Milička, J. (2018). Kapitoly z korpusové versologie – cesta správným směrem. Česká literatura 66(2), 287–289. **Rec. (2).**
- (140) Škrabal, M. (2018). Poetika statistického testování hypotéz. Svět literatury 57, 175–178. **Rec. (2).**
- (141) Dudek, P. (2018). Statistické testování hypotéz českého verše. Bohemica Olomucensia 10(2), 192–196. **Rec. (2).**
- (142) Gbúr, J. (2017). Kapitoly z korpusové versologie. Litikon 2(2), 309–312. **Rec. (2).**
- (143) -mort- (2017). Kapitoly z korpusové versologie. Host 13/2017, 28. **Rec. (2).**
- (144) Wiendl, J. (2015). Živá versologie. Česká literatura 63(2), 287–292. **Rec. (3).**
- (145) Matějka, I. (2015). Věda v sešitovém vydání. Literární noviny 15. 3. **Rec. (56).**
- (146) -účl- (2015) Tip do vaší knihovny. Čtenář – měsíčník pro knihovny 67/(7–8), dostupné z <<http://ctenar.svkkk.cz/clanky/2015-roc-67/0708-2015/tip-do-vasi-knihovny-146-2005.htm>> **Rec. (56).**
- (147) Darebný, J. (2014). Úvod do teorie verše. Bohemica litteraria 17(2), 151–153. **Rec. (3).**
- (148) Tureček, D. (2014). Potřeba versologie. Svět literatury 23/1, 278–280. **Rec. (3).**
- (149) Sadowski, W. (2013). Versologie mezi jazyky. Česká literatura 61(6), 951–956. **Rec. (29).**
- (150) Wachtel, M. (2013). Formal Methods in Poetics. Glottometrics 26, 104–109. **Rec. (29).**

Citace v kvalifikačních pracích

- (151) Místecký, M. (2018). Komplexní analýza aktivity ve vybraných dílech české poezie 20. století. Magisterská diplomová práce. Ostrava: FF OU. Cit. (2).
- (152) Jahelková, L. (2018). Roald Dahl's the Giraffe and the Pelly and Me: translation and analysis. Magisterská diplomová práce. Brno: PF MU. Cit. (3).
- (153) Aguirrezabal Zabaletak, M. (2017). Automatic Scansion of Poetry. Dizertační práce. Bilbao: Euskal Herriko Unibertsitatea. Cit. (19).
- (154) Darebný, J. (2017). Verso español y verso checo: traducción del teatro polimétrico de Calderón de la Barca. Dizertační práce. Brno: FF MU. Cit. (3).
- (155) Hron, J. (2017). Současné české překlady anglicky psané poezie pro děti. Magisterská diplomová práce. Praha: FF UK. Cit. (3).
- (156) Kapounová, K. (2017). Forma a obsah v „nové čínské básni“. Magisterská diplomová práce. Praha: FF UK. Cit. (3).
- (157) Hovorková, B. (2017). Seifertova básnická tvorba od roku 1965. Bakalářská diplomová práce. Ústí nad Labem: Katedra bohemistiky PF UJEP. Cit. (3).
- (158) Lafatová, A. (2017). Postava dudáka v české literatuře 19. století, srovnání s inscenací Strakonický dudák J. A. Pitínského z roku 2013. Bakalářská diplomová práce. Praha: PF UK. Cit. (43).
- (159) Mittmann, A. (2016). Escansão automática de versos em Português. Dizertační práce. Florianópolis: Centro Tecnológico, Universidade Federal de Santa Catarina, Brasil. Cit. (29, 16, 19).
- (160) Chodaková, P. (2016). Analyse prosodique de musiques urbaines en français et en tchèque. Dizertační práce. Praha/Paris: Ústav románských studií FF UK / Université Paris Diderot. Cit. (3, 29, 19, 22, 32).
- (161) Kuzminova, A. G. (2016). Razpredelenije častěj reči v nepolnoudarnych strokach češskogo četyrechstopnogo jamba na materiale proizvedenij Jaroslava Vrchlickogo. Magisterská diplomová práce. Moskva: Institut lingvistiki, Rossijskij gosudarstvennyj gumanytarnyj universitet. Cit. (24, 19).
- (162) Místecký, M. (2016). Souhláskové shluky jako eufonický prostředek. Bakalářská diplomová práce. Ostrava: Katedra českého jazyka FF OU. Cit. (28).
- (163) Mrázek, O. (2016). Arne X: výuková počítačová hra. Bakalářská diplomová práce. Ústav českého jazyka FF MU. Cit. (42).
- (164) Vacková, A. (2016). Vincenc Furch. Bakalářská diplomová práce. České Budějovice: Ústav bohemistiky FF JČU. Cit. (3).
- (165) Valentová, L. (2016). Básníci Charles Cros, Tristan Corbière a Jules Laforgue v pojetí Vladimíra Holana. Bakalářská diplomová práce. Praha: FF UK. Cit. (3).
- (166) Židek, T. (2016). Básnické dílo Stanislava Zedníčka. Bakalářská diplomová práce. Ústí nad Labem: PdF UJEP. Cit. (3).
- (167) Mastalski, A. S. (2015). Kognitywne podstawy teorii wiersza, semantyki wersyfikacyjnej oraz interpretacji wersologicznej. Wybrane problemy okulograficznej analizy kompozycji prozodyjnej. Dizertační práce. Kraków: Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie Wydział Filologiczny. Cit. (3, 26, 56)

- (168) Říha, J. (2015). Rým a strofika v českém verši, obzvláště u Jana Nerudy. Dizertační práce. Praha: Ústav české literatury a komparatistiky FF UK. Cit. (3, 23, 29, 56).
- (169) Hudečková, M. (2015). E. A. Poe – The Raven. Analysis of Sixteen Czech Translations Based on Philosophy of Composition. Bakalářská diplomová práce. Brno: Katedra anglického jazyka a literatury PF MU. Cit. (3).
- (170) Štěpinová, B. (2015). Analyse ausgewählter tschechischer Übersetzungen der Lyrik J. W. von Goethes. Plzeň: Katedra germanistiky a slavistiky FF ZU. Cit. (3).
- (171) Záchová, K. (2015). Korpusy a jejich využití ve výuce českého jazyka na ZŠ a SŠ. Bakalářská diplomová práce. Olomouc: Katedra bohemistiky FF UP. Cit. (3, 43).
- (172) Lošťák, O. (2015). Aspekty digitalizace textů. Bakalářská diplomová práce. Brno: Ústav hudební vědy FF MU. Cit. (20).
- (173) Zindulková, K. (2014). Tónický verš a čeština. Magisterská diplomová práce. Praha: Ústav české literatury a literární vědy FF UK. Cit. (3, 28).
- (174) Zapletalová, K. (2014). Překlady Jaroslava Vrchlického z románských literatur. Magisterská diplomová práce. České Budějovice: Ústav bohemistiky FF JČU. Cit. (3).
- (175) Hron, J. (2014). Komentovaný překlad vybraných kapitol z knihy *The Life and Death of Anne Boleyn*. Padstow: Blackwell Publishing, 2005. Bakalářská diplomová práce. Praha: Ústav translatologie FF UK. Cit. (3).
- (176) Šenkapoun, P. (2011). Písňový verš Karla Kryla a Jaromíra Nohavici. Dizertační práce. Olomouc: Katedra bohemistiky FF UP. Cit. (28).
- (177) Zindulková, K. (2011). Vzájemné působení metrických a fonologických faktorů při utváření básnického rytmu. Bakalářská diplomová práce. Praha: Ústav české literatury a literární vědy FF UK. Cit. (28).

Ohlasy v médiích

- (178) Matušková, L. (2017). Kapitoly z korpusové versologie. Web AV ČR, dostupné z <<http://www.avcr.cz/cs/pro-media/aktuality/Kapitoly-z-korpusove-versologie/>>.
- (179) Willoughby, T. (2017). Jde spočítat autor básně? MF Dnes – City Life 24. 2., 30–31.
- (180) Matušková, L. (2015). Korpus českého verše nabízí jedinečné možnosti analýzy poezie. Vědaproživot.cz – portál Akademie věd. Dostupné z <http://www.vedaprozivot.cz/sd/novinky/hlavni-stranka/150601_korpus_ceskeho_vedce.html>.
- (181) -účl- (2015) Unikátní projekt počítačnické versologie. Čtenář – měsíčník pro knihovny 67/7–8, dostupné z <<http://ctenar.svkkk.cz/clanky/2015-roc-67/0708-2015/unikatni-projekt-komputacni-versologie-146-2004.htm>>
- (182) Feryna, J. (2015). Nic nesplní Feryna, / jak ho potká dívčina. Splav! 7.
- (183) Svoboda, L. (2014). Korpus českého verše, Akademický bulletin, dostupné z <http://data.abicko.avcr.cz/sd/novinky/hlavni-stranka/news_1171.html>.
- (184) Kočí, P. – Šulek, M. (2014). Srdce, duše, láska, svět – nejfrekventovanější slova v klasické české poezii, Český rozhlas Radiožurnál, dostupné z <http://www.rozhlas.cz/zpravy/data/_zprava/srdce-duse-laska-svet-nejfrekventovanejsi-slova-v-klasicke-ceske-poezii—1416246>.

- (185) Bouchner, J. (2014). Korpus českého verše a jeho využití ve výuce. Metodický portál inspirace a zkušenosti učitelů, dostupné z <<http://spomocnik.rvp.cz/clanek/19305/KORPUS-CESKEHO-VERSE-A-JEHO-VYUZITI-VE-VYUCE.html>>.
- (186) Šmerglová, M. (2014). Počítačem na Nerudu. Týdeník Květy 48, 22–23.
- (187) Olivová, J. (2014). Korpus českého verše, Český rozhlas Vltava, dostupné z <http://www.rozhlas.cz/mozaika/veda/_zprava/korpus-ceskeho-verse—1416705>.

Granty

- (188) Stylometrická analýza básnických textů (2017–2019; GA ČR 17–01723S). [řešitel]
- (189) POSTDATA – Poetry Standardization and Linked Open Data (2016–2021; ERC Starting Grant 2015/679528; <http://postdata.linhd.es/>). [člen výzkumné komunity]
- (190) Dějiny a teorie českého verše 19. století (2011–2015; GA ČR P406/11/1825). [člen řešitelského týmu]

Pedagogická a popularizační činnost

- (191) Bermúdez-Sabel, H. – Martínez-Cantón, C. I. – Plecháč, P. (2019). Seminář korpusové versologie. Univerzita Palackého v Olomouci, 5. 3.
- (192) Kolár, R. – Plecháč, P. – Říha, J. – Dobiáš, D. (2018). Verzologický tým. Rozhovor. Litikon 3/1, 243–245.
- (193) Plecháč, P. (2018). Přednáška „Versologie a určování autorství“. Pražský lingvistický kroužek, 23. 4. 2018.
- (194) Plecháč, P. – Sela, A. (2018). Versification & Stylometry. Text Mining Boot Camp. Central European University, Budapest, 10.4.–20. 4.
- (195) Plecháč, P. (2018). The Corpus of Czech Verse and its online tools. Perspectives on Digital Scholarship for REEE Scholars. Committee on Libraries and Information Resources ASEES – Education & Access Subcommittee, 9. 3. 2018.
- (196) Plecháč, P. (2016). Rozhovor pro Český rozhlas Plus, 26. 7.
- (197) Plecháč, P. (2016). Quantitative research on versification: The Corpus of Czech Verse. Tecnologías digitales aplicadas al estudio de la poesía. Cursos de Verano. Universidad Nacional de Educación a Distancia, Madrid, 27. 6.–1. 7.
- (198) Plecháč, P. (2016). Rozhovor pro Český rozhlas Vltava, 30. 5.
- (199) Plecháč, P. (2016). Přednáška v Brněnském naratologickém kroužku, 7. 4.
- (200) Plecháč, P. (2016). Přednáška v Literárněvědné společnosti Olomouc, 3. 4.
- (201) Plecháč, P. (2016). Seminář pro doktorandy FF UP v Olomouci, 26. 2.
- (202) Plecháč, P. – Kolár, R. (2015). Rozhovor pro Český rozhlas Plus, 12. 12.
- (203) Plecháč, P. (2015). Rozhovor pro portál Host 7 dní online, 18. 11.
- (204) Kolár, R. – Plecháč, P. (2015). Prezentace Korpusu českého verše, knihy Úvod do teorie verše a publikace Báseň a počítač, V. kongres světové literárněvědné bohemistiky. Praha, 1. 7.

- (205) Kolár, R. – Plecháč, P. (2015). Přednáška o Korpusu českého verše. Klub přátel českého jazyka – Univerzita Karlova v Praze, 15. 4.
- (206) Plecháč, P. (2015). Přednáška o Korpusu českého verše. Univerzita Palackého v Olomouci, 15. 4.
- (207) Plecháč, P. (2014). Přednáška o Korpusu českého verše. Bohemistický kruh. Univerzita Karlova v Praze, 11. 12.
- (208) Plecháč, P. (2014). Přednáška o Korpusu českého verše pro studenty Univerzity Karlovy v Praze. Ústav pro českou literaturu AV ČR. Praha, 11. 12.
- (209) Kolár, R. – Plecháč, P. (2014). Přednáška Český verš pod mikroskopem. Univerzita J. E. Purkyně v Ústí nad Labem, 11. 11.
- (210) Plecháč, P. (2014). Rozhovor pro Český rozhlas Plus, 10. 11.
- (211) Plecháč, P. (2014). Rozhovor pro Český rozhlas Vltava, 5. 11.
- (212) Kolár, R. – Plecháč, P. (2014). Den s Korpusem českého verše (tisková konference / přednáška pro veřejnost). Ústav pro českou literaturu AV ČR. Praha, 4. 11.
- (213) Plecháč, P. (2014). Versologický workshop pro doktorandy FF OU v Ostravě, 10. 10.
- (214) Plecháč, P. (2014). Účinkování v pořadu Lovci záhad (epizoda Po stopách K. H. Máchy). Česká televize Děčko, 25. 5.
- (215) Kolár, R. – Plecháč, P. (2014). Přednáška Voda v české literatuře pro studenty Gymnázia prof. Jana Patočky. Ústav pro českou literaturu AV ČR. Praha, 13. 5.
- (216) Plecháč, P. (2012). Přednáška Versifikační systémy. Univerzita Pardubice, 19. 11.
- (217) Plecháč, P. (2012). Přednáška Sémantika daktylu v raném obrození. Pražský lingvistický kroužek. Praha, 14. 5.
- (218) Ibrahim, R. – Plecháč, P. (2011). Přednáška Počítače a poezie. Akademie věd / Památník národního písemnictví (Týden vědy a techniky). Praha, 8. 11.
- (219) Plecháč, P. (2011). Seminář Úvod do teorie verše, Univerzita Palackého v Olomouci, povinný seminář pro studenty České filologie, letní semestr 2011: 1h týdně.
- (220) Plecháč, P. (2010). Seminář Úvod do teorie verše. Univerzita Palackého v Olomouci, povinný seminář pro studenty České filologie, letní semestr 2010: 2h týdně.