

Filozofická fakulta Univerzity Palackého v Olomouci

Katedra obecné lingvistiky



Určování autorství Slezských písní

Bakalářská práce

Autor: Barbora Varmužová

Vedoucí práce: Mgr. Dan Faltýnek, Ph.D.

Olomouc 2021

Prohlášení

Prohlašuji, že jsem bakalářskou/magisterskou diplomovou práci „Určování autorství Slezských písní“ vypracovala samostatně a uvedla jsem veškerou použitou literaturu a veškeré použité zdroje.

V Olomouci dne 01.03.2021 Podpis

Abstrakt

Název práce: Určování autorství Slezských písní

Autor práce: Barbora Varmužová

Vedoucí práce: Mgr. Dan Faltýnek, Ph.D.

Počet stran a znaků:

Počet příloh: 2

Abstrakt (minimálně 900 znaků): Cílem této práce je využití nízko frekventovaného lexika v analýze autorství Slezských písní. První část práce uvádí čtenáře do kontextu dlouhodobých diskusí na toto téma a je věnovaná dvěma autorům, kteří se v tomto směru docházejí k opačným závěrům: Oldřich Králík zastává teorii, že básně Slezských písní napsal jeden autor a hovořit můžeme o dvou pseudonymech, které reprezentují období v jeho životě. Druhým autorem, na nějž navazujeme, je Jan Drozd, který zpopularizoval teorii o dvojím autorství Slezských písní a zastává názor, že nad sbírkou nadepsaný autor opsal část básní od svého přítele Ondřeje Boleslava Petra. Druhá část práce je věnována analýzám nízko frekventovaného lexika prováděným za pomoci lingvistického softwaru QUITA. Básně jsou na základě hypotéz z první části rozděleny na skupiny. Ty pak mezi sebou porovnáváme za použití několika metod určování autorství, a to v kontrastu k dalším sbírkám dobových autorů. Praktická část končí vyhodnocením výsledků a srovnáním obou teorií. Na základě analýzy nakonec usuzujeme, že nic nenasvědčuje tomu, že by srovnávané skupiny básní Slezských písní pocházely z pera různých autorů.

Klíčová slova: Petr Bezruč, Slezské písně, forenzní lingvistika, určování autorství, QUITA, funkční slova, hapax legomenon

Abstract

Title: Authorship identification of Slezské písně

Author: Barbora Varmužová

Supervisor: Mgr. Dan Faltýnek, Ph.D.

Number of pages and characters:

Number of appendices: 2

Abstract (900 characters): The aim of this thesis is to analyse poems of Slezské písně in a way that they have not been analysed before (by comparison of low-frequency lexicon). The first part of the thesis is focused on two different authors who have come up with conflicting opinions concerning the duality of poetic styles in Slezské písně by Petr Bezruč. While Oldřich Králík acknowledges that Vašek's writing style deteriorated throughout the years, he firmly believes that Vašek is the sole author and the individual differences between each style are attributed to his fifty year old career, Jan Drozd on the other hand popularized a controversial theory that Vladimír Vašek cannot be the mastermind behind the entire collection of poems, as individual poems are too different to be written by a single person. Drozd also claims that he had found enough evidence to support a theory that at least 20 poems have been written by Ondřej Boleslav Petr, a close friend of Vašek. The second part of this thesis is devoted to quantitative analysis done using linguistic software QUITA. Poems within the collection were put into different groups based on Drozd's theory and by using multiple methods of authorship attribution they are compared based on their similarities in low-frequency lexicon and function words, alongside with several collections of poems written by different Czech authors from a similar time period. The conclusion consists of reevaluation of theories explained in part one according to new findings from each analysis.

Keywords: Petr Bezruč, Slezské písně, forensic linguistics, authorship attribution, QUITA, function words, hapax legomenon

Poděkování

Chtěla bych poděkovat svému vedoucímu bakalářské práce doc. Mgr. Danu Faltýnkovi, Ph.D. za pomoc s výběrem tématu práce, poskytnutí většiny materiálů a za nekonečnou trpělivost při konzultacích, bez kterých by tato práce nemohla vzniknout. Taktéž chci poděkovat Mgr. Michalu Kosákovi, Ph.D. za rozdělení básní do skupin a Mgr. Vladimíru Matlachovi, Ph.D., autorovi programu QUITA, pomocí kterého byly texty analyzovány

Obsah

Úvod	7
Teoretická část	9
Kanonizace Slezských písní	10
Teorie trojjedinosti Vladimíra Vaška	11
Teorie dvojího autorství	12
Ondřej Boleslav Petr	12
Jan Drozd	13
Jaromír Nohavica	15
Kritika teorie dvojího autorství	16
Praktická část	17
Vzorek	19
Analýza kompletní slovní zásoby	21
Funkční slova	24
Analýza první sady funkčních slov	25
Analýza druhé sady funkčních slov	28
Hapax legomena	31
Analýza autorství Slezských písní za využití hapax legomen	32
Srovnání grafů – vyhodnocení	34
Analýza vzorků Slezských písní po sloučení vzorků VV a VV mimo jádro	35
Analýza hapax legomen po sjednocení vzorků VV a VV mimo jádro	37
Srovnání výsledků analýz po sjednocení skupiny VV a VV mimo jádro	38
Závěr	39
Literatura a zdroje	42
Seznam obrázků	45
Seznam příloh	45
Seznam tabulek	46

Úvod

Slezské písně jsou unikátní sbírkou v kontextu historie české literatury. Když roku 1899 přišla Janu Herbenovi do redakce časopisu Čas první zásilka s básněmi podepsanými jménem Petr Bezruč, nikdo netušil, jak složité bude toto dílo vydat. V průběhu let Herben dostal více než pětatřicet zásilek s básněmi různých témat. Identita autora těchto básní vyšla na povrch až o několik let později, kdy se ukázalo, že písně pocházely z pera Vladimíra Vaška, plachého úředníka z brněnské pošty, který se k básním nikdy nehlásil a neměl zájem o to je vydávat v knižní podobě. Krom Petra Bezruče se Vladimír Vašek skrýval ještě za dalšími pseudonymy, z těch hlavních to byl Smil z Rolničky a Bohumil Šarek. Básně vydané v Herbenově čase se ovšem těšily ohromné popularitě a byly vydány poprvé v roce 1909, a pak ještě mnohokrát později v různých edicích a úpravách. Mnoho čtenářů si v souvislosti s vydáním básní Petra Bezruče nemohlo nevšimnout, že už v roce 1902 došlo u autora k zásadním změnám v autorském stylu.

Poté, co Petr Bezruč poslal svou první sbírku básní Herbenovi pod pseudonymem Smil z Rolničky, se zdálo, že básně mají s těmi předchozími pramálo společného. Autor už nebyl průvodce děje, nýbrž vzdálený pozorovatel. Už nešlo o vlastenecké a buřičské básně. Nově psal o svém životě, o domově, o svých milostných zkušenostech. Ohromný rozdíl nastal i ve stylu a plynulosti textu. Tento nečekáný rozdíl v kvalitě byl natolik patrný, že ačkoliv básně vyšly v jedné sbírce, dodnes se rozlišuje „jádro Slezských písní“, tvořící básně poslané v prvních 24 zásilkách, a básně ostatní, stylově i tematicky naprosto rozdílné. Vašek byl navíc nechvalně známý tím, že svá díla k údivu čtenářů neustále upravoval. Z těchto důvodů se vedly mnohé diskuse o tom, jak by se Slezské písně měly vydávat, v jakém pořadí jednotlivých básní a v které textové podobě. Finální rozhodnutí proběhlo roku 1963 na Literárněvědní konferenci věnované Slezským písním v Opavě, kde bylo rozhodnuto, že se jako kanonický text bude považovat vydání z roku 1930 (Vodička 1965). Tím ale diskuse o tomto nezvyklém díle neskončily. Více než sto let od jeho prvního zveřejnění v časopise Čas básně inspirují čtenáře a nutí vědce k mnoha otázkám. Jak je možné, že se Vašek během pouhého roku tak změnil? Proč měl k dílu, které z něj udělalo legendu, za období komunismu dokonce národního umělce, tak odtažitý vztah a neustále se skrýval za pseudonymy? Proč od básní odvracel tvář a ze začátku odmítal, aby byly vydány v samostatné knize? A proč je, dle názorů fanoušků i literárních kritiků, naprosto nepochopitelně měnil, a tak jim ubíral na kvalitě?

Teoriím vysvětlujícím rozdíl mezi básněmi z jádra a básněmi novějšími se budeme věnovat v první části této práce. Budou představeny hlavní argumenty pro teorii, že je Vašek jediným autorem, také i argumenty tvrdící opak. Na základě těchto teorií budou v druhé části básně rozděleny na skupiny a budou formulovány hypotézy o jejich možné příbuznosti. Tyto skupiny dále budou porovnávány kvantitativními metodami za pomoci lingvistického softwaru QUITA. Hlavními prvky, na které se při analýze zaměříme, budou funkční slova a slova jedinečná – hapax legomena, porovnávány budou také na základě celého použitého slovníku. Pro přesnost výsledků budou básně porovnávány nejen mezi sebou, ale i se sbírkami jiných autorů obdobné délky. Na základě těchto vlastností budou podobnosti skupin básní zobrazeny v grafech. Tyto výsledky následně vyhodnotíme a interpretujeme vzhledem k otázkám vyloženým v teoretické části. Cílem této práce je studium Slezských písní způsobem, jakým dosud zkoumány nebyly – a to na základě využití atribučních vlastností nízko frekventovaných částí lexika. Práce si neklade za cíl dát definitivní důkaz potvrzující či vyvracející zavedenou teorii, a to že je Vladimír Vašek jediným autorem Slezských písní, má pouze ambici se k této teorii vyjádřit na základě nových skutečností.

Teoretická část

Teoretická část bude zaměřena na publikace významných literárních kritiků a textologů, kteří se problematice Slezských písní věnovali. Prvním z nich bude profesor Oldřich Králík, významný literární historik a textolog, který se dílem zabýval většinu svého života a je autorem mnoha stěžejních publikací na toto téma. Ve svých pracích se soustředil mimo jiné na vývoj osobnosti Vladimíra Vaška a jeho pseudonymů. Poté bude věnována pozornost Janu Drozdovi a jeho dílu Autoři Slezských písní, ve kterém v roce 2003 zpopularizoval teorii, že Vašek nemůže být jediným autorem celé sbírky. Okrajově budou zmíněni i další autoři, kteří taktéž k tomuto tématu přispěli.

Kanonizace Slezských písní

Vzhledem k složité historii Slezských písní nebylo snadné básně publikovat jako jednotné dílo, ke většině básní existovalo několik verzí, vytvořené buď časem pro zamezení cenzury, nebo nevysvětlitelně autorem. Po Vaškově smrti v roce 1958 se začaly konat konference a diskuse mezi literárními kritiky a historiky o tom, která z několika verzí Slezských písní má být kanonizovaná (Pešat, 1964). Dvě protichůdné strany v tomto sporu zastával Oldřich Králík a Miroslav Červenka. Profesor Oldřich Králík zastával názoru, že básně by měly být publikovány v jejich rukopisné podobě, případně v jejich prvních vydání, neboť veškeré pozdější úpravy narušovaly jejich kvalitu. Svůj názor podpořil dílem *Text Slezských písní* (1963), ve kterém básně ve své nejstarší podobě seřadil dle data vzniku. Věřil také, že takzvané jádro Slezských písní, tedy básně napsané do roku 1900, tvoří samostatný celek, kterému se Vaškova pozdější tvorba vydaná pod jinými pseudonymy nemůže rovnat. Druhý tábor tvořili Miroslav Červenka a Břetislav Štorek z Ústavu pro jazyk český, kteří s Králíkovou idealizací rukopisů nesouhlasili (Pešat, 1964). Autor se vyvíjel postupně, a i v jádru Slezských písní je možné nalézt prvky patetického stylu charakteristického pro Smila z Rolničky. Rozpolcenost k autorovi patří a není vhodné sbírku o tuto tak podstatnou vlastnost autora upírat. Podle Červenky je nutné respektovat změny, které Vašek ve svých básních provedl, i když pouze v omezené míře. Ke kanonizaci navrhoval vydání z roku 1928, které i navzdory změnám nijak nezjednotnilo kvalitu sbírky.

Diskuze nezůstala jen u kanonizace básní, jednou z hlavních otázek bylo, do jaké míry lze Petra Bezruče považovat za autora sbírky, když byla její část napsána pod jiným pseudonymem, a zda je vhodné sbírku publikovat jako jedno dílo, přestože sám autor se pod částí písní podepsal jiným pseudonymem. Jak tedy básně publikovat? Dle kvality, nebo jako celek? V prvním případě by bylo nutné upřednostnit jádro, které se dle podpůrců této možnosti jednoznačně vymezuje nad zbytek Vaškovy tvorby. Tato teorie si však nezískala mnoho zastánců, jelikož většina zúčastněných diskuze nesouhlasila s roztržením už tak fragmentovaného díla (Pešat, 1964). K rozuzlení došlo v roce 1963, kdy byla jako kanonická zvolena verze z roku 1930 obsahující jak jádro, tak pozdější tvorbu jediného autora sbírky Vladimíra Vaška. (Pešat 1964).

Teorie trojjedinosti Vladimíra Vaška

Oldřich Králík byl významným textologem, literárním historikem a bezručologem, který se zasadil o textologické zhodnocení sbírky a její vydání k podobě, jak ji známe i dnes. Dílu se věnoval v podstatě celý život a napsal o něm nemálo publikací. Především významná je jeho práce *Historický vývoj textu* (1967), ve kterém se mu podařilo básně seřadit podle data vzniku a vyjmenovat všechny úpravy, které na nich Vašek tvořil. Králík tyto změny odsuzoval, věřil, že je nejvhodnější kanonizovat básně v jejich nejstarší podobě a na změny zanevřít, jelikož většina oprav ubírala originálu na kvalitě. Králík byl také první, kdo přišel s takzvanou teorií trojjedinosti autora. Podle ní Vašek básně nevydával pod různými pseudonymy náhodně. Písně napsané pod jednotlivými jmény jsou diametrálně odlišné záměrně a nemohou se stylově zaměňovat. Každý Vaškův pseudonym má vlastní charakter a způsob tvorby a je nutno na něj takto nahlížet.

Ve svém díle Králík rozlišuje tzv. nacionalistické „jádro“ Slezských písní, vydané do roku 1900, sbírku popisných básní Smila z Rolničky, a básně ovlivněné světovou válkou od B.Šarka. Autorovu tvorbu lze takto rozdělit na jednotlivá období, která byť jsou od sebe tematicky i stylově vzdálená na míle daleko, přesto jsou všechna v určité míře bezručovská (Králík 1963). Králík dodává, že po roce 1900 osobnost Petra Bezruče umírá a Vašek se mění ve Smila z Rolničky, který píše milostné a popisné básně, ve kterých vzpomíná na domovinu a milostný život. Na druhou stranu Bohumila Šarka považuje za vyvrcholení tvorby autora unaveného válkou a pobytem ve vězení, kterým se Vašek ani nesnaží přiblížovat stylově ani tematicky ke své předchozí tvorbě. Jednotlivá období způsobená padesátiletým vývojem tvorby je třeba dle Králíka respektovat, jelikož patří do různých období autorova života a jeho tehdejších zážitků. Nic by nemělo nasvědčovat teorii, že by Vašek neměl být jediným autorem celé sbírky, jakoukoliv hypotézu tvrdící opak až do konce života Králík odmítal (Králík, Ficek 1978).

Rozdíl mezi kvalitou prvotních básní a všech ostatních si Králík vysvětluje průběhem jejich tvorby. Existují dva hlavní způsoby vysvětlení vzniku „jádra Slezských písní“. Podle tzv. Místecké teorie napsal Vašek většinu básní v letech 1891–1893 během svého pobytu v Místku. Na básně však zanevřel a rozhodl se je publikovat až poté, co na přelomu století zápasil s tuberkulózou. Za pobytu v Místku byl Vašek velmi aktivní, zde se seznámil také

s Ondřejem Boleslavem Petrem, bývalým učitelem a bohémem, se kterým se přátelil až do Petrovi náhlé smrti v roce 1893. Více se této osobnosti věnuje stejnojmenná kapitola.

Druhé vysvětlení vývoje básní předpokládá, že básně vznikly eruptivně náhlým přílivem inspirace v roce 1899 a Vašek je téměř ihned zaslal Herbenovi. Tato teorie podle Králíka lépe vysvětluje jedinečnost nejznámějších básní, ke kterým se sám Vašek ve své pozdější tvorbě už nikdy nepřiblížil. Místecká teorie je obecně málo uznávaná, neboť k ní chybí dostatečné důkazy. Explozivní povaha některých básní navíc nenasvědčuje, že by na ně mohl autor jen tak zapomenout a na sedm let je nechat schované. Proto se kritici kloní spíše k hypotéze, že básně vznikly nárazově ve velmi krátkém časovém intervalu kolem roku 1899 (Kosák, 2014). Tato teorie by podporovala i fakt, že Vašek o několik desítek let později básně několikrát opravoval, občas i tak, že se opravou vytratila jejich původní myšlenka nebo expresivita. Svými změnami se nikdy nedokázal vrátit do okamžiku jejich tvorby.

Na Králíkovy publikace nejednou reagoval další významný literární kritik a textolog Miroslav Červenka. Ve své publikaci *Slezské písně jako problém textologický a ediční* (2009) pokládá Králíkovy hypotézy za subjektivní a příliš zjednodušené. Přesné vymezení Vaškovy tvorby na jednotlivá období je podle něj nemožné, neboť Vaškovy stylistické prvky se prolínají napříč celou jeho tvorbou.

Teorie dvojího autorství

Výše zmíněný nenadálý úpadek v kvalitě Slezských písní dala nejednomu čtenáři nápad, že se krom stylu, tématu i kvality změnila i ruka, která básně psala. Uvažovalo se, zda Vašek své nejlepší básně neukradl svému otci, studentovi, příteli, nebo nějaké jiné neznámé osobě. K ani jedné z těchto tezí neexistovaly žádné důkazy, proto je většina kritiků při diskuzích o kanonizaci díla nebrala v potaz (Smolka, 2018).

Ondřej Boleslav Petr

Většina zastánců teorie dvojího autorství věří, že Vašek ty nejpodarenější básně z roku 1899 a 1900 opsal od svého přítele, zesnulého Ondřeje Boleslava Petra, se kterým se seznámil během svého pobytu v Místku. Ondřej Boleslav Petr byl bývalý učitel, později herec, který byl známý svou bohémskou povahou a protiněmeckými názory, za které byl vyloučen ze studií. Pro Vaška

představoval věrného přítele a mentora, který ho uvedl do tamějšího společenského dění (Drozd, 2003). Pro mnohé čtenáře tudíž představoval ideálního Petra Bezruče.

O vlivu Petra na Vaškovu tvorbu psal už Josef Vachala ve své knize věnované Vaškově životě na přelomu století nazvané *Místecký pobyt Vladimíra Vaška 1891-1893 a jeho přerod v Petra Bezruče* (Vachala, 1963). Tvrdil, že Petr byl literárně velmi aktivní a psal básně o Těšínsku, které stylem připomínaly dle svědků Slezské písně. Z jeho tvorby se však nic nedochovalo až na jednu krátkou báseň věnovanou rodnému statku. Vachalova kniha měla však pramalý ohlas, většina svědeckví, o která se opíral, pocházela z druhé až třetí ruky od osob, které už byly v době vydání knihy roky po smrti (Smolka, 2018).

Vztahu mezi Petrem a Vaškem se věnoval i Jiří Urbanec ve své knize *Mládí Petra Bezruče* (Urbanec, 1969). Urbanec se v knize snaží co nejpodrobněji popsat Vaškův život do roku 1900, soustředí se na jeho blízké okolí a pozornost věnuje i Petrovi, který podle něj nepopíratelně ovlivnil podobu Slezských písní. Urbanec nicméně odmítá možnost, že by některou z básní napsal právě Petr. Přesto argumentuje, že měl Petr na Vaška nepochybně velký vliv, který se odráží v raném období své tvorby. K tomu dodává, že je také možné, že Vašek použil některý z motivů z Petrovy tvorby anebo že některé básně napsali společně. Přesně se Petrův podíl na sbírce určit nedá, jelikož se z jeho tvorby nedochoval žádný vzorek vhodný ke komparaci (Urbanec, 1969).

Králík na Urbancovu publikaci reagoval chladně, s naprostým přesvědčením odmítá teorii, že by Petr mohl napsat třeba jen část nějaké písně. Jediná krátká báseň, kterou lze s přesností nazývat dílem Ondřeje Boleslava Petra podle Králíka nedokazuje, že by měl jakýkoliv literární talent. Veškeré dohady o tom, zda je Petr autorem Slezských písní považuje Králík za absurdní (Králík – Ficek, 1978). Možnost, že hrál Petr ve Vaškově tvorbě zásadní vliv, považuje pouze jako možný důkaz k doložení Místecké teorie zapříčiněný jejich přátelským vztahem až do Petrovi smrti.

Jan Drozd

Jan Drozd v roce 2003 navázal na studie Oldřicha Králíka, které rozvinul až do, v určité míře, extrémní podoby. Zatímco Králík rozlišuje Vaškovu tvorbu podle pseudonymu, pod kterým ji posílal Herbenovi, přičemž zastává názor, že autor dosáhl svého básnického vrcholu v roce 1900 a poté kvalita jeho básní upadla v takové míře, že poté už nedá hovořit o Bezručovi jako takovém, jelikož toto alter ego umřelo (Králík, 1963), Drozd tento přístup kritizuje, neboť je

podle něj nemožné, aby se autor během pár let tak drasticky změnil bez pádné příčiny. Tento propastný rozdíl mezi jádrem Slezských písní a pozdějšími básněmi vydanými pod jinými pseudonymy si Drozd vysvětluje tím, že se musí jednat o dva různé autory. Přičemž argumentuje, že druhým autorem je nepochybně právě již zmiňovaný přítel Ondřej Boleslav Petr.

Jan Drozd v zásadě s Urbancem souhlasí, Petr na Vaška nepochybně vliv měl a byl pro něj básnickou inspirací. Nicméně nezůstal jen u toho. Přiřítá mu autorství více než 20 básní ze sbírky. Králíkovy pochybnosti o literárním nadání Petra vyvrací tvrzením, že když se z jedné dochované básně nedá soudit talent Petra, stejně se z jakékoliv z Vaškových pozdějších výkonů nedá soudit hodnota Slezských písní. Svou teorii navíc rozvíjí i dál, Bezručovu tvorbu nedělí jen na jádro patřící Petrovi a ostatní básně, ve kterých se Vašek marně snažil Petrovi vyrovnat. Tvrdí, že Vašek k Petrovým básním, které mezi lety 1899–1900 Herbenovi poslal, údajně přimíchal i pár vlastních výtvorů. Přesně vymezuje 5 básní z onoho jádra, které se od zbytku liší, a tudíž údajně nemohly být napsány Petrem.¹ V těchto prvních pěti básních se se dle Drozda snažil Vašek Petrovi co nejvíce přiblížit, Drozd také neodmítá možnost, že jsou některé básně psány oběma autory společně (Drozd, 2003). Domnívá se, že veškeré další básně napsané po roce 1900 už patří výhradně Vladimíru Vaškovi. Tuto teorii Drozd podporuje podle něj nevylučitelnými argumenty týkajícími se rozdílů v životě a povaze obou autorů:

Petr byl známý rebel, kterého vyloučili za jeho protiněmecké chování ze školy. Poté působil jako učitel, jenže o místo přišel, když odmítal mluvit německy. Několik let cestoval jako kočovný herec a roku 1891 se vrátil zpátky do Místku, kde se seznámil s Petrem. Jeho život končí v roce 1893, podle Drozda nejspíše sebevraždou (Drozd 2003). Jeho životní příběh obsahuje všechny prvky, o kterých se Bezruč zmiňuje v jádru své sbírky. Naproti tomu Vašek žil klidný život jako poštovní úředník nejdříve v Místku a pak v Brně, mluvil v práci německy a měl i německou milenkou. Podle Drozda není možné, aby člověk takového charakteru byl schopný napsat buřičské básně o poněmčování Těšínska, které patří k neznámějším básním sbírky.

Sporné je autorství básní Vladimíra Vaška dle Drozda i proto, že básně posílal Janu Herbenovi v zásilkách anonymně, původně s nimi nechtěl mít nic společného a autorství odmítal přiznat. Po celý život nesouhlasil s tím, aby byly vydány v samostatné knize a nepomáhal nakladatelům

¹ Jedná se o básně Den Palackého, Hanácká ves, Kyjov, Plumlov, Ptení a Rektor Halfar.

s jejich rozdělením. Tento odtažitý postoj k dílu tak napovídá o jisté neintegritě básnického díla a jeho autora. V návaznosti na to Drozd upozorňuje na Vaškovo neustálé opravování a přepisování většiny básní až do jeho smrti v roce 1958. Změny, které prováděl, byly podle něj naprosto neopodstatnitelné a náhodné. Drozd vysvětluje, že Vašek neměl k básním žádnou citovou vazbu, nerozuměl jim, a tak nebyl schopný si uvědomit, že svými úpravami ničí literární hodnotu svého díla.

Drozdova kniha *Autoři Slezských písní* vydaná v roce 2003 odstartovala lavinu článků, ve kterých novináři a literární vědci reagovali na jeho kontroverzní teorii. Vesměs ale byly reakce na teorii dvojího autorství Slezských písní od odborné veřejnosti natolik negativní, že Drozd o dva roky později vydal *Otevřený dopis*, ve kterém se na kritiky obrací s dalšími argumenty. Za svým názorem si stojí a vyzývá k vytvoření skupiny nestranných expertů, kteří by byli ochotni se této tématice dále věnovat. Je vhodné dodat, že v této novější publikaci už netvrdí, že Vašek básně ukradl nebo slepě opisoval z toho, co našel. Uznává, že se Vašek nejspíš jen rozhodl pokračovat v dědictví svého zesnulého přítele. Sám Vašek o Petrovi nikdy nemluvil a jeho smrti už neexistují důkazy, které by mohly Drozdovu teorii potvrdit, či vyvrátit. Hypotéza, že sbírku napsal Petr spolu s Vaškem, si i dnes najde své zastánce (především v laické veřejnosti), ale i odpůrce. S Drozdem souhlasí například zpěvák Jaromír Nohavica, který se otázce autorství Slezských písní věnuje už několik desítek let a svým zkoumáním navazuje na Drozdovu práci.

Jaromír Nohavica

Ostravský písničkář Jaromír Nohavica se Slezskými písněmi zabývá už téměř čtyřicet let. Kniha ho zajímá už od raného dětství, tudíž k ní má silný emotivní vztah. Jako písničkář pocházející ze stejného kraje, který některé Bezručovy básně zhudebnil, má navíc na sbírku jiný úhel pohledu než odborná veřejnost (Jaroušek, 2014). Všimá si melodičnosti a přesných reálií, kterým většina literárních historiků zabývajících se sbírkou nevěnovala pozornost.

Na myšlenku, že básně mohou být napsány dvěma různými autory, ho přivedla v roce 1980 kniha *Jen Jedenkrát: zázilky Času 1899-1914* (1980). Po přečtení sám uznává, že nebyl schopný pochopit, že by tak rozličné básně mohl mít na svědomí pouze jeden autor. Poté, co nenalezl odpovědi v odborné literatuře, se rozhodl záležitost zkoumat na vlastní pěst (Motýl, 2014). Většinu svých argumentů zakládá na vlastních pozorováních, porovnává děje básní s reáliemi

z života Petra i Vaška, obchází archívy a knihovny, a dokonce navštěvuje města, kterou podrobně porovnává s jejich opisem v básních. Hledání důkazů potvrzujících autorství jádra Slezských písní Ondřeje Boleslava Petra považuje za svého koníčka, tudíž se zabývá sběrem materiálu, nikoliv jejich publikací v literárních magazínech. Výsledky svého více než třicetiletého zkoumání dosud nikdy oficiálně nepublikoval. Příležitostně své argumenty prezentuje pro studenty středních škol na přednáškách, raritně na koncertech. Nejnověji se problematice věnoval na začátku roku 2020 během svého třídního koncertu v pražské Viole (Vetešková, 2020).

Nohavica ve svých argumentech navazuje na Drozdovu publikaci, kterou rozšiřuje o výstupy z vlastního zkoumání. Tvrdí například, že Vašek neměl cit pro hudbu (viz Závodský 1977), proto není možné, aby napsal „melodické“ básně typu Maryčka Magdónova. Taktéž za svého pobytu na Slezsku neměl čas kraj dobře poznat, což se prokazuje na básních, které nekorespondují s reálnými místopisy (Hrabyň např.). Na básních se jasně ukazuje vztah autora ke své domovině. Vašek, který strávil většinu svého života v Brně, neměl čas ke krajině přilnout. (Jiroušek 2014). Spoustu věcí přepisoval z deníku, který mu po Petrově údajné sebevraždě zůstal, mnohdy však nepřesně. Jako příklad jeho chybných přepisů uvádí příjmení Maryčky Magdónové - Magdón se zrodil z polského příjmení Magdoń, ve kterém Vašek, neznalý polštiny, špatně opsal čárku (Kosák, 2012). Přesný výpis jeho argumentů je nicméně nemožný, jelikož je nikde nepublikoval a na svých přednáškách uznal: „Bude to pouze úvod do tématu, nepřímé důkazy, zamyšlení. Nejde o to, abyste mé závěry přijali. Chci, ať se nad nimi zamyslíte,“ (Jiroušek, 2014). V roce 2020 v rozhovoru pro rozhlas uznává, že se jedná o obsáhlé, dosud neprozkoumané téma, co si vyžaduje několikaleté studium a Nohavicův cíl není své argumenty jednoznačně prokázat, nýbrž podnítit veřejnost k tomu, aby se o sbírku nadále zajímala, případně nalézt někoho, kdo se tématem bude zabývat podrobněji. (viz Nohavica 2020 in: irozhlas).

Kritika teorie dvojího autorství

Přestože se o autorství sbírky polemizovalo léta, nikdy nebyly předloženy argumenty, které by mohly jednoznačně prokázat, že část básní napsal Ondřej Boleslav Petr. Na Jana Drozda i Jaromíra Nohavicu reagovala většina odborné veřejnosti negativně (. V této podkapitole budou představeny argumenty vyvracující teze Drozda i Nohavici, které v průběhu let přednesl Zdeněk Smolka, který se zúčastnil několika diskusí s oběma autory

Proti Drozdovi ale stále stojí valná většina textologů, Smolka, který se diskutuje s Drozdem účastnil osobně, roku 2008 napsal článek, ve kterém se zpětně vrací k celému problému a jeden po druhém Drozdovy argumenty vyvrací (viz Smolka 2008 in: Host). Také nejnovější kritické vydání Slezských písní z roku 2014 v nakladatelství Academia možnost druhého autora naprosto popírá a označuje všechny teorie tohoto typu jako „pusté a nepodložené hypotézy“ (Kosák in Bezruč, 2014).

Nohavicovy hypotézy nebyly v literárních kruzích přijaty, čehož si je písničkář vědom. V článku Petr Bezruč plný záhad (2018) napsaným po setkání s Nohavicou na konferenci konané 21-22.9.2017 reaguje na jeho přednášku právě na toto téma a zpochybňuje jeho argumenty.

Jakožto hudebník si Nohavica při prezentaci svých hypotéz zakládá na svém citu pro melodii, díky které si je schopný uvědomovat rozdíly mezi kvalitou „melodičnosti“ jednotlivých písní. Smolka nicméně tvrdí, že básně jako takové hudební melodii vzhledem ke své formě nemají. Autor při psaní básně nebere v potaz možnost, že by se je jednou někdo rozhodl zhudebnit, proto je Vaškův údajný neexistující hudební sluch jako argument zpochybňující jeho autorství irelevantní. (Smolka, 2018).

Dalším z častých hypotéz Drozda i Jaromíra Nohavici je vztah mezi básněmi a fakty z života jejich domnělého autora. Vašek byl dle několika různých svědectví plachý úředník, zatímco Petr byl známý buřič a bohém. Jejich rozdílné životy se dle obou autorů na básních musely prokázat. Smolka uznává, že se v básních objevují s životopisecké prvky, jenže přesto se nejedná o přesný popis reality. Hranice mezi skutečností a metaforou je v literatuře občas velmi těžko zpozorovatelná. Bez autora přesného potvrzení není možné brát skutečnosti popsané v jeho díle za jednoznačná fakta. Není proto možné, je považovat za prvky určující autora.

Celkově se dá shrnout, že Smolka označuje Nohavicovy teorie jako nedostatečně podložené. Kritizuje ho za ovlivňování publika a přednášení vědecky neprokazatelných domněnek a ignorování veškeré kritiky od odborné veřejnosti. (Smolka, 2018)

Praktická část

V následující části práce provedeme analýzy podobnosti vzorků básní Slezských písní kvantitativními metodami určování autorství. Básně budou analyzovány na základě zastoupení kompletní slovní zásoby, funkčních slov a nízko frekventovaného lexika (Faltýnek – Matlach – Owsianková, 2020). Veškeré analýzy budou prováděny v programu QUITA (Quantitative Index Text Analyzer).² Nejdříve budou sbírky srovnány na základě jejich celkové slovní zásoby bez jakýchkoliv jejich specifikací. Následovat bude analýza podobnosti na základě použitých funkčních slov. Více o funkčních slovech a jejich významu v určování autorství bude uvedeno v podkapitole Funkční slova. Tento postup bude proveden pro větší přesnost dvakrát – vždy pro skupinu o 50 slovech extrahovaných z celkového souboru Slezských písní. Po analýze funkčních slov budou vzorky básní srovnávány na základě výskytu hapax legomen – slov použitých v textu pouze jednou. Tomuto specifickému typu slov bude věnována podkapitola Hapaxy. Výsledky jednotlivých analýz budou vyobrazeny v grafech a vzájemně srovnávány. Na závěr bude vyhodnoceno, která z analýz se jeví jako nejvěrohodněji vystihující konstelace srovnávaných textů.

Pro analýzy jsme básně rozdělili do skupin popsaných v teoretické části této práce – viz teorie dvojího autorství Jana Drozda. Zajímáme se o totožné autorství těchto skupin básní, ptáme se také na to, zda analyzované vzorky spadají do tzv. „jádra Slezských písní“. Srovnávané skupiny básní připravil pro naši analýzu dr. Michal Kosák (viz Kosák 2005, Kosák – Fleišman 2012). Současně se vzorky básní Slezských písní budeme analyzovat soubory textů čtveřice dalších českých básníků. Tito autoři byli vybráni jako komparační skupiny odlišných autorských stylů (pracujeme vždy se dvěma sbírkami od každého z autorů), nebylo naším záměrem je vybrat pouze na základě jejich časové či stylové blízkosti s Bezručem, ačkoliv jsme k tomuto faktu přihlédlí (tzn. že např. nekomparujeme Bezruče s díly mnohem pozdějšími nebo prozaickými). Téma vzorků v tomto případě nebylo bráno v potaz, jelikož tematické odlišnosti se nacházejí i mezi jednotlivými skupinami výchozí sbírky Slezských písní.

Cílem praktické části práce není zcela nepochybně dokázat, že byla některá část Slezských písní plagiovaná. Chceme pouze básně porovnat za pomoci konkrétních metod určování autorství, z nich metoda hapaxová doposud nebyla na Bezručovo dílo použita, jelikož je všeobecně nová (Faltýnek – Matlach, 2020). Výsledkem zkoumání by mělo být seskupení textů napsaných jedním autorem na základě podobnosti analyzovaných rysů, pokud možno v každé z analýz.

² <https://kol.ff.upol.cz/quita/>

V případě, že se některá z pěti skupin Slezských písní jasně vyčlení od ostatních, mohlo by nás to vést k názoru, že pochází z pera jiného autora než ty ostatní, což by podporovalo teorii Jana Drozda. Pokud by se tak stalo, bylo by vhodné sbírku vyhodnotit dalšími kvantitativními metodami určování autorství v budoucích pracích a teorii dvojího autorství dále prověřit pro dosažení přesnějšího výsledku.

Vzorek

Porovnávat budeme básně v jejich nejstarší doložené podobě, abychom zamezili ovlivnění výsledků mnohými opravami a změnami, které Vladimír Vašek v průběhu let prováděl. Srovnávané skupiny básní pro analýzu připravil dr. Michal Kosák (viz Kosák 2005, Kosák – Fleišman 2012). Slezské písně budou rozděleny do následujících skupin:

- Skupina OBD obsahující básně přisuzované Ondřeji B. Petrovi
- Skupina VV obsahující básně přisuzované Vladimíru Vaškovi zařazené do jádra sbírky
- Skupina PB obsahující zbylé blíže neurčené básně z jádra sbírky
- Skupinu VV mimo jádro obsahující zbylé básně přisuzované Vaškovi
- Skupina PB mimo jádro obsahující zbylé blíže neurčené básně

Skupina + počet slov	Seznam básní
PB 3466 slov	Bernard Žár, Červený květ, Čtenáři veršů, Dědina nad Ostravicí, Domaslovice, Dombrová I, Dva hrobníci, Dvě dědiny, Chycený drozd, Jedna melodie, Koniklec, Národní maškaráda, Oni a my, Opava, Par nobile, Paskovské rybníky, Pětvald I, Pětvald II, Pole na horách, Praga caput regni, Setkání, Střebovský mlýn, Sviadnov I, Tošonovice, Úspěch
PB mimo jádro 4243 slov	1864-1904, Beze mne a se mnou, Didus ineptus, Dragon Hanys Blendovský, Holderlin nad Neckarem, Kalina (Kde jsi vzala...), Kalina (Na hrobech...), Kalina (Nad řekou stojíš...), Kaštany, Krásné pole, Labutinka, Lazy, Melč, Mohelnice, Nápis na hrob bojovníkův, Nový mlýn pod Brnem, Pluh, Plumlov II, Polská Ostrava, Pyšný Janek, Sedm Havranů, Smrt Caesarova, Souputnice, Starček, Zem pod horama, Žil jsem jak stepní kůň svobody, Žně

OBP 5129 slov	70000, Blendovice, Dombrová II, Hrabyň, Hučín, Já, Jen jedenkrát, Kdo na moje místo, Kovkop, Leonidas, Ligoťka Kameralna, Markýz Géro, Maryčka Magnodnova, Michalkovice, Motýl, Návrat, Ondráš, Ostrava, Osud, Slezské lesy, Škaredý zjev, Uhni mi z cesty, Vrbice, Z Ostravy do Těšína, Žermanice
VV 871 slov	Den Palackého, Hanácká ves, Kyjov, Plumov, Ptení, rektor Halfar
VV mimo jádro 2661 slov	Děrné, Dvě mohyly, Hanys Horehled', Hradec-Podolí, Idyla ze mlýna, Mladá žena, Papírový Mojšl, Střebovský mlýn, Valčice

Tabulka 1: Seznam básní Slezských písní v jednotlivých skupinách dle podezření o dvojím autorství dle Jana Drozda

Texty Slezských písní srovnáváme s básnickými sbírkami čtyř českých básníků. Seznam jednotlivých básní těchto sbírek je uveden v příloze této práce. Autory, jejich díla a počet slov uvádíme v Tabulce 2.

Autor	Dílo	Počet slov
František Gellner	Po nás at' přijde potopa	2946
	Radosti života	4213
František Halas	Ladění	5565
	Tvář	4005
Jiří Wolker	Host do domu	4596
	Těžká hodina	5565
Stanislav Kostka Neumann	Bohyně, světice, ženy 1	5418
	Bohyně, světice, ženy 2	4199

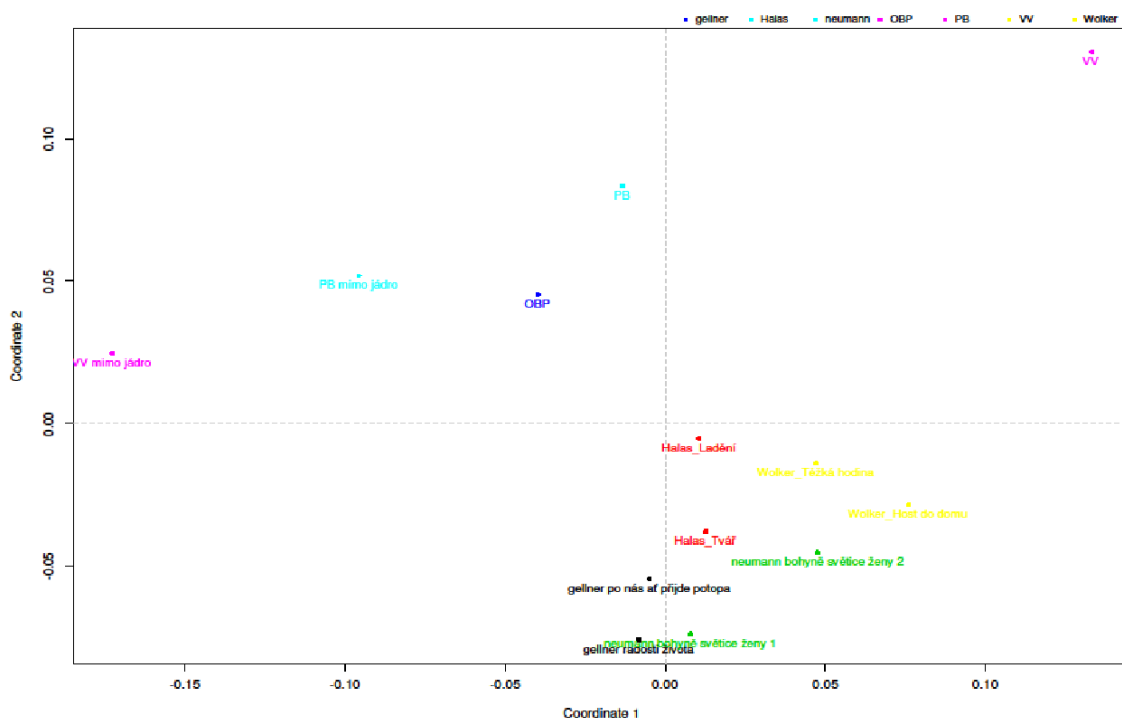
Tabulka 2: Seznam autorů a jejich děl použitých ke komparaci autorských stylů básní Slezských písní.

Jak již bylo zmíněno výše, díla nebyla vybrána na základě podobného data vzniku ani tématu se Slezskými písněmi. Jelikož se analýza bude zaměřovat na prvky slovní zásoby rozlišující idiolekt jednotlivých autorů, nikoliv například obsah, délku verše, ani jeho rytmus, pro potřeby práce považujeme za nutné především to, aby byly skupiny básní/sbírky obdobné délky. V našem případě se valná většina porovnávaných vzorků pohybuje v rozmezí 3500–5500 slov. Nejmenší porovnávanou sbírkou je *Po nás ať přijde potopa* od Františka Gellnera, která má pouze okolo 2946 slov. Ze Slezských písní tvoří nejmenší skupiny básně přiřazené k Vladimíru Vaškovi – VV a VV mimo jádro, které mají dokonce méně než 3000 slov (konkrétně méně než 871 slov pro VV a 2 661 slov pro skupinu VV mimo jádro). Vzhledem k této skutečnosti budou na konci praktické části analýzy provedeny znovu: tentokrát budou porovnávány jen básně tvořící SP, ve kterých bude korpus textů od Vladimíra Vaška (VV a VV mimo jádro) sjednocen, abychom zamezili ovlivnění výsledků délkou porovnávaných materiálů. Nemůžeme totiž vyloučit, že právě délka je rozhodující vlastností seskupování textů. Třetí nejmenší vzorek představuje skupina textů z jádra SP přiřazených Petru Bezručovi, která má okolo 3466 slov. Naproti tomu dílo *Bohyně, světice, ženy* od Stanislava Kostky Neumanna muselo být kvůli délce rozděleno na dva soubory.

[Analýza kompletní slovní zásoby](#)

V prvním kroku naší analýzy přistoupíme k porovnání vzorků na základě jejich celkového slovníku. To nám umožní sledovat podobnost vzorků na základě frekvence jednotlivých jazykových prostředků v plné šíři jejich zastoupení. Jsme si vědomi, že tento typ analýzy je značně ovlivněn tématem textů (druhem zastoupených plnovýznamových slov), stylem textu (druhem zastoupených funkčních slov a gramatických tvarů) a délkou textu. Při interpretaci k těmto okolnostem přihlídneme. Je také třeba upozornit na skutečnost, že z hlediska srovnání kompletní slovní zásoby textů můžeme přihlížet rovněž k zastoupení a frekvenci n-gramů slovních forem (Tesar – Poesio – Strnad – Jezek 2006.). V této analýze se ovšem přidržíme pouze slovníku samostatných slovních forem.

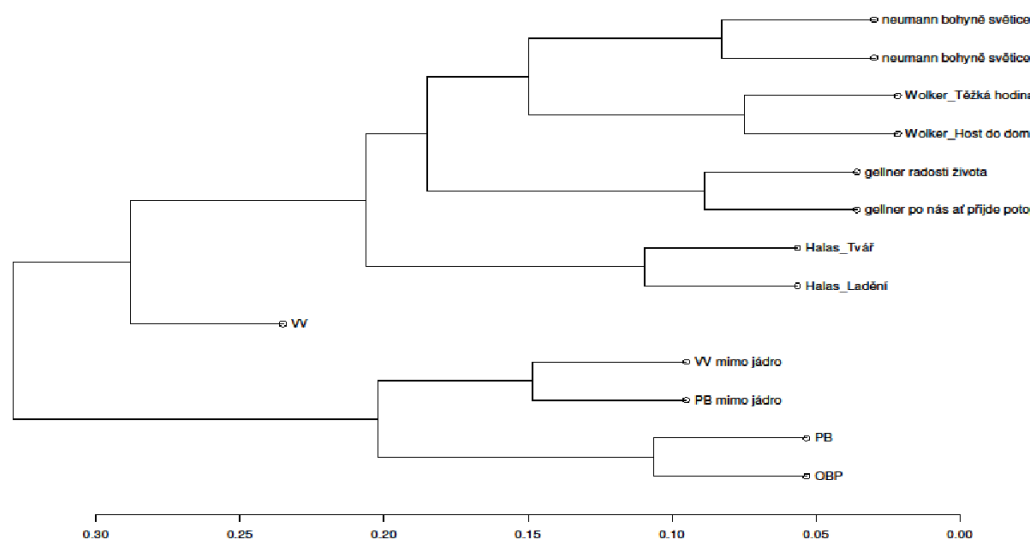
Při porovnávání celého slovníku textů se na výsledcích projevuje tematika textu. Dá se předpokládat, že při vyhodnocování podobnosti vzorků tímto způsobem se budou vzorky s příbuzným tématem od různých autorů podobat více než texty jednoho konkrétního autora s tématem rozdílným. Texty v porovnávaném korpusu nebyly zvoleny dle tematické ani stylové příbuznosti, jak už bylo zmíněno výše, obsahují tedy básně různých typů: vlastenecké, milostné, vitalistické aj. Je tedy pravděpodobné, že metoda porovnávání slovníku textů v celé jejich šíři nebude dostatečně efektivní z hlediska autorské atribuce, že budou texty zařazeny do skupin na základě obsahu. Z tohoto důvodu bude větší důraz kladen na výsledky analýz funkčních slov a hapaxů, tyto přístupy totiž z pravidla nejsou ovlivněny tématem vzorku (Baayen a kol. 1996).



Obrázek 1: Analýza Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě kompletní slovní zásoby, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení pomocí metrického vícerozměrného škálování (viz Torgerson 1952).

V obrázku 1 si lze všimnout dvou jasně vymezených skupin – sbírky Slezské písně v levé horní části grafu a shluku sbírek zbylých porovnávaných děl vpravo dole. Skupina básní VV se jako jediná umístila v pravém horním rohu, je ovšem umístěna v nejbližším postavení ke

zbylým vzorkům Slezských písní. Tato skutečnost může být zapříčiněna velikostí vzorku, jedná se totiž o vzorek s nejmenším počtem slov. Ostatní skupiny se umístily blízko sobě.



Obrázek 2: Analýza Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě kompletní slovní zásoby, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení pak pomocí hierarchického shlukování.

Podobnost jednotlivých slovníků všech básnických sbírek byla vyobrazena také pomocí metody hierarchického shlukování. Výsledkem je dendrogram, na kterém jsou texty s nejpodobnějšími jazykovými prvky umístěny do dvojic. Na Obrázku 2 si můžeme všimnout, že se sbírky od stejného autora umístily bez výjimky u sebe. Toto zobrazení nám poskytuje přesnější pohled na srovnání textů Slezských písní. Nejvíce podobné se dle celkového obsahu slovníku zdají být básně patřící do kanonického jádra díla. Dále se pak básně sbírky, jež podle literárních kritiků do jádra nespádají, shlukly vedle sebe v následující větvi dendrogram. Stejně jako u předchozího zobrazení se jasně vymezila skupina VV, která tvoří samostatný přechod mezi Slezskými písněmi a porovnávanými sbírkami jiných autorů. Znovu můžeme usuzovat na to, že se tak stalo na základě odlišné délky vzorku VV. To ostatně prověříme dále na základě sloučení tohoto vzorku se vzorkem VV mimo jádro.

Analýza celkového obsahu slovníku vybraných textů se překvapivě ukázala jako přesná. Dvojice sbírek od stejného autora se vždy umístily u sebe. V návaznosti na Drozdovu teorii dvojího autorství nám tato analýza neposkytuje žádný důvod k předpokladu, že by některá ze

skupin básní Slezských písní měla být napsána jiným autorem než zbytek sbírky. Protiargumentem k tomuto závěru je samozřejmě již několikrát zmíněná tematická blízkost jednotlivých autorských textů. V následujících krocích budou skupiny analyzovány na základě specifitějších prvků slovní zásoby, tak abychom i vliv tematiky textu mohli omezit.

Funkční slova

V následujícím oddíle provedeme analýzu dvou sad funkčních slov ve sledovaných vzorcích básní. Z jednotné skupiny SP bude extrahováno celkem 100 funkčních slov. Ve skupinách po 50 budou tato slova následně analyzována a vyhodnocována metodou Burrowsovy delty, to znamená za pomoci manhattanské metriky použité na z-skóre vybraných položek slovníku. Tento přístup k určování autorství je jedním z nejúspěšnějších postupů v počítačové stylistice (Burrows 2003, Juola 2008, 241–249). Více o tomto typu analýzy píše například Dupal 2017. I k této analýze bude použit software QUITA. Dříve, než přejdeme k samotné analýze, je třeba vysvětlit, co funkční slova jsou a proč jsou vhodná k rozpoznávání autorského stylu.

Frekvence funkčních slov je všeobecně považována za jeden z nejvěrohodnějších stylometrických prvků (Kestemont, 2014). Průkopníky používání funkčních slov k vymezení autorského stylu jsou Mosteller a Wallace, kteří ve své studii z roku 1964 zavádějí pojem funkční slova. Poprvé se tak namísto plnovýznamových slov nebo např. nízko frekventovaných autorsky specifických slov zaměřují na vzájemné poměry slov, která jsou všeobecně velmi frekventovaná, ale nesou pouze gramatický význam. Mezi taková slova patří například zájmena, spojky, předložky, modální a pomocná slovesa a členy. Ve své studii používají tento typ slov k tomu, aby odhalili autora Listů federalistů (The Federalist papers). Použití slov, která nenesou nociónální význam, se v jejich studii ukázalo jako velmi přínosné. Výhody využití právě funkčních slov pro atribuci autorství dále uvádí Binongo ve svém článku z roku 2003. Můžeme uvést některé z nich:

- Slova nesoucí pouze gramatický význam mají podle něj větší význam při určování autorství, jelikož nejsou závislá na tématu ani žánru porovnávaných děl.
- Funkční slova tvoří určitou „kostru“, která na sebe navazuje slova plnovýznamová (content words). Nejsou navíc ovlivněna dobou vzniku textu – to znamená, že Binongo nepředpokládá výraznou proměnu těchto slov v průběhu vývoje jazyka, tzn. prudké

změny v používání gramatických jevů, vyjadřovaných za spoluúčasti funkčních slov, uživateli jazyka v průběhu času.

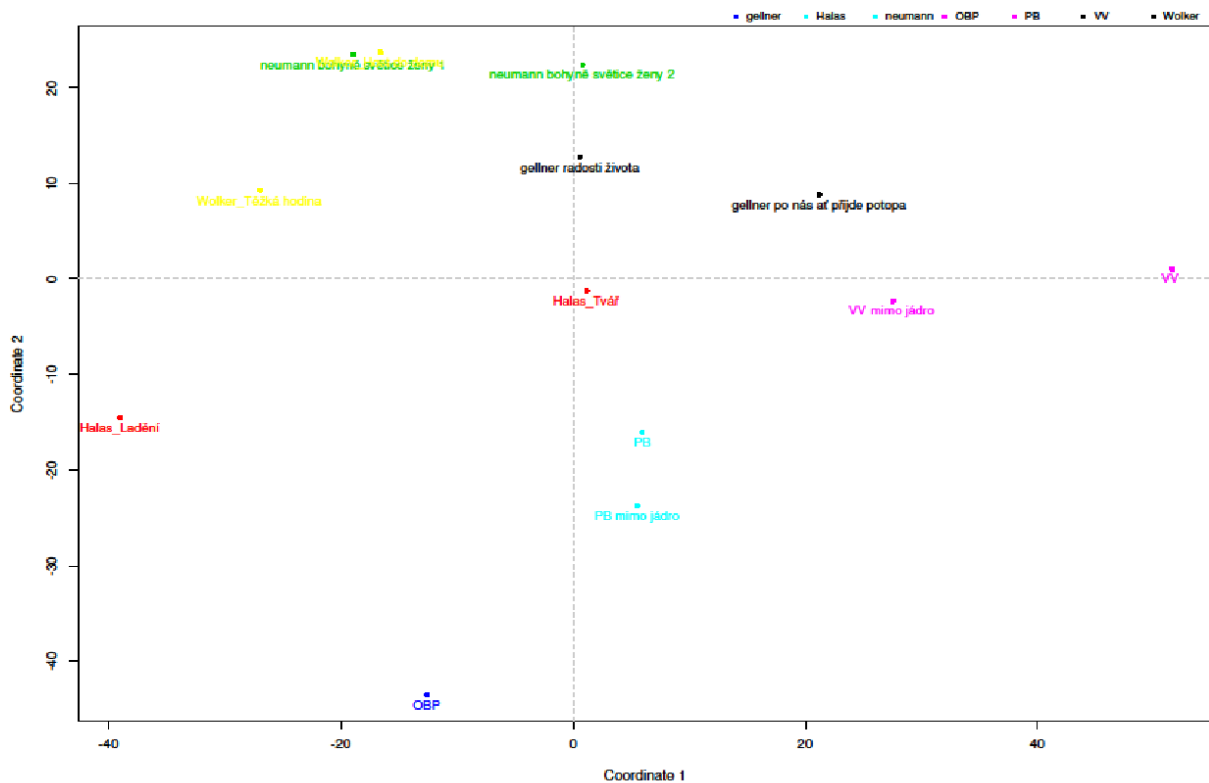
- Další výhodou je fakt, že se ve většině případů nedají ohýbat. V našem případě výjimku tvoří např. zájmena a pomocná slovesa. Tato výhoda je směřována k určování autorství v jazycích bez větší flexe, což je právě případ Binonga.
- Nedají se zaměňovat, jejich počet je mnohem stabilnější než u slov plnovýznamových.
- Poslední výhodou funkčních slov je, že se jejich použití nedá vědomě ovlivňovat, jsou pro autora skrytým prostředkem jeho individuálního stylu. Takovou vlastnost většině slovní zásoby nepřipisujeme, protože je spojena s vyjadřováním obsahu sdělení a zasazením tohoto obsahu do kontextu, což podléhá vůli autora a výběr slov je tak kontrolován vědomou volbou (Binongo 2003, Kestemont 2014).

Studii zaměřených na srovnávání textů na základě nejfrekventovanějších funkčních slov existuje mnoho, liší se i navrhovaná velikost vzorku funkčních slov, která se má pro co nejpresnější analýzu použít. Například Juola a Baayen ve svém článku z roku 2003 použili k analýze 164 nejčastějších jednotek včetně diakritických znaků. V naší analýze bude korpus básní analyzován na základě dvou sad nejfrekventovanějších funkčních slov obsažených v sjednoceném souboru Slezských písní. Jelikož se jedná o básně, ve kterých se diakritické znaky objevují jen zřídka, nebude na ně brán zřetel. Každá sada bude obsahovat 50 jednotek.

Analýza první sady funkčních slov

být	a	se	v	na	s	to	co	když	k
za	po	do	jak	z	si	jen	jako	už	tak
jež	že	já	ty	i	pod	ve	o	nad	mít
až	tu	však	pro	tam	kde	bez	kdo	od	až
ještě	my	by	u	ten	můj	jenž	svůj	bych	před

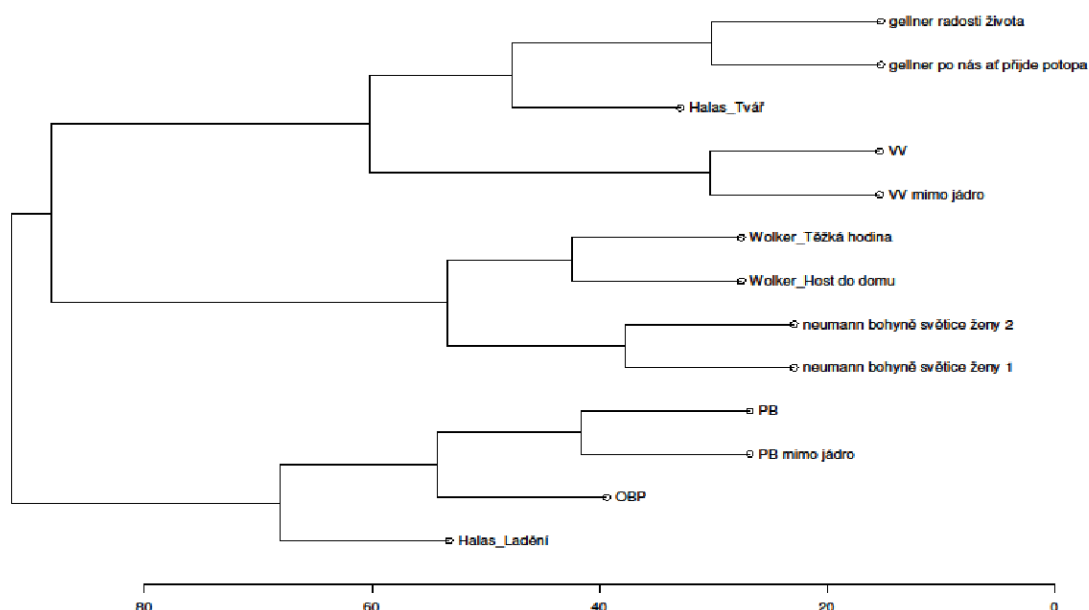
Tabulka 3: Seznam padesáti nejfrekventovanějších funkčních slov ve sbírce Slezské písně.



Obrázek 3: Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě z-skóre padesáti funkčních slov pomocí manhattanské metriky, zobrazení pomocí metrického vícerozměrného škálování.

Srovnání vzorků básní na základě frekvence padesáti nejčastějších funkčních slov vyskytujících se v díle Slezské písně se ukázalo jako dobře orientující z hlediska autorské atribuce. Texty určitého autora se vždy objevují blízko sebe. Překvapivá je vzhledem k tomu dvojice děl Františka Halase, jejíž prvky se jako jediné z celého korpusu neumístily v těsné blízkosti. Dalším nečekaným detailem je splynutí první části sbírky *Bohyně, svěťice ženy* od Stanislava Kostky Neumanna s *Hostem do domu* od Jiřího Wolker. Navzdory tomuto prolínání se sbírek odlišných autorů se ale v obou případech druhá část dvojice nalézá nedaleko. Taktéž si lze všimnout skupiny OBP, která se od ostatních textů, včetně těch patřících do sbírky Slezských písní, vzdálila až na spodní okraj levé části grafu. Zároveň můžeme konstatovat, že vzhledem k ne zcela jednoznačnému přiřazení dvojic autorských textů, které nám slouží ke konfrontaci stylu Slezských písní (Gellner, Halas, Neumann, Wolker), se metoda funkčních slov jeví být méně úspěšnou ve srovnání s výše provedeným srovnáním za užití celého slovníku (shlukování zde bylo z tohoto pohledu bezchybné). Protiargumentem je zde opět tematická blízkost textů, předpokládáme, že při použití jiných děl by situace nebyla vzhledem k celkovému slovníku

takto přehledná a že naopak autorský styl by se projevil v z-skóre funkčních slov. Nakonec můžeme oponovat tak, že pro texty v rozsahu básnických sbírek, jaké analyzujeme zde, jsou z hlediska analýzy autorství vzhledem k stylu autora funkční slova méně bezpečná než plný rozsah slovníku – to je nutné dále prověřovat jinde.



Obrázek 4: Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě frekvence padesáti funkčních slov pomocí manhattanské metriky, zobrazena za pomoci hierarchického shlukování.

Zobrazení pomocí hierarchického shlukování nám podává přesnější náhled na výsledky analýzy padesáti funkčních slov ve vzorcích, než tomu bylo u grafu MDS. V dendrogramu se jasně spárovaly dvojice textů od autorů Františka Gellnera, Vladimíra Vaška, Jiřího Wolker a Stanislava Kostky Neumanna. Obě skupiny spojené s pseudonymem Petr Bezruč se taktéž vyskytují na jedné větvi blízko sebe. Lze si všimnout, že skupina OBP se, jak již vyplývá z předchozího zobrazení pomocí metrického vícerozměrného škálování, nejvíce podobá skupinám PB a PB mimo jádro. Mimo společnou větev se ovšem v grafu umístila díla Františka Halase – na rozdíl od ostatních autorů neutvoří těsný autorský pár, každé z děl se nachází na opačné straně dendrogramu. Z analýzy vyplývá, že se *Tvář* vzhledem k funkčním slovům podobá více dílům Františka Gellnera, zatímco *Ladění* patří více ke většině skupin tvořících Slezské písně.

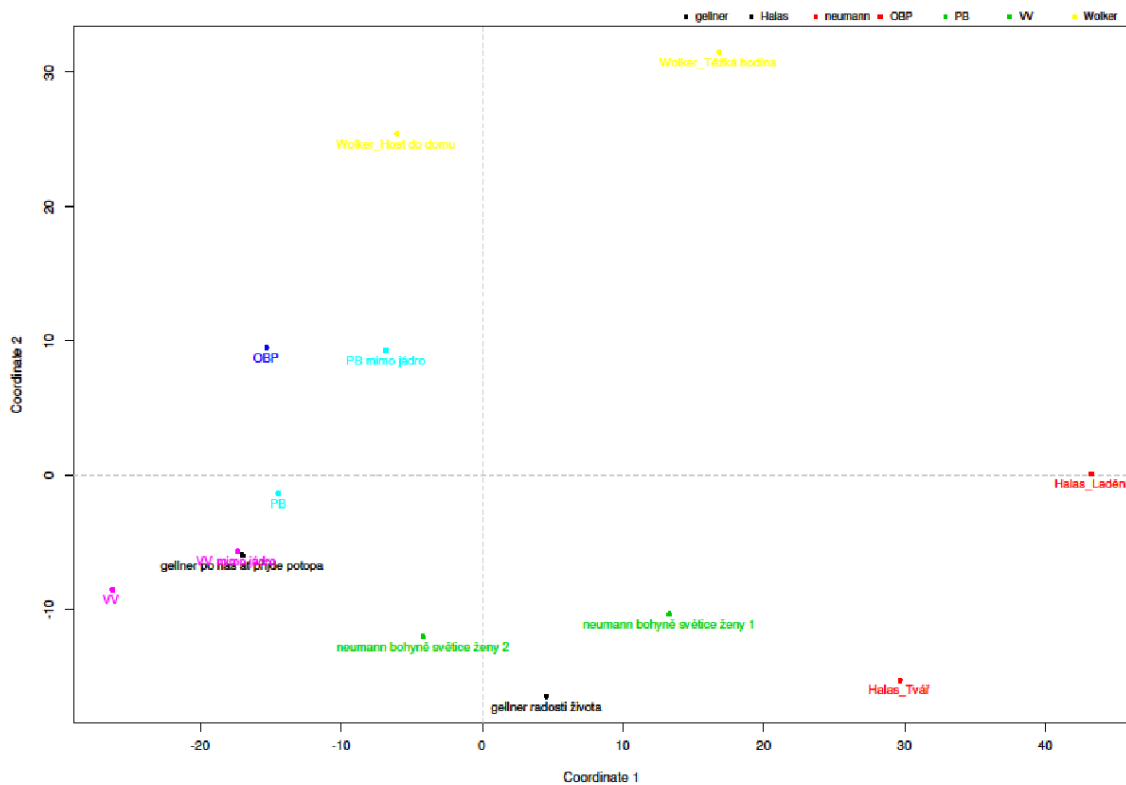
První analýza funkčních slov ukazuje na odlišnou charakteristiku básní ve skupině VV (Vladimíra Vaška), které nejsou v blízkosti dalšího klastru vzorků slezský písní. Tento fakt ale nebudeme brát jako určující, protože analýza v dalších shlukováních nenaplnila očekávání v autorské atribuce v jednoznačném stupni určitosti.

Analýza druhé sady funkčních slov

Jako další provedeme analýzu podobnosti na základě výpočtu z-skóre funkčních slov pomocí manhattanské metriky znovu stejným způsobem. Tentokrát budeme zkoumat zastoupení a frekvence druhé sady padesáti nejčastějších funkčních slov ze sbírky Slezské písně (tzn. 51 až 100 nejčastěji zastoupených funkčních slov). Jako v předchozím kroku budou výsledky analýzy zobrazeny ve dvojici grafů, metrického škálování a hierarchického shlukování.

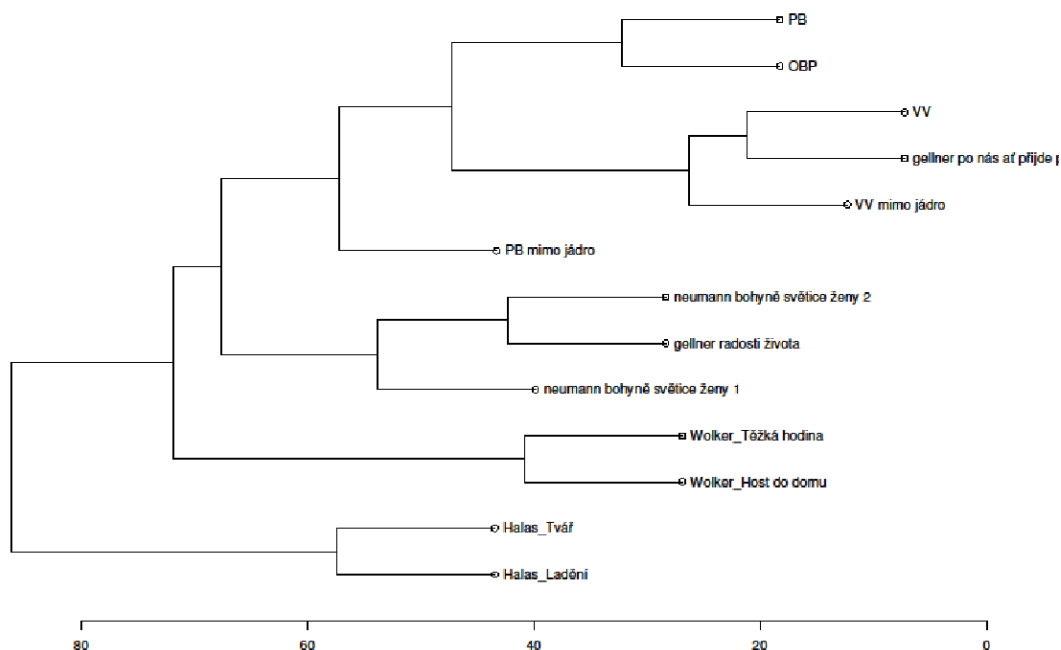
ta	tvoje	než	ale	přes	ne	zas	on	mezi	již
ať	jejich	teď	proč	ona	vy	jednou	snad	jeho	pak
jenom	tady	nic	sám	vše	také	naše	to	kdy	jestli
ach	který	první	ani	či	li	vždy	každý	necht'	její
jest	všech	nikdy	kam	ke	níž	zde	víc	tolik	trochu

Tabulka 4: Seznam padesátého prvního až stého nejfrekventovanějšího funkčního slova ve sbírce Slezské písně.



Obrázek 5: Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě z-skóre padesátého prvního až stého funkčního slova za pomoci manhattanské metriky, zobrazení pomocí metrického vícerozměrného škálování.

Obrázek 5 ukazuje podobnosti jednotlivých sbírek na základě frekvence druhé sady nejčastějších funkčních slov extrahovaných z celého korpusu Slezských písní. Zobrazení vytvořené metrickým vícerozměrným škálováním přineslo zajímavé, i když očekávatelné, poznatky. Jediný pár, který se vyskytuje ve stejné části dendrogramu, tvoří díla Františka Halase, která se obě nachází v pravém dolním rohu. Skupiny textů Jiřího Wolker se taktéž umístily relativně blízko sebe v horní části grafu. S ohledem na skupiny SP si můžeme všimnout, že dílo *Po nás at' přijde potopa* od Františka Gellnera se umístilo v těsné blízkosti od skupiny VV mimo jádro. Můžeme tedy konstatovat, že méně frekventovaná funkční slova přináší o autorském původu textů mnohem méně přesvědčivý obraz než funkční slova nejfrekventovanější. Tento názor by mělo podtrhnout následné zobrazení v dendrogramu.



Obrázek 6: Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) na základě z-skóre padesátého prvního až stého funkčního slova za pomoci manhattanské metriky, zobrazení za pomoci hierarchického shlukování.

Graf vytvořený metodou hierarchického shlukování nám poskytuje přesnější pohled na seskupení jednotlivých sbírek dle jejich podobnosti ve frekvenci druhé sady nejčastějších funkčních slov. Můžeme si všimnout, že touto metodou byly správně přiřazeny sbírky Františka Halase a Jiřího Wolкера. Taktéž básně patřící do kanonického jádra Slezských písní se umístily u sebe ve vrchní části dendrogramu. Zobrazení ale ukazuje neodpovídající závěry u děl Františka Gellnera, která prokazují větší míru podobnosti funkčních slov se sbírkami Stanislava Kostky Neumanna a Vladimíra Vaška než mezi sebou. Přítomnost těchto sbírek v grafu narušila dvojice jinak správně shlukovaných sbírek zbylých autorů. Pozornost také budí umístění skupiny básní Petra Bezruče mimo jádro. Vzorky básní Vladimíra Vaška VV a VV mimo jádro se oproti analýze první skupiny funkčních slov nacházejí v centru celé skupiny Slezských písní. Skupina básní PB mimo jádro se tentokrát nachází na periférii shluku Slezských písní. Celkově nám tedy analýza funkčních slov, v případě obou použitých vzorků, nedává jednotný pohled na stylovou podobnost zkoumaných vzorků. Znovu ovšem můžeme konstatovat evidenci toho,

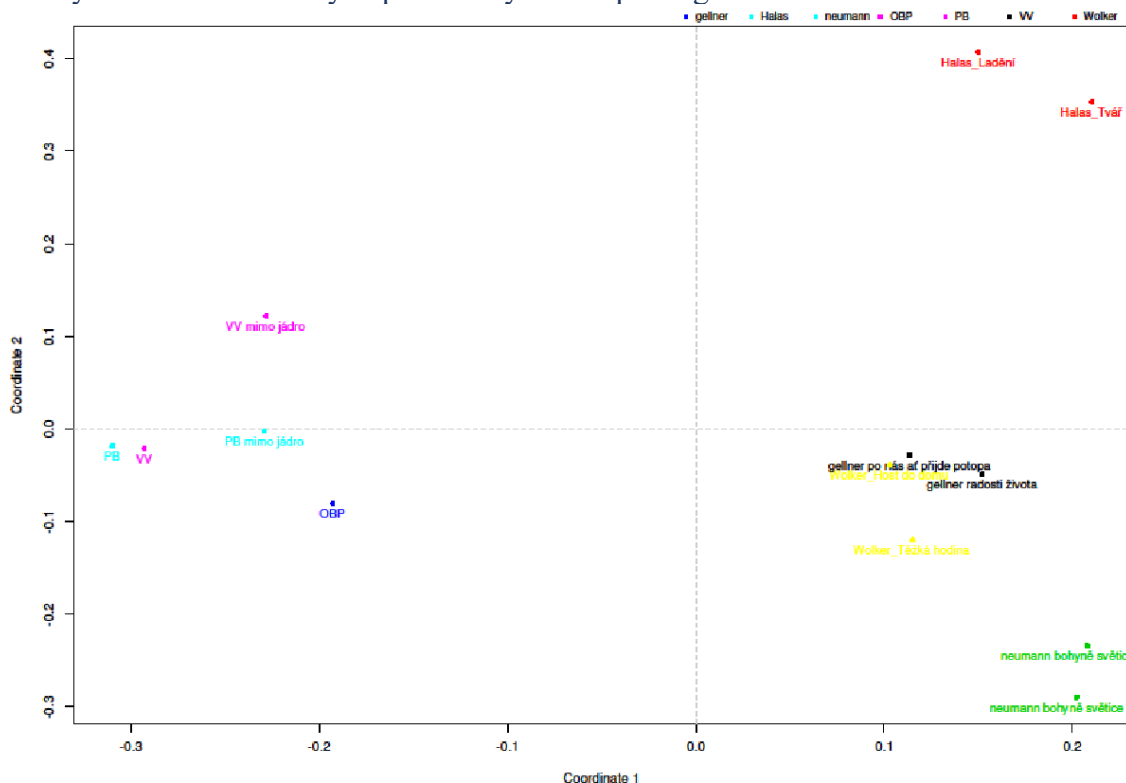
že žádný z výsledků analýzy přesně nenaznačuje nic o stylové odlišnosti vytvořených skupin básní Slezských písní, jak je naznačují hypotézy o sporném autorství sbírky.

Hapax legomena

V následujících analýzách provedeme srovnání vzorků na základě zastoupení slov vyskytujících se v nich pouze jednou – hapax legomen. Sbírký budou porovnány podobně jako v případě analýzy celého slovníku na základě využití kosinové nepodobnosti a zobrazení metrického vícerozměrného škálování nebo hierarchického shlukování. Stejně jako v předešlé kapitole je nutné nejdřív komentovat, proč budeme ke srovnání používat zvolený materiál.

Pod pojmem hapax legomena rozumíme slova jedinečná, tedy taková, která se v textu objevují jen jednou. Po analýze na základě funkčních slov tak nabízí hapaxy opačný úhel pohledu na možnosti srovnávání textů. Zatímco vysoce frekventovaná plnovýznamová slova (content words) v sobě nesou informace o tématu textu, Faltýnek a Matlach (2020) ukazují, že hapaxy mají společné vlastnosti jako funkční slova (viz předchozí podkapitola) v tom smyslu, že jsou do značné míry nezávislé na tématu textu, a to vzhledem k autorské volbě synonymních prostředků, kterými mohou být jednotlivé hapaxy a fráze, v nichž se vyskytují, zastoupeny (podrobněji k výskytu autorských hapaxů v určitých frázích viz Faltýnek 2020). Předpokládáme také (viz Baayen a kol. 1996), že autoři hapax legomena volí v textu do značné míry nevědomě, že jejich volba v souboru synonymních jazykových prostředků není kontrolovaná, nebo není kontrolovaná v celém rozsahu slov jedinečných vyskytujících se v textu. Kromě postavení na opačné straně frekvenčního spektra mají hapax legomena ještě jeden podstatný rozdíl oproti funkčním slovům. Funkčních slov je k atribuci používáno několik desítek (viz Mosteller – Wallace, 1964), případně stovek (Juola – Baayen 2003), hapax legomena představují komplex desetitisíců samostatných vlastností (v našem případě), na základě kterých autory porovnáваме. Jak již bylo zmíněno výše, jedná se o slova v textu přítomná bez vědomé kontroly autora, tím pádem se jedná o velmi rozsáhlý a jedinečný korpus ke komparaci autorských stylů. Musíme pouze prokázat, že přítomnost v textu nemá náhodnou podobu, danou kontextem výpovědi, ale že se jedná o autorovy preferované jazykové vlastnosti. Prokáže se to v následné analýze.

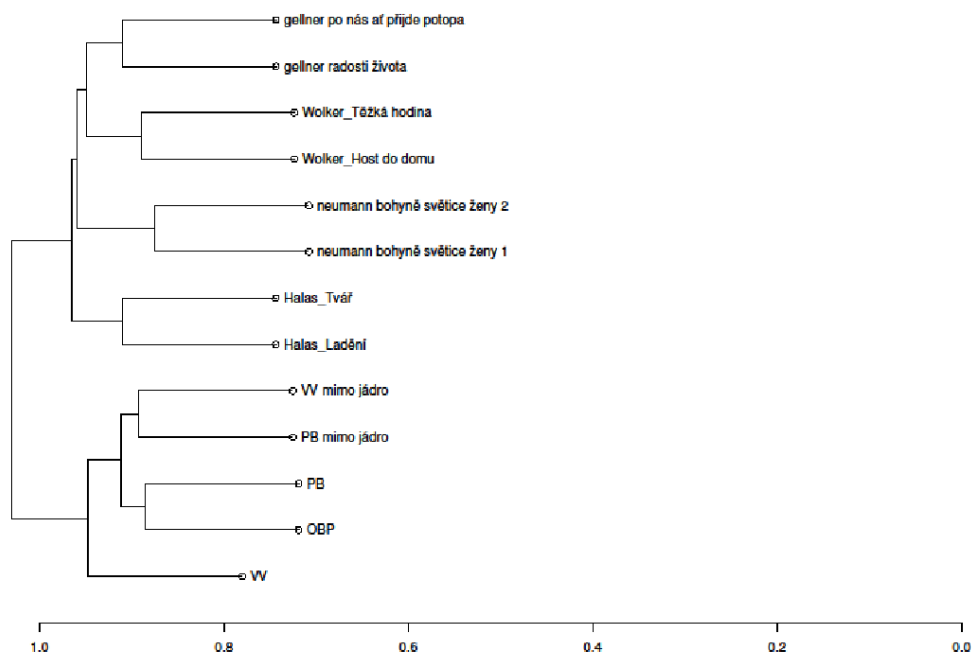
Analýza autorství Slezských písní za využití hapax legomen



Obrázek 7 – Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) za využití hapax legomen a kosinové nepodobnosti, zobrazení za pomoci metrického vícerozměrného škálování.

Analýza autorství za využití slov jedinečných, pouze jednou v textu se vyskytujících tzv. hapax legomen, se ukázala jako velmi úspěšná. Všechny skupiny básní se v grafu shlukují v těsné blízkosti. Autorské klastry jsou zřetelněji odděleny, než tomu bylo v případě analýzy celkové slovní zásoby, mnohem zřetelnější je výsledek ve srovnání s analýzou obou skupin funkčních slov (viz podkapitoly výš, Obrázky 1–4). Sbírký jednoho autora vždy tvoří jasně identifikovatelný pár. Můžeme si všimnout dvojice obsahující texty Stanislava Kostky Neumanna v pravém dolním rohu a paralelně pak dvojice děl Františka Halase v pravém horním rohu. Také díla Františka Gellnera a Jiřího Wolkerera se v grafu seskupily. Jednoznačné seskupení se neprojevovalo u dvou skupinách o dvou dílech. Host do domu Jiřího Wolkerera se nachází velmi těsně u skupiny Gellnerových textů. V této souvislosti můžeme říci, že komparační metoda selhala. Skupiny básní patřících do Slezských písní se všechny drží v levé části grafu. Dá se tedy říct, že všechny se od sebe nachází v relativní blízkosti. V tomto směru

je výsledek uspokojivější či jednoznačnější než v případě funkčních slov a celkové slovní zásoby.



Obrázek 8 – Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2) za využití hapax legomen a kosinové nepodobnosti, zobrazení za pomoci hierarchického shlukování.

V Obrázku 8 byly výsledky analýzy hapax legomen zobrazeny pomocí dendrogramu vytvořeného metodou hierarchického shlukování. Tato metoda nám poskytuje přesnější pohled na míru podobnosti jednotlivých textů. Z grafu je patrné, že se díla od jednoho z autorů vždy umístila na společné nejbližší větvi dendrogramu. Zaměříme-li pozornost na skupiny patřící do Slezských písní, můžeme si všimnout, že se jasně vymezila ta část básní, které patří do kanonického jádra sbírky, a skupina, která do jádra dle kritiků nepatří. Obě skupiny ovšem obsahují takové množství společných hapaxů, takže se od sebe více neoddálily. Nejmenší prvek korpusu slezských písní, skupina VV, je dle výsledků analýzy stejně podobná oběma zbylým párům. Pochybnosti o autorském přiřazení Hosta do domu ke Gellnerově skupině se ukázalo jako vedlejší jev použití zobrazovací metody MDS, který na vlastní výsledek analýzy nemá vliv. Celkově proto můžeme analýzu hapax legomen Slezských písní prohlásit za překvapivě úspěšnou, jelikož srovnávané autorské páry jednoznačně seskupuje. Zároveň ukazuje, že

všechny skupiny patřící do sbírky Slezské písně, si jsou dostatečně podobné, aby se dalo usuzovat, že je napsal tentýž autor.

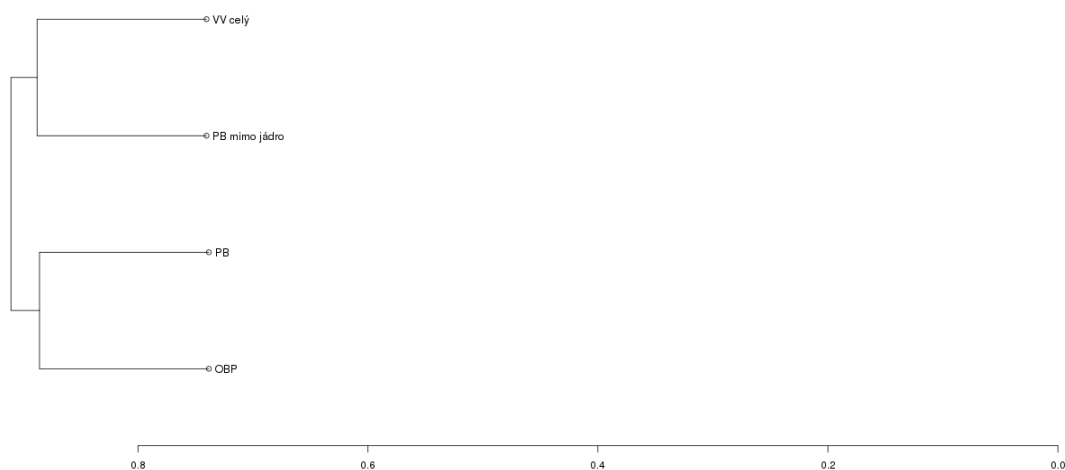
Srovnání grafů – vyhodnocení

Korpus Slezských písní a s nimi srovnávaných básnických sbírek byl podroben několika analýzám, které se běžně používají v oblasti určování autorství, některé metody byly naopak nové a nabízejí na problematiku nový pohled. Konkrétně byly sbírky porovnávány na základě jejich celkového slovníku, funkčních slov a slov, která se v každém díle vyskytují jen jednou – hapax legomen. Překvapivě nejvíc odchylek od autorských skupin se při analýze objevilo při způsobu porovnávání textů na základě slov, jež nenesou žádný obsah, tzn. slov funkčních. Byla provedena analýza dvou sad nejčastějších funkčních slov vyskytujících se ve sbírce Slezské písně, a to vždy po padesáti prvcích. Za pomoci manhattanské metriky aplikované na z-skóre funkčních slov byla následně vyhodnocena podobnost vzorků. Tato metoda má zaručovat srovnání podobnosti textů nezávisle na jejich tématu a stylu. V oboru určování autorství se jedná o velmi častou a populární metodu, její výpočet má také několik různých variant (Evert a kol. 2017). Porovnání oběma sadami funkčních slov byly mimo autorskou skupinu seskupeny texty Františka Gellnera a Františka Halase (viz Obrázek 4 a 6).

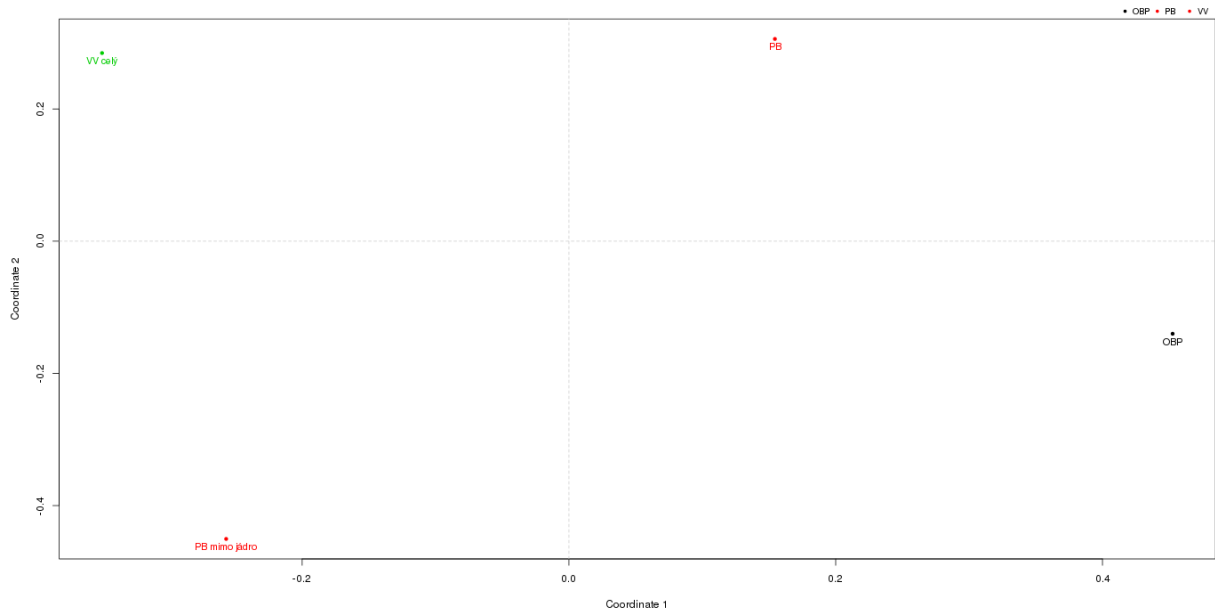
Co se týče analýzy funkčních slov a sbírky Slezské písně, tak se při analýze prvních padesáti nejčastějších funkčních slov vzorky básní výchozí sbírky rozdělily na dvě části. Básně, které byly dle teorie Jana Drozda zahrnuty ve skupině patřící Vladimíru Vaškovi, nemají dle výsledků první analýzy mnoho společných prvků se zbytkem SP, avšak mezi sebou navzájem podobné jsou. Skupiny básní PB a PB mimo jádro vytvořily dvojici, ke které se v grafu blíží skupina OBP. Analýza druhé sady funkčních slov ukázala, že nejvíce podobné jsou dvojice PB a OBD, VV a VV mimo jádro. Skupina PB mimo jádro byla vyhodnocena jako podobná s ostatními vzorky Slezských písní, nicméně ne tolik jako dílo *Po nás ať přijde potopa* od Františka Gellnera, která má dle výsledků analýzy víc společných prvků s básněmi Vladimíra Vaška než s druhým dílem od tohoto autora. Tyto výsledky nepřinášejí konzistentní pohled na atribuci Slezských písní a všeobecně neprokazují spolehlivé vlastnosti atribuční procesu (to nemusí být dáno samotnou metodou funkčních slov, platí to ale pro náš textový materiál).

Metody srovnávání sbírek na základě jejich celkového slovníku a na základě obsahu slov jedinečných (hapaxů) nám překvapivě poskytly přesvědčivější míru spolehlivosti než funkční slova. V obou případech se skupiny básní od jednoho z autorů vždy sdružily bez výjimek. Skupiny Slezských písní taktéž vytvořily paralelní dvojice – jako nejpodobnější byly vyhodnoceny skupiny patřící do tzv. „srdce“ sbírky, tedy skupiny PB a OBD, druhý pár vytvořily básně, které do tohoto jádra nezapadají – VV mimo jádro a PB mimo jádro. Ostatní texty vedené v analýze pod názvem VV, mají dle výsledků obou metod dostatek společných prvků se zbytkem Slezských písní, ale zároveň ne tolik, aby vytvořily s některou ze skupin dvojici. Tato skutečnost může být zapříčiněna velikostí tohoto vzorku, VV obsahuje jen šest básní o celkové velikosti pouhých 871 slov. Následně tedy budou texty analyzovány znovu, tentokrát však budou skupiny VV a VV mimo jádro sjednoceny pod názvem VV celý.

Analýza vzorků Slezských písní po sloučení vzorků VV a VV mimo jádro

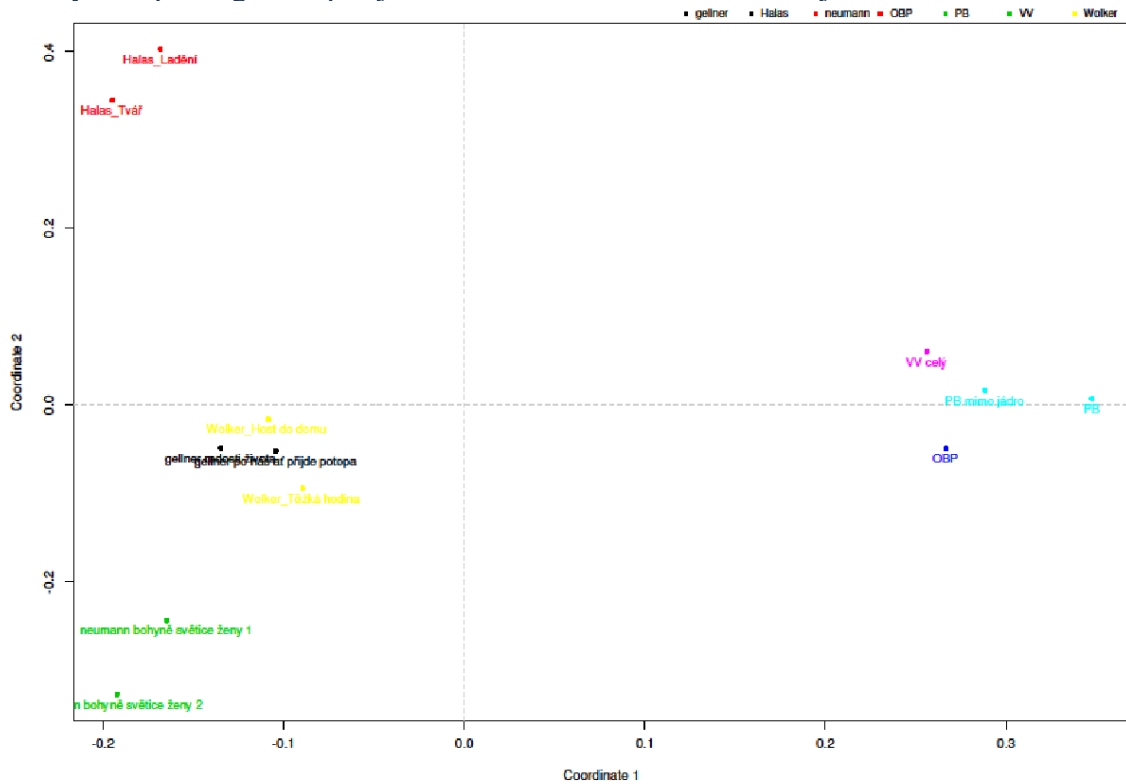


Obrázek 9: Analýza vzorků básní Slezských písní (viz Tabulka 1, vzorek básní VV celý obsahuje sloučené vzorky VV a VV mimo jádro) za užití celkového slovníku, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení za pomoci hierarchického shlukování.

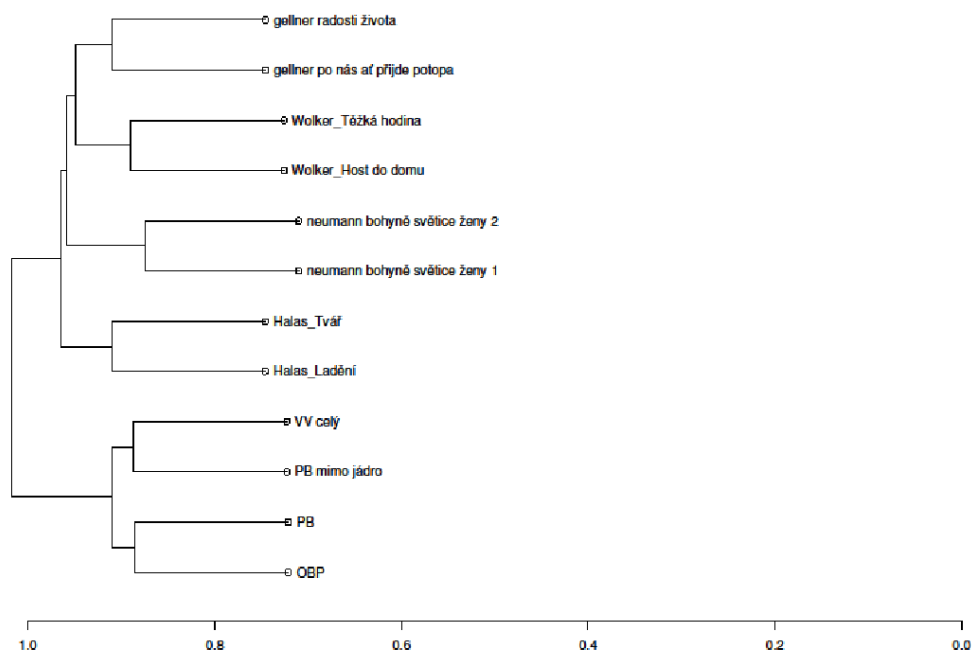


Obrázek 10: Analýza vzorků básní Slezských písní (viz Tabulka 1, vzorek básní VV celý obsahuje sloučené vzorky VV a VV mimo jádro) za užití celkového slovníku, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení za pomoci metrického vícerozměrného škálování.

Analýza hapax legomen po sjednocení vzorků VV a VV mimo jádro



Obrázek 10: Analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2, vzorek básní VV celý obsahuje sloučené vzorky VV a VV mimo jádro z Tabulky 1) za užití hapax legomen, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení za pomoci metrického vícerozměrného škálování.



Obrázek 11: A analýza vzorků básní Slezských písní a dalších básnických sbírek (viz Tabulka 1 a Tabulka 2, vzorek básní VV celý obsahuje sloučené vzorky VV a VV mimo jádro z Tabulky 1) za užití hapax legomen, srovnání bylo provedeno za užití kosinové nepodobnosti, zobrazení za pomoci hierarchického shlukování.

Srovnání výsledků analýz po sjednocení skupiny VV a VV mimo jádro

Sjednocení skupiny VV a VV mimo jádro se ukázalo jako velmi účinné. Vzhledem k předchozím analýzám a častému klastrování vzorku VV na okraji skupiny Slezských písní se zdá, že sloučení vzorků Vladimíra Vaška odstranilo vliv délky na zobrazované podobnosti. O celkovém počtu slov 3532 se nová skupina VV celý svou délkou přibližuje zbylým skupinám patřícím jak do množiny Slezských písní, tak k ostatním porovnávaným sbírkám. Při porovnávání skupin na základě jejich celkového slovníku byla množina Slezských písní přesně a jednoduše rozdělena na dvojici obsahující básně z kanonického jádra sbírky OBP a PB a na dvojici VV celý a PB mimo jádro. Tyto vzorky zároveň prokazují velmi těsnou stylovou blízkost – jejich pozice není příliš ovlivněna srovnávanými vlastnostmi zbylých básnických sbírek zahrnutých do analýzy. Při analýze hapaxů je nová skupina VV celý shlukována v těsné blízkosti s ostatními texty korpusu Slezských písní. Jako při předchozí analýze se vzorky

Slezských písní oddělily na pár tvořící jádro sbírky a pár tvořící zbylé básně. Díla ostatních autorů se znovu seskupila k sobě. Předpoklad, že se skupina VV od ostatních vzdalovala ne kvůli nízké míře podobnosti, ale malému počtu slov, se nám tedy podařilo potvrdit. Především u analýzy hapax legomen, která představují velmi širokou síť vlastností – vzhledem k převažujícím názorům v oblasti určování autorství náhodných vlastností – se těsné shlukování sbírek jeví jako přesvědčující o autorské identitě.

Závěr

Již od svého prvního vydání v roce 1899 upoutává soubor básní známý dnes pod názvem Slezské písně zájem čtenářů. Básně byly nejprve vydávány v časopise Čas a knižní publikace se dočkaly až v roce 1910. Nejednen čtenář pak mohl zaznamenat, že s každým novým vydáním sbírky se rapidně snižovala kvalita básní. Přestože se jedná o poměrně neurčitý popis těchto novějších upravených verzí básní, byl úpadek v kvalitě básní potvrzen několika literárními historiky, viz teoretická část práce. Fakt, že se autor Slezských písní podepisoval pod novější básně jiným jménem, a také jeho odtažitý postoj ke starší tvorbě, kterou mnohokrát opravoval, přepisoval a bez pádného důvodu měnil, vedlo mnohé ke zvažování teorie, že jmenovaný Petr Bezruč není autorem celé sbírky a minimálně část básní vydávaných během přelomu devatenáctého a dvacátého století, těch „nejkvalitnějších“, napsal někdo jiný.

V teoretické části této práce jsou představeny dvě základní a protichůdné teorie vysvětlující drastické stylové změny v Bezručově tvorbě. Jako první je představena teorie Oldřicha Králíka, který se sbírkou zabýval několik desítek let a vydal o ní několik stěžejních děl, které ovlivnily podobu, pod jakou dnes sbírku známe. Základ jeho teorie tvoří přesvědčení, že s přibývajícímí léty se Vaškova tvorba jednoduše postupně měnila. Nový pseudonym dával autorovi svobodu, možnost se odpoutat od slávy předešlého díla a zkusit něco nového. Po roce 1900 je podle Králíka Petr Bezruč, část autorovy osobnosti, mrtev, přestože Vladimír Vašek píše básně ještě dvacet let poté. Všechny teorie o tom, že by byl autorem básní i někdo jiný, Králík nikdy nepovažoval za opodstatněné.

V kontrastu k teorii Oldřicha Králíka je v teoretické části práce představena také teorie Jana Drozda, který ve svém díle Autoři Slezských písní (2003) publikuje názor, že autorů sbírky musí být víc. Argumentuje, že část básní napsal Vaškův zesnulý přítel Ondřej Boleslav Petr, učitel z Frýdku-Místku, jehož osobnost dle svědků a záznamů více odpovídá prvním známým

básním Slezských písní. Tuto teorii podkládá mnohými argumenty. Drozdova publikace rozpoutala v roce 2003 zájem, který se nicméně po pár letech vytratil kvůli absolutnímu nedostatku materiálů prokazující byt' jen některou z těchto hypotéz. Dnes věří, že za Slezskými písněmi stojí víc než jeden autor, jen málokdo. Hlavním zastáncem této hypotézy je i dodnes písničkář Jaromír Nohavica, který se tematikou taktéž zabývá už několik desítek let. Jelikož se sám Nohavica uznává, že se tematikou zabývá jen jako svému koníčku, jeho argumenty nikdy nebyly oficiálně publikovány, a zůstávají ve formě mluvených přednášek pro laickou veřejnost, je obtížné se jimi tedy dopodrobna zabývat, což potvrdil mimo jiné i Zdeněk Smolka. V nejnovějším kritickém vydání sbírky Slezské písně vydaném Ústavem pro českou literaturu Akademie věd z roku 2014 je možnost, že Vašek část básní zdědil, opsal, či dokonce ukradl, jednoznačně zamítnuta.

Zatímco teoretická část práce pouze vyjmenovává postoje jednotlivých textologů k tématu, praktická část představuje konkrétní způsoby určování autorství a analýzy autorství Slezských písní. Celá sbírka je nejprve rozdělena do skupin dle hypotéz o jejich vzniku představených v teoretické části, následně je porovnávána spolu se sbírkami dalších čtyř autorů, kteří stojí v analýzách proti vzorkům básní Slezských písní jako komparační materiál. Předpokládáme, že pokud by měly Slezské písně více autorů, mohl by nám tento komparační vzorek pomoci vyznačit jejich stylovou nejednotnost – „rozdělit“ vzorky na základě odlišných stylových vlastností ze skupiny Slezských písní a přidružit je ke srovnávaným dílům. Tento přístup je sice spekulativní, pokud ale vezmeme v potaz předpoklad disciplíny určování autorství, že autorský styl je souborem nevědomě používaných jazykových prostředků, které si můžeme představit jako konkrétní množinu v textu identifikovatelných jevů, dává nám to dobré důvody k jeho aplikaci. Znamená to totiž, že pokud nejsou jazykové prostředky nekonečné, což v případě slovních forem nejsou, pak při větším rozsahu komparovaného textu v analýze, v našem případě básnických sbírek, se při jeho libovolném výběru odkrývají odlišné vlastnosti textů, u nichž se ptáme na jejich autorský původ.

V práci používáme několik přístupů k analýze autorství. Jako první je provedena analýza na základě podobnosti celé slovní zásoby. Následně je zaveden pojem funkční slova, který je pro rozpoznávání idiolektu autorů pro mnohé vědce klíčový. Skupiny básní jsou následně analyzovány na základě dvojí sady funkčních slov obsažených v celku Slezských písní. Určování autorství prostřednictvím právě těchto slov, která nejsou ovlivněna žánrem, stylem ani tematikou, se v praktické části ale neprokázalo jako spolehlivé, a to navíc ve srovnání se

zbylými dvěma metodami, které používáme. Při analýzách obou sad došlo k chybnému shlukování básnických sbírek známých autorů (Františka Gellnera a Františka Halase). Poslední provedená analýza se týká slov, která se v každém textu vyskytují jen jednou, tedy hapax legomen. Tato slova, byť jsou z hlediska frekvence opakem funkčních slov, mohou sloužit k atribuci textu, a to v podobné míře spolehlivosti jako jiné metody. V případě Slezských písní se tato metoda osvědčila jako nejpřesnější, nevedla k žádnému chybnému shlukování textů známých autorů. Naopak těsněji než ostatní metody shlukovala všechny vzorky Slezských písní.

Cílem této práce nebylo na základě nových výsledků přesně určit, zda byla sbírka Slezské písně napsána jedním autorem, nebo dvěma. K tomu nemáme dostatečnou sbírku materiálů pro každého z možných autorů. Práce si kladla za cíl tento soubor textů analyzovat kvantitativními metodami určování autorství a aplikovat metodu hapax legomen, která v kauze Slezských písní doposud nebyla použita. V souvislosti s tím jsme provedli několik souvisejících analýz užívaných k identifikaci autora textu. Po jejich provedení je nutno konstatovat, že ani jeden z přístupů neidentifikoval odlišnost autorského stylu dle teze Jana Drozda, v současnosti propagované Jaromírem Nohavicou. Námi analyzované vzorky Slezských písní nevykazují dostatečnou míru odlišnosti na to, aby se dalo poukazovat na jejich odlišný původ, nemáme tudíž jediný důvod spekulovat o tom, že autorem díla je jiný člověk než Vladimír Vašek. V případě, že by měli zastánci teorie dvojího autorství pravdu a soubor básní, které patří do jádra sbírky, skutečně patřil Ondřeji Boleslavu Petrovi, a rozdíl mezi idiolektem Petra a Vaška byl tak ohromný a jednoznačný, jak zastánci této teorie tvrdí, pak by se tato skupina analyzovaná pod jménem OBP měla jasně oddělovat od ostatních skupin sbírky. To jsme však ani při jedné z analýz nepozorovali.

V rámci této práce byla provedena jen vybraná skupina analýz sloužících k atribuci autora. Je možné, že dalším studiem a zkoumáním jinými metodami určování autorství by se mohlo podařit téměř stoletou záhadu Slezských písní rozluštit a finálně určit, zda se Vladimír Vašek svým přítelem Petrem inspiroval, nebo skutečně část jeho tvorby převzal. V této práci se nám podařilo pouze prověřit, zda nemohou vybrané velmi řídké se vyskytující prvky textu identifikovat jeho autora. Ukázalo se to být možné, a to ve srovnání s metodami, které jsou prozatím chápány jako standardní.

Literatura a zdroje

Bezruč, P. ([2014]). *Slezské písně*. Ústav pro českou literaturu AV ČR. <http://www.digitalniknihovna.cz/mzk/uuid/uuid:37fff290-9aa9-11e6-9325-005056827e52>

BAAYEN, H., – HALTEREN, H., – TWEEDIE, F. (1996). *Outside the cave of shadows: using syntactic annotation to enhance authorship attribution*. *Literary and Linguistic Computing*, 11(3), 121–132.

BURROWS, J. (2003). *Questions of Authorships: Attribution and Beyond*. *Computers and the Humanities*, 37(1), 5–32.

Drozd, J. (2003). *Autoři slezských písní*. Tilia. <http://www.digitalniknihovna.cz/mzk/uuid/uuid:1c34c8c0-341f-11e5-8851-005056827e51>

DROZD, Jan. (2005) *Otevřený dopis vědeckým literárním institucím, aneb, Autoři Slezských písní*. V Šenově u Ostravy: Tilia.

Evert, S. – Proisl, T. – Jannidis, F. – Reger, I. – Pielström, S. – Schöch, C. – Vitt, T. (2017): *Understanding and explaining Delta measures for authorship attribution*. *Digital Scholarship in the Humanities*, Volume 32, Issue suppl_2, December 2017, ii4–ii16, <https://doi.org/10.1093/llc/fqx023>.

FALTÝNEK D. (2020): *Celkom iste sa prúde na to, že niektoré slová sa opakujú doslovne: Horeckého hypersyntax*. *Jazykovedný časopis*, 71, 2, 185–196.

FALTÝNEK, D. – MATLACH, V. (2020): *Hapax remains: Authorial features of textual cohesion in authorship attribution*. DOI: 10.13140/RG.2.2.19865.24164

Pět studií o Petru Bezručovi. (1947). Universitní fond. <http://www.digitalniknihovna.cz/mzk/uuid/uuid:8e782000-a382-11e3-9d7d-005056827e51>

Bezruč, P., Králík, O., Ficek, V., & Pallas, L. (Eds.). (1967). *Slezské písně Petra Bezruče: historický vývoj textu*. Profil.

JIROUŠEK, Martin (2014): Nohavica opět na stopě Slezských písní. Bezruč není jediný autor, řekl studentům, Ostravan.cz [cit. 2021-03-15]. Dostupné z <https://www.ostravan.cz/16868/nohavica-opet-na-stope-slezskych-pisni-bezruc-neni-jediny-autor-rekl-studentum/>

JUOLA, P. (2008). Authorship attribution. *Foundations and Trends in Information Retrieval*, 1(3).

JUOLA, P. and BAAYEN, H. (2003), A controlled-corpus experiment in authorship identification by cross-entropy. *Literary and Linguistic Computing*, 20, 59–67.

KRÁLÍK, Oldřich a Viktor FICEK. (1978) *Kapitoly o Petru Bezručovi*. Ostrava: Profil., Dostupné také z: <http://www.digitalniknihovna.cz/mzk/uuid/uuid:28430c50-8240-11e7-b92d-005056827e51>

KRÁLÍK, Oldřich. (1963) *Text Slezských písní*. Ostrava: Krajské nakladatelství, Knihovna Památníku Petra Bezruče.

BEZRUČ, Petr a Vladimír MACURA (1980): *Jen jedenkrát: zázilky Času 1899-1914*. Praha: Československý spisovatel, Klub přátel poezie. Základní řada. Dostupné také z: <http://www.digitalniknihovna.cz/mzk/uuid/uuid:5709c24a-ab96-4052-a33d-d3c6e1d45584>

Kestemont, Mike. (2014). Function Words in Authorship Attribution. From Black Magic to Theory?

Mosteller, F., and Wallace, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Reading, MA: Addison Wesley Publishing Company.

MOTÝL, Ivan (2014): O pravdě a lži, in Týden, cit. 2021-03-14. Dostupné z http://www.nohavica.cz/cz/jn/rozhovory/rozh_cz_62.htm

Tesar R., Poesio M., Strnad V., Jezek K.: "Extending the Single Words-Based Document Model: A Comparison of Bigrams and 2 – Itemsets". In Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng'06), Amsterdam, Netherlands, ACM press, ISBN 1-59593-515-0, pages 138-146, September 2006. (<http://doi.acm.org/10.1145/1166160.1166197>)

Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, 17, 401–419. <https://doi.org/10.1007/BF02288916>

VETEŠKOVÁ, Michaela (2020) Kdo napsal Slezské písně? Rozhovor. In: rozhlas. Rádio. 28. ledna 2020, [cit. 2021-03-15] Dostupné také z: <http://irozhl.as/8SP>

SMOLKA, Zdeněk (2008): Jak jsme se hádali o autorství Slezských písní, in Host. Brno: Spolek přátel vydávání časopisu HOST, Roč. XXIV., Č. 3.

SMOLKA, Zdeněk in Petr Bezruč – bard první, co promluvil. (2018) Ostrava: Ostravská univerzita, ISBN 978-80-7599-004-4

PEŠAT, Zdeněk (1964): CO S BEZRUČOVÝMI SLEZSKÝMI PÍSNĚMI? *Česká Literatura*, vol. 12, č. 1, 59–63., Dostupné také z: www.jstor.org/stable/43332574.

PLECHAC, Petr & BOBENHAUSEN, Klemens & HAMMERICH, Benjamin. (2019). Versification and authorship attribution. Pilot study on Czech, German, Spanish, and English poetry. *Studia Metrica et Poetica*.

Evert, S., Jannidis, F., Proisl, T., Vitt, T., Schöch, C., Pielström, S., Reger, I. (2016). Outliers or Key Profiles? Understanding Distance Measures for Authorship Attribution. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 188-191.

Urbanec J. (1969). Mladá léta Petra Bezruče: život a dílo 1867-1903. Ostrava: Profil (Knihovna Památníku Petra Bezruče v Opavě).

DUPAL, Pavel. Statistické metody ve stylometrii, Praha (2017). Diplomová práce. Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky, Katedra matematiky. Vedoucí diplomové práce Nikola Kaspříková.

KESTEMONT, M. (2014). Function words in authorship attribution. from black magic to theory? In Proceedings of the 3rd workshop on computational linguistics for literature (clfl)

(pp. 59–66)

Seznam obrázků

Obrázek 1: Analýza slovníku vybraných básnických sbírek pomocí modelu Bag-of-words a kosinové nepodobnosti, zobrazené pomocí hierarchického shlukování.

Obrázek 2: Analýza vybraných básnických sbírek na základě frekvence 50-100 nejfrekventovanějších funkčních slov za pomoci manhattanské metriky, zobrazeno metrickým vícerozměrným škálováním.

Obrázek 3: Analýza vybraných básnických sbírek na základě frekvence 50-100 nejčastějších funkčních slov za pomoci manhattanské metriky, zobrazeno hierarchickým shlukováním.

Obrázek 4 - Analýza hapax legomen za užití kosinové nepodobnosti zobrazené pomocí vícerozměrného škálování.

Obrázek 5 - Analýza hapax legomen za užití kosinové nepodobnosti zobrazené pomocí hierarchického shlukování.

Obrázek 6: Analýza slovníku sbírky SP metodou bag-of-words za užití kosinové nepodobnosti zobrazená pomocí hierarchického shlukování.

Obrázek 7: Analýza slovníku sbírky SP metodou bag-of-words za užití kosinové nepodobnosti zobrazená pomocí hierarchického shlukování.

Obrázek 8: Analýza hapax legomen vybraných skupin básní a sjednocené skupiny VV celý za užití kosinové nepodobnosti zobrazené pomocí metody metrické vícerozměrné škálování.

Obrázek 9: Analýza hapax legomen vybraných skupin básní a sjednocené skupiny VV celý za užití kosinové nepodobnosti zobrazené pomocí hierarchického shlukování

Seznam příloh

Příloha 1: Porovnávání skupiny sbírky Slezské písně (OBP, PB, PB mimo jádro, VV, VV mimo jádro)

Příloha 2: Ostatní porovnávané sbírky (F.Gellner – Po nás ať přijde potopa, Radosti života; F.Halas – Ladění, Tvář; S.K.Neumann – Bohyně, světice, ženy 1-2; J.Wolker – Host do domu, Těžká hodina)

Seznam tabulek

Tabulka 1: Seznam básní SP v jednotlivých skupinách
Tabulka 2: Seznam autorů a děl
Tabulka 3: Seznam 1-50 nejfrekventovanějších funkčních slov ve sbírce Slezské písně.
Tabulka 4: Seznam 51-100 nejfrekventovanějších funkčních slov ve sbírce Slezské písně.