

Leksička gustoća i raznolikost u Šimićevu pjesništvu*

IRINA BUDIMIR

Sveučilište u Mostaru, Filozofski fakultet

University of Mostar, Faculty of Humanities and Social Sciences

E-mail: irina.budimir@ff.sum.ba

DOI: <https://doi.org/10.47960/2712-1844.2026.12.333>

UDK: 811.163.42'373

821.163.42 A.B.Šimić: 811.163.42'373

Izvorni znanstveni rad

Primljeno: 21. travnja 2026.

Prihvaćeno: 29. lipnja 2026.

SANDRA MALETIĆ

PDS Mediji i komunikacija (doktorandica), Sveučilište Sjever, Koprivnica

Postgraduate Doctoral Study in Media and Communication (PhD student), University North

E-mail: samaletic@unin.hr

IVAN MARKIĆ

Sveučilište u Mostaru, Fakultet strojarstva, računarstva i elektrotehnike

University of Mostar, Faculty of Mechanical Engineering, Computing and

Electrical Engineering

E-mail: ivan.markic@fsre.sum.ba

Sažetak

Na temelju dviju mjera leksičkoga bogatstva, leksičke gustoće i raznolikosti, cilj je rada ustanoviti različitosti leksičke strukture pojedinih faza poezije Antuna Branka Šimića, s obzirom na to da se u literaturi navode različite podjele njegova pjesničkoga stvaralaštva, od dviju do četiriju faza. Analizirani korpus čine pjesme iz njegove impresionističke faze (1913.-1917.) te pjesme iz ekspresionističke zbirke *Preobraženja* (1920.) i ciklusa *Pjesme o tijelu* (1920.-1925.).

Vodeći se Šimićevom promjenom stila i pjesničkog izričaja u navedenim fazama i primjenjujući mjere leksičkoga bogatstva (gustoća i raznolikost), dobiveni rezultati pokazuju leksičke i druge različitosti koje idu u prilog tezi o mogućim trima fazama. Dobiveni rezultati usporedili su se s postojećim rezultatima tih dviju mjera drugih hrvatskih pjesnika. Istraživanje pokazuje

* Rad je pod naslovom "Mjere rječničkoga bogatstva u različitim fazama nastanka Šimićeve poezije" izlagan na međunarodnoj znanstvenoj konferenciji *Percepcija smrti u novome mileniju: Od pjesničke apstrakcije do medijske atrakcije*, održanoj 6. i 7. studenoga 2025. na Filozofskom fakultetu Sveučilišta u Mostaru, a povodom 100. obljetnice smrti Antuna Branka Šimića.

da analiza leksičke gustoće i raznolikosti donosi nova saznanja o leksičkim obilježjima među promatranim skupinama Šimićevih pjesama, koje se prema stupnju leksičke složenosti međusobno razlikuju, a jasno se razlikuju i od rezultata mjera leksičkoga bogatstva drugih uspoređivanih pjesnika.

Ključne riječi: A. B. Šimić; poezija; leksička raznolikost; leksička gustoća.

Lexical Density and Diversity in the Poetry of Antun Branko Šimić

Original scientific article

Received: 21 April 2026

Accepted: 29 June 2026

Summary

The works of Antun Branko Šimić represent an inexhaustible source of numerous scholarly studies dealing, among other things, with language and style, the analysis of syntactic stylistic figures, lexis, phraseology, and similar topics. The corpus of Šimić's poetry was computationally processed more than a decade ago, resulting in parameters that demonstrate the lexical sophistication of his poetic language. Concordances, an alphabetical dictionary with frequencies, as well as frequency and reverse dictionaries, were compiled. In addition to these measures, the informativeness and content richness or poverty of a text can also be measured through lexical density and lexical diversity.

Computer-based quantitative methods (such as MATTR, MTLD, TTR, etc.) are used to calculate these parameters, depending on the length of the text. The richer a text is, the more demanding it becomes for reading. The quality of lexical analysis of written texts directly affects the accuracy of assessing these two measures. Owing to them, it is possible to calculate the reading difficulty of certain texts. They may reveal, for example, how demanding Šimić's poetry is for readers and why some poems are more difficult to understand than others.

Based on two measures of lexical richness—lexical density and lexical diversity—the aim of this paper is to determine the differences in the lexical structure of particular phases of Šimić's poetry, considering that the existing literature has proposed different classifications of poetic development, ranging from two to four phases.

The analysed corpus consists of poems from his Impressionist phase (1913–1917), as well as poems from the Expressionist collection *Preobraženja* (*Transformations*, 1920) and the cycle *Pjesme o tijelu* (*Poems about the Body*, 1920–1925).

Guided by Šimić's stylistic and poetic transformations across these phases, and applying measures of lexical richness (density and diversity), the obtained results indicate lexical and other differences that support the thesis of three possible phases in his poetic development. The results were compared with existing results of these two measures in the works of other Croatian poets. The research demonstrates that analysis of lexical density and diversity provides new insights into the lexical features of the observed groups of Šimić's poems, which differ from one another according to the degree of lexical complexity, while also clearly differing from the results obtained for the lexical richness of the other poets included in the comparison.

Keywords: A. B. Šimić; poetry; lexical diversity; lexical density.

1. Uvod

Književni teoretičari koji su se bavili fazama stvaralaštva A. B. Šimića navode da je najprije stvarao impresionističke pjesme (1913.-1917.). Tu njegovu fazu čine "formalno besprijeorne pjesme u duhu Matoševa artizma, s naglaskom na strogu formu, vezani stih i bogatu, čistu rimu."¹ Drugu fazu predstavljaju njegove ekspresionističke pjesme. Impresionističku fazu (1913.-1917.) čine "formalno besprijeorne pjesme u duhu Matoševa artizma, s naglaskom na strogu formu, vezani stih i bogatu, čistu rimu."² Krmpotić³ navodi da je Šimić uza svoju početničku poeziju pisao "opet po uzoru na Matoša, feljtonističku prozu i poneku kritiku. To njegovo šegrtovanje u književnosti koje možemo smatrati za prvu fazu razvoja, trajalo je otprilike tričetiri godine". I sam Šimić kaže: "od 13 do 16, 17 svoje godine bio sam Matoševac."⁴

U svojim pjesmama pisanim slobodnim stilom, nastalim uglavnom u trogodišnjemu razdoblju, od 1917. do 1920. godine, Šimić "upotrebljava tzv. nepoetski vokabular, traga za ekspresivnim ritmom, uvodi

¹ KREŠIMIR BAGIĆ, "'Drama svijeta' kao 'drama jezika'", u: VLADO PANDŽIĆ (ur.), *Zbornik radova s Međunarodnoga skupa o hrvatskom književniku Antunu Branku Šimiću*, Matica hrvatska, Drinovci, 2008., str. 361-374.

² *Isto*, str. 362.

³ ANTUN BRANKO ŠIMIĆ, *Pjesme i eseji*, (priredila i pogovor napisala Vesna Krmpotić), Matica hrvatska, Zagreb, 1964., str. 157.

⁴ *Isto*.

privatni pjesnički pravopis, osamostaljuje ključne riječi u stihove."⁵ Njegovu želju za novim stilom, pristupom i izričajem Krmpotić dokazuje pjesnikovim riječima: "pred kraj 1916-ste počeo sam se upoznavati s pokretom takozvanog ekspresionizma iz *Sturma* i prestao sam naglo gledati u Matošu svoj ugled. Uzalud bih htio danas – pa i uz najbolje pamćenje – opisati taj prijelaz, zapravo taj prekid: jer upoznavanje s ekspresionizmom izazvalo je u meni najprije negaciju, negaciju svega onoga što sam do tada mislio o umjetnosti, smatrao dobrom u umjetnosti."⁶ Smatra se da je zbirka pjesama *Preobraženja*: "antiakademska, lišena sentimentalne retorike, krajnje jednostavnih pjesama – tadašnja kritika tu zbirku smatra nepoezijom, točnije, izlomljenom prozom."⁷

Ciklus *Pjesme o tijelu* (1920.-1925.) predstavljaju Šimićeve "najzrelije i najgušće pjesme (...) pjesme neoklasicističke faze karakterizira hladan, analitički pristup doživljaju (...) stihovi koji su nastali poslije 1920-te smireni su i blagi. U toj fazi Šimić se nije služio likovnim elementom u pjesmama."⁸

Šimićevo pjesništvo samo u dvanaest godina prolazi iznimno intenzivne promjene, kakve se rijetko događaju i u puno duljim razdobljima razvoja hrvatske poezije.⁹ K. Bagić navodi da je u Šimićevo opusu moguće razlikovati četiri mikropoetike i stilske faze – od impresionističkih početaka pod utjecajem Matoša do inovativnih lirskih koncepata i kasnijih jasno diferenciranih pjesničkih razdoblja. Međutim, ne navodi izričito razliku između treće i četvrte: "Dalo bi se, u pomnijoj analizi, pokazati da su pjesme koje je pisao između 1917. i 1919. stilski, kompozicijski i motivski različite od pjesama u zbirci 'Preobraženja', a ove pak da su različite od pjesama koje su nastale nakon nje (1921-1925)."¹⁰

⁵ K. BAGIĆ, *nav. dj.*, str. 363.

⁶ A. B. ŠIMIĆ, *nav. dj.*, str. 158.

⁷ *Isto*, str. 159.

⁸ A. B. ŠIMIĆ, *nav. dj.*, str. 161.

⁹ K. BAGIĆ, *nav. dj.*, str. 364.

¹⁰ *Isto*.

Dakle, teoretičari književnosti ne slažu se u podjeli Šimićeve poezije na faze, odnosno o broju tih faza, raspon je od dvije (impresionistička i ekspresionistička) ili pak četiri mikropoetike i četiri relativno osamostaljiva pjesnička stila. Stoga ovaj rad istražuje i moguću periodizaciju Šimićeva pjesništva uspoređivanjem mjera leksičke gustoće i leksičke raznolikosti. Čini to analizirajući različitim računalnim jezičnim alatima dvije neistražene mjere leksičkoga bogatstva koje pokazuju uporabu samoznačnih riječi ili samoznačnica te broj različitih riječi ili natuknica u odnosu na ukupan broj riječi u Šimićevu stvaranju. Naime, razvoj korpusne lingvistike posljednjih desetljeća i izrada nacionalnih jezičnih korpusa omogućio je primjenu korpusnih, tj. računalnih metoda za proučavanja jednoga jezika. Oni su vrijedan resurs za proučavanje jezične strukture, semantike i uporabe u različitim kontekstima.

Osim općega jezičnog materijala, nacionalni korpusi mogu sadržavati i specijalizirane podkorpuse posvećene određenim jezičnim žanrovima ili stilovima, poput poezije. U kontekstu pjesništva računalnom obradom mogu se proučavati i analizirati pjesme različitih autora, leksička, frazeološka, stilska obilježja, jezični i autorski obrasci, ritam i metrika, žanr i sl. Računalnim jezičnim alatima kojim se sastavljaju korpusi i različiti podkorpusi jednoga jezika, mogu se analizirati i jezični mikrosustavi, poput pjesničkog opusa. Zahvaljujući korpusnim alatima, npr. određivanju anotacije, mogu se izračunati morfosintaktičke oznake riječi čiji omjeri i odnosi pokazuju neke od mjera leksičkoga bogatstva kakva teksta. Neke su vrste riječi znatno obavjesnije i povezanije sa stvarnošću te tako objektivno pokazuju raznolikost i obavijesnost uporabnoga rječnika.¹¹

Bogatstvo leksika može se utvrditi raznim mjerama, a u literaturi se najčešće ističu leksička istančanost, leksička gustoća, leksička raznolikosti i brojanje pogrešaka.¹² U ovom radu koristit će se samo mjere leksičke gustoće i raznolikosti. Korpus ovog istraživanja temeljio se

¹¹ VICTORIA JOHANSSON, "Lexical diversity and lexical density in speech and writing: a developmental perspective", u: *Lund University, Dept. of Linguistics and Phonetics Working Papers*, 53, Lund, 2008., str. 65-66; <https://journals.lub.lu.se/LWPL/article/view/2273> (14. 1. 2026.).

¹² JOHN READ, *Assessing vocabulary*, Cambridge University Press, Cambridge, 2000., str. 202.

na uzorku Šimićeva pjesništva, određenim prema fazama nastanka pojedinih pjesama. Za potrebe ovoga rada određena su tri podkorpusa: prvi su 42 pjesme njegove impresionističke faze, drugi pjesnička zbirka *Preobraženja* (1920.) s 48 pjesama i treći čini ciklus od 38 pjesama *Pjesme o tijelu*.

Cilj je rada leksičkom raznolikošću (LR) i leksičkom gustoćom (LG) s obzirom na vrijeme nastajanja Šimićeve poezije ispitati koliko su dobiveni kvantitativni podatci relevantni za opis njegova pjesničkoga razvoja. Prema navedenim mjerama u radu će se odgovoriti na sljedeća istraživačka pitanja:

1. Može li se prema leksičkoj gustoći i leksičkoj raznolikosti odrediti broj različitih faza u Šimićevu pjesništvu?
2. Kakva je leksička složenost Šimićeva pjesništva u odnosu na slična istraživanja hrvatskih pjesnika, uz ograničenja koja proizlaze iz metodologije pjesničkih korpusa drugih pjesnika?
3. Razlikuju li se interpunkcijska obilježja u fazama Šimićeva pjesništva s obzirom na leksičku gustoću i leksičku raznolikost?

2. Leksičke jedinice u kvantitativnoj analizi teksta: pojava, natuknica, obličnica i različnica

Prema njihovoj pojavljivanju u tekstu razlikuju se četiri različita vida riječi potrebna za izračun različitih mjera. Svako pojavljivanje riječi shvaćene kao niz slova između dviju bjelina naziva se *pojava* (eng. *tokens*). *Natuknica*, koja se naziva i lema, jedinstven je izraz riječi kako se pojavljuje u rječnicima, njezin citatni oblik. Tako Z. Jelaska i V. Baričević pokazuju da je u rečenici *Izišao sam od Oca i odlazim Ocu* ukupno 7 pojava i 6 natuknica (*izići, biti, od, otac, i, odlaziti*), pri čemu natuknica *Otac* ima 2 pojavnice.¹³

Osim pojavnice i natuknice postoje još dva pojma. *Obličnice* su različiti oblici jedne riječi ili natuknice. Primjerice, hrvatske imenice, a tako i druge imenske riječi, imaju sedam padeža u jednini i množini, što znači da imaju ukupno četrnaest obličnica. Pri tomu se tek neke

¹³ ZRINKA JELASKA – VALENTINA BARIČEVIĆ, "Leksička jednostavnost i značenjska složenost rječnika Ivanova evanđelja", u: *Lahor: časopis za hrvatski kao materinski, drugi i strani jezik*, 13, Zagreb, 2012., str. 102-137.

obličnice izrazom razlikuju od drugih, npr. *mačka* (N), *mačku* (A), *pas* (N), *psom* (I), *nit* (N), *nit* (G). Druge su obličnice izrazom jednake, npr. *mačkama*, *psima* (D, L, I mn.), *psa* (G, A).

Različnice su riječi različita izraza, što uključuje i različite izraze jedne riječi, tj. natuknice. Stoga svaka promjenjiva riječ redovito ima više obličnica nego različenica jer ni jedna natuknica nema toliko različitih izraza koliko ima oblika, odnosno obličnica. Tako natuknica *mačka* ima osam različenica (*mačka*, *mačke*, *mački*, *mačku*, *mačko*, *mačkom*, *mačaka*, *mačkama*), *pas* sedam (*pas*, *psa*, *psu*, *psom*, *psi*, *pasa*, *psima*), a *nit* samo tri (*nit*, *nit*, *nitima*). U spomenutoj rečenici natuknica *Otac* ima dvije različnice: *Oca* i *Ocu*. Različnice se računalno jednostavno izdvajaju, no obličnice je potrebno ručno razlikovati kada im je izraz jednak pripisujući svakoj njezin gramatički sadržaj, npr. genitiv jednine *vrapca* i akuzativ jednine *vrapca*. Istoizraznice ili homonimi u tekstovima predstavljaju jednu različnicu, npr. prijedlog *kod* od nominativa i akuzativa imenice *kod* pa ih je potrebno ručno razdvajati.

Razlikovanje različitih riječi i pojava u različitim rečenicama nije jednako zahtjevno. Primjerice, u engleskome je jeziku jednostavnije nego u hrvatskom zato što u njemu mnoge vrste riječi nemaju tako velik broj oblika, odnosno obličnica, kao u hrvatskom,¹⁴ npr. *cat* (jd.), *cats* (mn.). Vrste riječi s više obličnica morfološki su složenije, pa je hrvatski složeniji od engleskoga. Vrste riječi s više različenica (oblici iste riječi različita izraza) i fonološki su složenije, pa je hrvatski zbog svoje sklonidbe znatno složeniji od engleskoga.

Samoznačnice su riječi koje nose značenje, tj. riječi koje uspostavljaju izravni sadržajni odnos s izvanjezičnom stvarnošću.¹⁵ Samoznačnice (eng. *lexical words*) čine imenice, glagoli, pridjevi, prilozi, uzvici i brojevi (kada se smatraju posebnom vrstom riječi, iako su zapravo samo značenjska skupina koja okuplja različite vrste riječi).

¹⁴ ZRINKA JELASKA i dr., *Hrvatski kao drugi i strani jezik*, Hrvatska sveučilišna naknada, Zagreb, 2005., str. 142.

¹⁵ ZRINKA JELASKA, "Dvojčane podjele vrsta riječi: samoznačnost i promjenljivost", u: LADA BADURINA – VINE MIHALJEVIĆ (ur.), *Jezična skladanja. Zbornik o šezdesetogodišnjici prof. dr. Ive Pranjkovića*, Udruga đaka Franjevačke klasične gimnazije – Visoko u Zagrebu, Zagreb, 2010., str. 108.

Suznačnice poglavito služe za uspostavljanje sintaktičkih i semantičkih odnosa među samoznačnicama, a to su prijedlozi, veznici, zamjenice, pomoćni glagoli, članovi.¹⁶

Analiza odnosa samoznačnica i suznačnica važna je za određivanje leksičke gustoće. Odnos natuknica i pojavaonica i njihovo tumačenje važno je za leksičku raznolikost.

2.1. O mjerama leksičkoga bogatstva

Leksička gustoća (eng. *lexical density*) označuje odnos između samoznačnica i ukupna broja riječi, odnosno pojavaonica u nekome tekstu. Izračunava se kao omjer samoznačnih riječi u odnosu na ukupan broj riječi. Izražava se vrijednostima u rasponu od 0 do 1. Tekstovi s velikim brojem samoznačnica imaju visoku, a tekstovi s manjim brojem samoznačnica nisku leksičku gustoću. Što je leksička gustoća nekoga teksta veća (bliža vrijednosti 1), tekst ima veći broj samoznačnica te je za njegovo razumijevanje potreban i bogatiji rječnik.¹⁷ Naime, taj odnos pokazuje koliko je tekst "bogat značenjem", tj. "gustoća s kojom se informacije prenose".¹⁸ Dakle, leksička je gustoća pokazatelj obavijesnosti teksta, a prosječni broj pojavaonica pokazatelj je duljine teksta.¹⁹

Leksička raznolikost (eng. *lexical diversity*) označuje raznolikost vokabulara, tj. koliko različitih riječi (natuknica) autor upotrebljava u odnosu na ukupan broj riječi (pojavaonica). Leksička raznolikost računa se omjerom natuknica i pojavaonica. Omjer natuknica i pojavaonica pokazuje koliko različitih riječi ima neki tekst, ili drugim riječima, koliko je čiji rječnik raznolik. "Što je mjera leksičke raznolikosti

¹⁶ Isto, str. 109.

¹⁷ ŽELJANA RADIĆ – JELENA KUVAC KRALJEVIĆ – MELITA KOVAČEVIĆ, "Udžbenik kao poticaj ili prepreka leksičkomu razvoju", u: *Lahor: časopis za hrvatski kao materinski, drugi i strani jezik*, 9, Zagreb, 2010., str. 43-59.

¹⁸ MICHAEL ALEXANDER KIRKWOOD HALLIDAY, *An Introduction to Functional Grammar*, Edward Arnold, London, 1985., str. 62; MARIJANA ČEŠI, *Ovladanost standardnim hrvatskim jezikom na kraju obveznoga obrazovanja*, doktorski rad, Filozofski fakultet, Zagreb, 2015., str. 87.

¹⁹ M. ČEŠI, *nav. dj.*, str. 88.

manja (bliža vrijednosti 0), u tekstu je prisutan manji broj različitih riječi, a većina se riječi ponavlja."²⁰

Ako je broj natuknica velik i ako je raznovrstan, onda je vokabular bogat. Dakle, odnosom natuknica i pojava (eng. *Type-Token Ratio* ili TTR) mjeri se visoka ili niska leksička raznolikost.²¹ Budući da se vrijednost TTR-a sustavno smanjuje s povećanjem duljine teksta, njegova uporaba kao jedine mjere pri usporedbi tekstova različitih duljina nije preporučljiva.²² Zato su u ovom radu odabrane mjere leksičke raznolikosti koje su se pokazale postojanijima i manje osjetljivima na duljinu teksta. To su pomični prosječni omjer natuknica i pojava (eng. *Moving Average Type-Token Ratio*, MATTR) i mjera tekstualne leksičke raznolikosti (eng. *Measure of Textual Lexical Diversity*, MTLD).²³

Iako je leksička raznolikost jedan od pokazatelja jezične kompetencije u pisanju, leksičko-stilske varijacije među tekstovima, a u novijim istraživanjima i kreativnosti pisanja, ne treba ju tumačiti kao samostalan pokazatelj.²⁴ Leksički bogatiji tekst, zahtijeva poznavanja

²⁰ Ž. RADIĆ – J. KUVAČ KRALJEVIĆ – M. KOVAČEVIĆ, "Udžbenik kao poticaj ili prepreka leksičkomu razvoju", str. 48.

²¹ J. READ, *Assessing vocabulary*, str. 203; Z. JELASKA – V. BARIČEVIĆ, "Leksička jednostavnost i značenjska složenost rječnika Ivanova evanđelja", str. 116.

²² PHILIP M. MCCARTHY – SCOT JARVIS, "MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment", u: *Behavior Research Methods*, 42/2, New York, 2010., str. 381-392, <https://doi.org/10.3758/BRM.42.2.381>; KRISTOPHER KYLE i dr., "Evaluating evidence for the reliability and validity of lexical diversity indices in L2 oral task responses", u: *Studies in Second Language Acquisition*, 46/1, Cambridge, 2024., str. 278-299, <https://doi.org/10.1017/S0272263123000402>.

²³ FRED ZENKER – KRISTOPHER KYLE, "Investigating minimum text lengths for lexical diversity indices", u: *Assessing Writing*, 47, Amsterdam, 2021., <https://doi.org/10.1016/j.asw.2020.100505>; KRISTOPHER KYLE i dr., "Evaluating evidence for the reliability and validity of lexical diversity indices in L2 oral task responses", str. 288; YVES BESTGEN, "Measuring Lexical Diversity in Texts: The Twofold Length Problem", u: *Language Learning*, 74/3, Hoboken, NJ, 2024., str. 638-671, <https://doi.org/10.1111/lang.12630>.

²⁴ SCOTT A. CROSSLEY, "Linguistic Features in Writing Quality and Development: An Overview", u: *Journal of Writing Research*, 11/3, Antwerp, 2020., str. 415-443, <https://doi.org/10.17239/jowr-2020.11.03.01>; BIRSU KANDEMIRCI i dr., "What is creative in childhood writing? Computationally measured linguistic characteristics explain much of the variance in subjective human-rated

nje većega broja različitih riječi. S povećanjem broja pojavnica bilo kojega teksta razumljivo je da se nakon određene brojke smanjuje omjer natuknica i pojavnica jer se mnoge natuknice preklapaju, tj. ponavljaju.²⁵ Što je omjer natuknica i pojavnica niži, tekst ima manje različitih riječi i po toj je mjeri jednostavniji.²⁶

Kako bi se došlo do navedenih mjera, potrebno je doći do ulaznih parametara, a to su ukupan broj: pojavnica (eng. *tokens*), samoznačnica (eng. *lexical words*) i natuknica (eng. *types*). Z. Jelaska i V. Baričević ističu da podjednaka leksička raznolikost nekih tekstova ne upućuje na jednako zahtjevan rječnik korišten u njima te pojašnjavaju: "Tri su univerzalne vrste riječi: imenice, glagoli i pridjevi (Marković 2012) znatno obavjesnije, povezanije s izvanjezičnom stvarnošću. I ostale su samoznačnice (nazivane i leksičkim riječima), posebno prilozi, obavjesnije, nego suznačnice (nazivane i gramatičkim riječima), kojima je glavna uloga uspostavljanje odnosa među riječima u rečenici. Ako je u dva teksta jednak omjer pojavnica i natuknica, tekst s više samoznačnica bit će leksički zahtjevniji od teksta s manjim udjelom samoznačnica."²⁷ Autorice ističu da leksička raznolikost i leksička gustoća ne govore ništa o značenjskoj složenosti riječi samih, ili u veza s ostalim riječima, "one govore o odabranomu rječniku teksta u odnosu na broj i vrstu pojavnica."²⁸

Iako to nije izravna tema ovoga rada, potrebno je objasniti i mjeru istančanosti. **Leksička istančanost** (eng. *lexical sophistication*) odnosi se na upotrebu istančanih i naprednih riječi,²⁹ odnosno rjeđih,

creativity scores", u: *Learning and Individual Differences*, 118, Amsterdam, 2025., <https://doi.org/10.1016/j.lindif.2025.102626>.

²⁵ VACLAV BREZINA, *Statistics in Corpus Linguistics: A Practical Guide*, Cambridge University Press, Cambridge, 2018.

²⁶ VALENTINA BARIČEVIĆ – MARTINA KEKELJ, "Udio vrsta riječi u tekstu na primjeru Lukina i Ivanova evanđelja", u: *Lahor: časopis za hrvatski kao materinski, drugi i strani jezik*, 8, Zagreb, 2009., str. 170-182.

²⁷ Z. JELASKA – V. BARIČEVIĆ, "Leksička jednostavnost i značenjska složenost rječnika Ivanova evanđelja", str. 117.

²⁸ *Isto*, str. 122.

²⁹ MINKYUNG KIM – SCOTT A. CROSSLEY – KRISTOPHER KYLE, "Lexical Sophistication as a Multidimensional Phenomenon: Relations to Second Language Lexical Proficiency, Development, and Writing Quality", u: *The*

naprednijih i kontekstno primjerenih leksema u tekstu, što znači na odmak od općega svakodnevnoga vokabulara prema preciznijem i stilistički prikladnijem leksičkom izboru.³⁰ Drugim riječima, ona označuje razinu razvijenosti i naprednosti leksičkog izbora u tekstu, a odnosi se ponajprije na učestalost i udio rjeđih, stilistički ili semantički složenijih riječi u odnosu na opći, visokočestotni leksik.³¹ Obično se izražava omjerom riječi prema učestalosti u tekstu. Zajedničkom primjenom leksičke istančanosti i leksičke raznolikosti, koja prikazuje broj natuknica, ostvaruje se cjelovit i međusobno komplementaran uvid u leksičko bogatstvo nekog autora.

Brojanje pogrešaka (eng. *number of errors*) je jedna od mjera leksičkoga bogatstva koja može pokazati jezični razvoj ili stupanj ovladanosti vokabularom. U analizi npr. učeničkih tekstova može se očitovati kao pogrešan izbor riječi s obzirom na stil, gramatičku kategoriju i sl. Tom mjerom može se otkriti koliko učenik razumije značenje riječi i njihove funkcije u komunikaciji.³² Analizom broja pogrešaka omogućuje se uočavanje jezičnih obrazaca, kontrastivna analiza te praćenje odstupanja od norme. Ta se mjera često koristi u analizi učeničkih radova ili tekstova učenika koji usvajaju neki strani jezik.³³ Takve pogreške važan su pokazatelj razvoja leksičke kompetencije jer otkrivaju na koji način učenici povezuju značenja, prenose obrasce iz materinskoga jezika i usvajaju pravila uporabe riječi u stranome ili materinskom jeziku.³⁴

Modern Language Journal, 102/1, Hoboken, NJ, 2018., str. 120-141, <https://doi.org/10.1111/modl.12447>.

³⁰ J. READ, *Assessing vocabulary*, str. 203.

³¹ BATIA LAUFER – PAUL NATION, "Vocabulary size and use: Lexical richness in L2 written production", u: *Applied Linguistics*, 16/3, London, 1995., str. 307-322; J. READ, *Assessing vocabulary*, str. 145-160; P. M. MCCARTHY – S. JARVIS, *nav. dj.*, str. 381-392; DOUGLAS BIBER – SUSAN CONRAD – RANDI REPPEN, *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, Cambridge, 1998.

³² J. READ, *Assessing vocabulary*, str. 203.

³³ NIVES MIKELIĆ PRERADOVIĆ, "Označavanje pogrešaka u CroLTeC-u (računalnom učeničkom korpusu hrvatskog kao stranog jezika)", u: *Rasprave Instituta za hrvatski jezik*, 46/2, Zagreb, 2020., str. 899-920, <https://doi.org/10.31724/rihjj.46.2.24>.

³⁴ VIŠNJA PAVIČIĆ TAKAČ, *Razvoj leksičke kompetencije u inome jeziku*, Sveučilište Josipa Juraja Strossmayera, Filozofski fakultet, Osijek, 2019., str. 38.

2.2. *Dosadašnja istraživanja mjera leksičkoga bogatstva na hrvatskome jeziku*

Mjere leksičkoga bogatstva bile su predmet istraživanja u hrvatskome jeziku u vrlo širokom rasponu s obzirom na složenost i veličinu leksičke građe. Prva su istraživanja obuhvatila njihovu primjenu na biblijskim tekstovima,³⁵ dok su se noviji radovi ponajprije usmjerili na različite skupine govornika i vrste diskursa. Tako su mjere leksičkoga bogatstva primjenjivane u istraživanjima početnica za hrvatski jezik,³⁶ vrjednovanja leksičke kompetencije učenika starije osnovnoškolske dobi,³⁷ u analizama pisanog pripovjednog diskursa govornika s razvojnim jezičnim poremećajem,³⁸ kao i u istraživanjima povezanosti leksičke raznolikosti, koherentnosti teksta i emocionalne valencije pripovijedanja.³⁹

U ovom pregledu posebno se izdvajaju istraživanja usmjerena na lirski leksički registar hrvatskih pjesnika, na temelju čijih će se rezultata provesti usporedba dobivenih vrijednosti s rezultatima ovog istraživanja. Analizirana su morfološka i leksička obilježja poetskoga diskursa Vesne Parun⁴⁰ te ranog stvaralaštva Josipa Pupačića, posebice

³⁵ V. BARIČEVIĆ – M. KEKELJ, "Udio vrsta riječi u tekstu na primjeru Lukina i Ivanova evanđelja", str. 172; Z. JELASKA – V. BARIČEVIĆ, "Leksička jednostavnost i značenjska složenost rječnika Ivanova evanđelja", str. 128.

³⁶ Ž. RADIĆ – J. KUVAČ KRALJEVIĆ – M. KOVAČEVIĆ, "Udžbenik kao poticaj ili prepreka leksičkomu razvoju", str. 45.

³⁷ M. ČEŠI, *Ovladanost standardnim hrvatskim jezikom na kraju obveznoga obrazovanja*, doktorski rad, str. 100.

³⁸ GORDANA HRŽICA – SARA KOŠUTAR – MATEA KRAMARIĆ, "Rječnička raznolikost pisanih tekstova osoba s razvojnim jezičnim poremećajem", u: *Hrvatska revija za rehabilitacijska istraživanja*, 55/2, Zagreb, 2019., str. 14-30.

³⁹ JELENA KUVAČ KRALJEVIĆ – ANA MATIĆ ŠKORIĆ – MATEJA GABAJ, "Personal narratives of school-age children: A cross-sectional developmental study", u: *Folia Phoniatrica et Logopaedica*, 75/6, Basel – New York, 2023., str. 412-430.

⁴⁰ MARIJANA BAŠIĆ – SANJA BARIČEVIĆ, "Morfološka i leksička obilježja poetskoga diskursa Vesne Parun", u: *Hum: časopis Filozofskog fakulteta Sveučilišta u Mostaru*, 20, Mostar, 2018., str. 29-65.

prve objavljene zbirke *Mladići* (1955.).⁴¹ Računalno⁴² je obrađen i Šimićev pjesnički registar, pri čemu su analizom mjere leksičke istančanosti konstruirane priređene konkordancije, čestotni abecedni rječnik te čestotni i odostražni rječnik, temeljen na korpusu od ukupno 229 pjesama, bez istaknute periodizacije. Čestotnost Šimićeva pjesništva bila je izazov i prije računalne obradbe, pa je Franičević 1964. godine ručno brojeći (uz razumljive pogreške) dao pregled Šimićevih najčešćih riječi. Izdvojio je riječi *tijelo*, *duša*, *čovjek* kao čestotno primarne, dok neke riječi – npr. *Bog* – uopće nije naveo zbog političkih okolnosti u kojima je radio.⁴³

3. Metodologija

Za potrebe rada koristilo se računalnom metodom obrade tekstova koja je organizirana kao ponovljiv postupak nad skupom dokumenata, pri čemu su svi tekstovi najprije kopirani u čiste ".txt" datoteke u UTF-8 kodiranju. Time je osigurano da analiza uvijek kreće iz istoga, usporediva ulazna formata, neovisno o tome iz kojeg su izvora tekstovi preuzeti, a ujedno se smanjuje mogućnost tehničkih odstupanja (npr. problemi s prikazom znakova) koja mogu utjecati na računalnu obradbu. Metoda je operacionalizirana kroz skriptu Python koja automatizira cijeli tijek obradbe i izračuna. U skripti se služi specijaliziranim modulima: CLASSLA za jezičnu obradu hrvatskoga teksta i LexicalRichness za izračun standardnih mjera leksičke raznolikosti.⁴⁴ U konfiguraciji koja uključuje tokenizaciju, morfosintaktičko označavanje i lematizaciju koristilo se modulom CLASSLA,

⁴¹ SANJA BARIČEVIĆ, "Neka obilježja Pupačićeva leksika u zbirci pjesama *Mladići*", u: *Croatica et Slavica Iadertina*, 20/2, Zadar, 2024., str. 117-141, <https://doi.org/10.15291/csi.4614>.

⁴² VJERA LOPINA, *Računalna obrada jezika i stila u pjesništvu A. B. Šimića*, doktorska disertacija, Filozofski fakultet u Zagrebu, Zagreb, 2012.

⁴³ *Isto*, str. 22.

⁴⁴ NIKOLA LJUBEŠIĆ – LUKA TERČON – KAJA DOBROVOLJC, "CLASSLA-Stanza: The Next Step for Linguistic Processing of South Slavic Languages", u: ŠPELA ARHAR HOLDT – TOMAŽ ERJAVEC (ur.), *Proceedings of the Conference on Language Technologies and Digital Humanities (JT-DH-2024)*, Institute of Contemporary History, Ljubljana, 2024., str. 252-253, <https://doi.org/10.5281/zenodo.13936406>; <https://pypi.org/project/classla/> – CLASSLA.SI, "classla", (21. 1. 2026.).

pri čemu dokumentacija naglašava da lematizacija zahtijeva prethodno izvršenu tokenizaciju i POS-označavanje.⁴⁵

Nakon jezične obradbe, skripta iz dobivena prikaza izvodi opće leksičke pokazatelje na razini dokumenta, prvenstveno broj riječi i broj različitih riječi te daje pregled učestalosti riječi kao potporu tumačenju rezultata. Uz to se na općoj razini promatra i odnos samoznačnica naspram suznačnica kako bi se dobio indikator "sadržajnosti" teksta, bez zadiranja u pojedinosti implementacije razvrstavanja. Leksička raznolikost procjenjuje se primjenom više komplementarnih mjera.⁴⁶ Tradicionalna mjera TTR i robusnija mjera MTLD izračunate su pomoću biblioteke *LexicalRichness* koja u dokumentaciji navodi, primjerice, da se MTLD računa s pragom (zadano 0.72).⁴⁷ Kako bi usporedbe bile stabilnije i manje osjetljive na duljinu teksta, skripta dodatno računa i kliznu varijantu TTR-a (MATTR), tj. pomični prosjek TTR-a u prozorima fiksne veličine u cijelomu dokumentu, što je pristup koji *LexicalRichness* također podržava kao zasebnu mjeru.⁴⁸

Postupak je postavljen tako da se za svaki dokument automatski generira standardiziran tekstualni izvještaj pa su svi rezultati dobiveni pod istim uvjetima i u istom formatu. Time se dobiva konzistentna osnova za usporedbe među dokumentima i za izvještavanje u radu, a cijeli proces ostaje ponovljiv jer se nad bilo kojim novim tekstom (*txt*) može primijeniti isti analitički tijek. Nakon toga, pristupilo se analizi POS-oznaka u sva tri podkorpusa. Ukupan broj pojava obuhvaća sve riječi osim rečeničnih i pravopisnih znakova te simbola (PUNCT, SYM). Dodatno se provjerilo koje su vrste riječi označene oznakom X u svakom podkorpusu. Oznaku X dobivaju riječi koje su nerazumljiv materijal, krnje riječi i neanalizirani

⁴⁵ N. LJUBEŠIĆ – L. TERČON – K. DOBROVOLJC, *nav. dj.*, str. 261-262.

⁴⁶ P. M. MCCARTHY – S. JARVIS, *nav. dj.*, str. 381.

⁴⁷ P. M. MCCARTHY – S. JARVIS, *nav. dj.*, str. 384-385; LUCAS SHEN, "Attributes and Methods in *LexicalRichness*", u: *LexicalRichness Documentation*, <https://lexicalrichness.readthedocs.io/en/latest/> (14. 1. 2026.).

⁴⁸ MICHAEL A. COVINGTON – JOE D. MCFALL, "Cutting the Gordian Knot: The Moving-Average Type-Token Ratio (MATTR)", u: *Journal of Quantitative Linguistics*, 17/2, London, 2010., str. 96, <https://doi.org/10.1080/09296171003643098>.

strani izrazi.⁴⁹ U prvoj je fazi ukupno 18 riječi (0,41%) s oznakom X, npr. *tintan*, *lune*. U drugoj fazi nije pronađena nijedna oznaka X (0.00%), dok je u trećoj fazi ukupno 23 riječi (0,42%), npr. *epoque*, *rois*. Ručno dodavanje tih riječi u skupinu samoznačnica ne bi bitno promijenilo dobiveni rezultat leksičke gustoće. Kako je cilj istraživanja utvrditi leksičku gustoću i leksičku raznolikost različitih Šimićevih faza, uzima se ukupan broj natuknica nakon normalizacije na mala početna slova.

4. Raščlamba

U ovome se poglavlju prikazuju rezultati leksičke gustoće i leksičke raznolikosti u različitim fazama Šimićeva pjesništva te njihove međusobne razlike. Uspoređivanjem rezultata tih odnosa zaključuje se o leksičkoj zahtjevnosti i mogućnosti da se odrede faze prema postojećim tumačenjima Šimićeve poetike. Razmatraju se i odabrani interpunkcijski znakovi kao dodatni pokazatelji promjena u pjesničkom izrazu i odabranoga broja faza.

4.1. Leksička gustoća

U svim analiziranim fazama Šimićeva pjesništva odnos samoznačnica i suznačnica pokazuje promjene relevantne za opis leksičke gustoće. U prethodnim istraživanjima Šimićeva leksičkoga registra već je uočeno da među najčešćim riječima prevladavaju suznačnice poput *i*, *u*, *se*, *na* i *je*, dok se samoznačnice pojavljuju rjeđe na početku popisa čestotnosti.⁵⁰ Naši rezultati potvrđuju takav odnos, ali istodobno pokazuju i razlike među fazama Šimićeva pjesništva, što je prikazano u tablici 1.

⁴⁹ <https://universaldependencies.org/u/pos/all.html> – UNIVERSAL DEPENDENCIES CONTRIBUTORS, "Universal POS tags" (14. 1. 2026.).

⁵⁰ V. LOPINA, *Računalna obrada jezika i stila u pjesništvu A. B. Šimića*, str. 22.

Tablica 1. Broj samoznačnica i suznačnica Šimićeve stvaralaštva

Faza	Broj samoznačnica	%	Broj suznačnica	%	100 %
1. rane pjesme	2897	66,02	1491	33,98	4388
2. zbirka <i>Preobraženja</i>	2357	64,10	1320	35,90	3677
3. ciklus <i>Pjesme o tijelu</i>	3315	59,89	2220	40,11	5535
UKUPNO	8569	63,01	5031	36,99	13.600

U Šimićeve pjesništvu samoznačnice čine oko tri petine teksta, a suznačnice oko dvije petine. Promatraju li se apsolutni brojevi, najviše je samoznačnica u ciklusu *Pjesme o tijelu* (3315), ali to proizlazi iz činjenice da je to najopsežniji potkorpus (5535 pojavnica). Zbog toga je za tumačenje važnija leksička gustoća. Ona se tijekom Šimićeve stvaralaštva postupno smanjuje, odnosno pada. Iznosi 0,660 u ranim radovima, 0,641 u zbirci *Preobraženja* te 0,599 u ciklusu *Pjesme o tijelu*. Te razlike u leksičkoj gustoći nisu velike u apsolutnom smislu, ali unutar istog autora i istoga (pjesničkog) registra deskriptivno su relevantne jer omogućuju precizniju usporedbu obavijesne zasićenosti. K tomu se može reći da predstavljaju tri različite faze njegova stvaralaštva.

Tablica 2. pokazuje da je pad leksičke gustoće između prve i druge faze blag, dok je između zbirke *Preobraženja* i ciklusa *Pjesme o tijelu* izraženiji, što znači da se između impresionističke faze i ciklusa *Pjesme o tijelu* leksička gustoća smanjuje. Ti rezultati mogu se povezati s drukčijom organizacijom leksičkoga registra novije poezije, ali i s njezinom manjom leksičkom zahtjevnosti.

Tablica 2. Promjene leksičke gustoće

Usporedba faza	Leksička gustoća	Postotni pad (%)
1. – 2.	0,660 → 0,641	2,88
2. – 3.	0,641 → 0,599	6,55
1. – 3.	0,660 → 0,599	9,24

Usporedba s leksičkom gustoćom u pjesništvu drugih hrvatskih pjesnika dodatno potvrđuje tvrdnju o tri Šimićeve faze, bez obzira na moguće metodološke poteškoće. Kao što je rečeno, tradicionalne mjere leksičkoga bogatstva, osobito TTR, pokazale su se izrazito

osjetljivima na duljinu teksta,⁵¹ a istraživanja varijacije registara⁵² i funkcionalnolingvistički pristupi⁵³ pokazuju da leksička distribucija ovisi o komunikacijskoj funkciji i žanru, ograničavajući izravnu usporedivost rezultata. Poetski, narativni ili religijski diskursi razlikuju se po stupnju ponavljanja riječi, njihovu značenju i različitim stilovima. Stoga su, u nastojanju da se osigura veća metodološka usporedivost rezultata, u ovome istraživanju uspoređivane vrijednosti unutar istoga diskursa proučavanoga autora te dvoje pjesnika kod kojih su iste mjere izračunate u prijašnjim istraživanjima.

U poeziji Vesne Parun leksička je gustoća visoka i iznosi 0,71.⁵⁴ Takav rezultat upućuje na bogat leksički inventar s velikim udjelom samoznačnica, što pridonosi izrazito slikovitom i značenjski razvedenom izrazu, tipičnomu za lirsku poeziju snažnoga metaforičkog i emocionalnog naboja. U pjesništvu Josipa Pupačića leksička je gustoća niža te iznosi 0,62,⁵⁵ što omogućuje interpretaciju da se radi o izraženijoj narativnosti njegova pjesničkoga registra (tablica 3.). Takvi se rezultati mogu tumačiti stilskim razvojem pojedinog autora i namjenom teksta. Budući da je riječ o istome žanru tj. pjesništvu, niža se leksička gustoća u Šimićevim kasnijim pjesmama može promatrati kao pokazatelj jednostavnijega jezičnog izraza (u smislu manjeg broja samoznačnica, a samim time i smanjenja obavijesnosti), što je u skladu s pojednostavljivanjem jezika karakterističnim za njegov novi poetski izričaj. Zbog razlika u opsegu korpusa, izboru

⁵¹ FIONA J. TWEEDIE – R. HARALD BAAYEN, "How Variable May a Constant Be? Measures of Lexical Richness in Perspective", u: *Computers and the Humanities*, 32/5, Dordrecht, 1998., str. 323-352, <https://doi.org/10.1023/A:1001749303137>; YVES BESTGEN, "Measuring Lexical Diversity in Texts: The Twofold Length Problem", u: *Language Learning*, 74/3, Hoboken, NJ, 2024., str. 638-671, <https://doi.org/10.1111/lang.12630>.

⁵² DOUGLAS BIBER, *Variation across speech and writing*, Cambridge University Press, Cambridge, 1988.; DOUGLAS BIBER, *Dimensions of Register Variation: A Cross-Linguistic Comparison*, Cambridge University Press, Cambridge, 1995.

⁵³ MICHAEL A. K. HALLIDAY, *An Introduction to Functional Grammar*, Edward Arnold, London, 1985.

⁵⁴ M. BAŠIĆ – S. BARIČEVIĆ, "Morfološka i leksička obilježja poetskoga diskursa Vesne Parun", str. 32-34.

⁵⁵ S. BARIČEVIĆ, "Neka obilježja Pupačićeva leksika u zbirci pjesama Mladići", str. 121-123.

tekstova i metodologiji izračuna, taj se odnos ipak može tumačiti samo okvirno, ali je dovoljan pokazatelj da je u Šimića doista prikladno govoriti o trima fazama.

Tablica 3. Leksička gustoća troje hrvatskih pjesnika



Razlike u leksičkoj gustoći nisu velike u apsolutnom smislu, ali su opisno važne i unutar samih Šimićevih faza i pojedinačnih autora. Najveća je podudarnost između Pupačića (0,62) i druge Šimićeve faze (0,64), potom Pupačića i treće Šimićeve faze (0,59). Vesna Parun (0,71) leksički je bitno složenija od impresionističke Šimićeve faze (0,66), najviše od treće faze. Dakle, na istraživačko pitanje o leksičkoj složenosti *Šimićeva pjesništva u odnosu na slična istraživanja hrvatskih pjesnika* s obzirom na veličinu i sastav korpusa, izbor pjesama, načina lematizacije i označavanja vrsta riječi te korištenje mjera, može se odgovoriti samo okvirno: Šimićeve uspostavljene dvije novije faze slične su pjesništvu Josipa Pupačića, pri čemu je druga faza leksički gušća od njega, a treća manje gusta. Drugim riječima, ovi podatci služe kao orijentacija u opisu položaja Šimićevih faza u unutar širega hrvatskoga pjesničkoga registra.

4.2. Leksička raznolikost

U rezultatima jezične obradbe Šimićeva pjesništva treba razlikovati broj natuknica prije normalizacije (npr. *Himnos* i *himnos* prikazuju se kao dvije natuknice) i nakon normalizacije (npr. *Himnos* i *himnos* prikazuju se kao jedna natuknica). U analizi leksičke raznolikosti polazi se od ukupna broja pojavnica i natuknica nakon normalizacije za svaku utvrđenu fazu (tablica 4.).

Tablica 4. Ukupan broj natuknica i pojavnica prema fazama

Faze Šimićeve poezije	Broj pojavnica	Broj natuknica prije i nakon normalizacije
1. impresionistička faza	4388	1304 (1300)
2. zbirka <i>Preobraženja</i>	3677	968 (966)
3. ciklus <i>Pjesme o tijelu</i>	5535	1294 (1288)
Ukupno	13.600	3566 (3554)

Budući da apsolutni broj natuknica sam po sebi ne pokazuje raznolikost leksika, razlike se među fazama tumače mjerama leksičke raznolikosti koje su prikladnije za usporedbu tekstova različitih duljina (tablica 5.).

Tablica 5. Promjene leksičke raznolikosti u Šimićevoj poeziji

Šimićeve poezije	LR MTLD (0,72)	LR MATTR (20)	LR MATTR (50)
1. impresionistička faza	146,529	0,926	0,851
2. zbirka <i>Preobraženja</i>	95,899	0,912	0,817
3. ciklus <i>Pjesme o tijelu</i>	56,153	0,867	0,755

Leksička raznolikost najveća je u Šimićevoj impresionističkoj fazi, a zatim se smanjuje u kasnijim fazama. Sve tri mjere upućuju na isti smjer promjene: najviše vrijednosti bilježe se u prvoj fazi, niže su u zbirci *Preobraženja*, a najniže u ciklusu *Pjesme o tijelu*. Budući da su MATTR i MTLD prikladniji za usporedbu tekstova različitih duljina, pad se ne može jednostavno svesti na razliku u opsegu potkorpusa, nego se može povezati i sa stilskim izborom u kasnijim fazama Šimićeve pjesništva. Usporedba leksičke raznolikosti (MTLD) među trima fazama navedena je u tablici 6.

Tablica 6. Pad leksičke raznolikosti po fazama

Usporedba faza	MTLD	Postotni pad (%)
1. – 2.	146,53 –> 95,89	-34,6
2. – 3.	95,89 –> 56,15	-41,4
1. – 3.	146,53 –> 56,15	-61,7

Pad leksičke raznolikosti između prve i druge faze (-34,6%) povećava se između zbirke *Preobraženja* i ciklusa *Pjesme o tijelu* (-41,4%). Veliki pad od 61,7% između prve i treće faze upućuje na znakovito

sužavanje raspona različitih leksema. Drugim riječima, u kasnijim fazama natuknice se češće ponavljaju. Takav je kvantitativni rezultat u skladu s postojećim tumačenjima Šimićeve pjesničkoga razvoja.⁵⁶

Dakle, prema navedenim se mjerama pokazuje da se doista može govoriti o trima fazama Šimićeve stvaralaštva. Od prve se Šimićeve faze leksička gustoća i leksička raznolikost postupno smanjuju, pri čemu je silazni trend najuočljiviji u ciklusu *Pjesme o tijelu*. Impresionistička faza pokazuje najveću leksičku zahtjevnost, druga srednju, a treća faza najmanju. Jasno, u poeziji niža leksička gustoća i raznolikost ne znače nužno i manju značenjsku složenost jer se, primjerice, semantička zgusnutost, eliptičnost, metaforičnost izraza mogu ostvarivati i s manjim brojem različitih riječi.

Rezultati dobiveni dvjema mjerama leksičkoga bogatstva, leksičkom gustoćom i raznolikošću, svojom preciznošću podupiru postojeća tumačenja Šimićeve poetike – dosljedno snižavanje dviju proučavanih mjera pokazuje prijelaz od rane, leksički razvedenije faze, prema kasnijoj poeziji reducirana izraza, slobodnoga stiha i jače koncentracije oko manjeg broja tematski i simbolički opterećenih riječi. Takvo čitanje podudarno je s opisima Šimića kao pjesnika koji traži "bit", "pravu i jedinu, nezamjenljivu riječ-simbol",⁵⁷ kao i s tumačenjima njegove kasnije poezije kao poezije sažetosti, unutarnje napetosti i sukoba tijela i duha.⁵⁸

4.3. Šimićeve interpunkcija po fazama

U proučavanju Šimićeve poetike brojni su autori istaknuli specifičnu uporabu, ali i izostavljanje interpunkcije, koje postaje jedno od ključnih obilježja njegove poezije. On svjesno "krši pravila", namjerno izostavlja, primjerice, zareze, a svaki interpunkcijski znak ima svoje značenje. Stoga će se u ovom dijelu rada provjeriti potvrđuje li i uporaba odabranih interpunkcijskih znakova predloženu podjelu po analiziranim fazama.

⁵⁶ MIROSLAV ŠICEL, *Povijest hrvatske književnosti XX. stoljeća*, Knjiga IV. Hrvatski ekspresionizam, Naklada Ljevak, Zagreb, 2007., str. 44.

⁵⁷ *Isto*, str. 44.

⁵⁸ K. BAGIĆ, *nav. dj.*, str. 367.

Slaven Jurić naglašava: "Izostavljanje interpunkcijskih znakova i svođenje deiktičnih i priloških oznaka na minimum ili njihovo potpuno uklanjanje, reduciranje lirskoga subjekta na neantropomorfni glas koji svijet promatra 'svevidećim' okom, bliži Šimićeveu liriku iz toga razdoblja tipu simultanizma koji bih, za razliku od globalnoga nazvao 'apstraktnim simultanizmom'."⁵⁹ I István Lukács⁶⁰ navodi da u zreloom pjesničkom razdoblju Šimić napušta tradicionalne vezane forme i interpunkciju te uvodi slobodno organizirane stihove koji stvaraju posebnu napetost i odražavaju doživljaj apsurdnosti svijeta. "Šimić je porabio (posebno u pjesmama) specifičnu interpunkciju. Odbacio je uglavnom točke, zareze i druge rečenične znakove, a ostavljao ih je samo gdje je morao izbjeći nerazumijevanje smisla pjesme."⁶¹

U ovom dijelu rada analizirat će se najčešći interpunkcijski znakovi korišteni u trima fazama: točka, zarez, dvotočka i upitnik. Ostali znakovi nisu predmet ovog istraživanja.

U prvoj je fazi najviše interpunkcijskih znakova, s 357 točaka, 354 zareza, 32 dvotočke i 14 upitnika. Zbirka *Preobraženja* sadrži najmanje interpunkcijskih znakova, s 36 točaka, 16 zareza, 32 dvotočke i 19 upitnika. U trećoj fazi raste broj interpunkcijskih znakova, sa 190 točaka, 140 zareza, 44 dvotočke i 48 upitnika. Ti podatci upućuju na Šimićeveu težnju k ekonomiji izraza: iz prve faze s rečenicama koje sadrže interpunkcijske znakove u većoj mjeri prelazi se prema fazama u kojima interpunkcija postaje selektivnija, a broj interpunkcijskih znakova i njihova funkcija prilagođeni su novomu poetskomu izrazu.

⁵⁹ SLAVEN JURIĆ, "Šimićev simultanizam u teoriji i praksi", u: VLADO PANDŽIĆ (ur.), *Zbornik radova s Međunarodnoga skupa o hrvatskom književniku Antunu Branku Šimiću*, Matica hrvatska, Drinovci, 2008., str. 407-418.

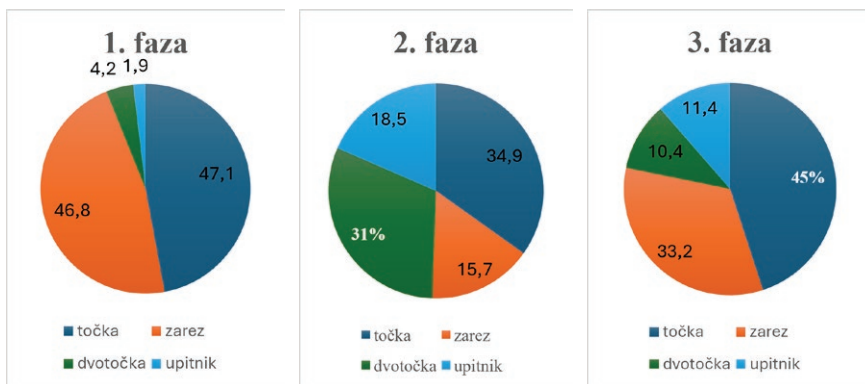
⁶⁰ ISTVÁN LUKÁCS, "Siromašni pjesnici o siromasima (Attila József i A. B. Šimić)", u: VLADO PANDŽIĆ (ur.), *Zbornik radova s Međunarodnoga skupa o hrvatskom književniku Antunu Branku Šimiću*, Matica hrvatska, Drinovci, 2008., str. 301-323.

⁶¹ MAZLLLOM KUMNOVA, "Antun Branko Šimić kod Albanaca", u: VLADO PANDŽIĆ (ur.), *Zbornik radova s Međunarodnoga skupa o hrvatskom književniku Antunu Branku Šimiću*, Matica hrvatska, Drinovci, 2008., str. 107-120.

Na grafikonu 1. ističe se udio točke, zareza, dvotočke i upitnika u odnosu na ukupan broj tih četiriju interpunkcijskih znakova (prva faza N = 757, druga faza N = 103, treća faza N = 422). Analizirano je i pojavljivanje tih interpunkcijskih znakova na 1000 pojava, što je omogućilo dobivanje usporedivih podataka (grafikon 2.).

Među najčešćim interpunkcijskim znakovima su točka, zarez, dvotočka i upitnik. U impresionističkoj Šimićevoj fazi najčešće se pojavljuju točka i zarez (47%); u drugoj fazi (*Preobraženja*) prevladavaju točka (35%) i dvotočka (31%), a u ciklusu *Pjesme o tijelu* – u trećoj fazi, ponovno su najčešći točka (45%) i zarez (33%). Posebno je zanimljivo da zbirka *Preobraženja* sadrži znatno veći udio dvotočaka i upitnika u odnosu na ostatak analizirana korpusa, a najmanji udio točaka i zareza.

Grafikon 1. Struktura interpunkcijskih znakova u fazama Šimićeve stvaralaštva



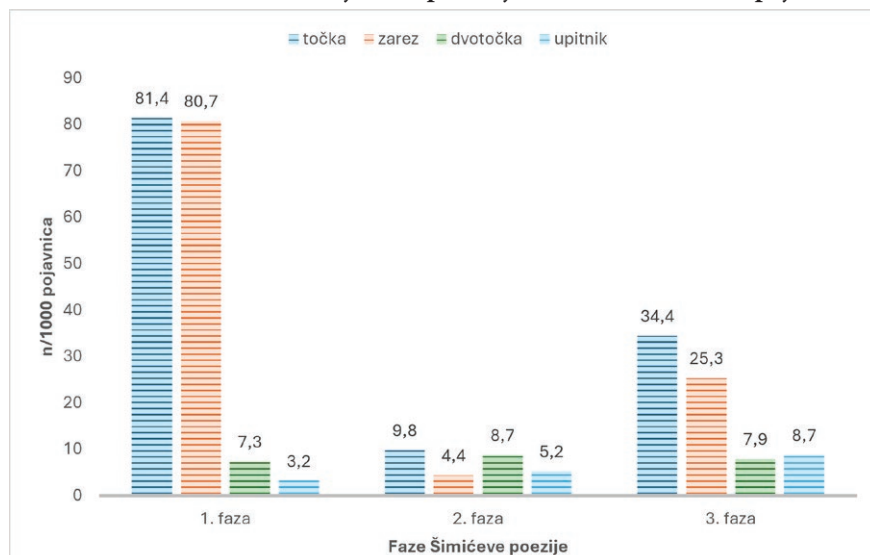
Nakon 1917. godine Šimić "počinje pisati slobodnim stihom, upotrebljava tzv. nepoetski vokabular, traga za ekspresivnim ritmom, uvodi privatni pjesnički pravopis, osamostaljuje ključne riječi u stihove."⁶² Pod "privatnim pjesničkim pravopisom" razumije se Šimićevo svjesno odstupanje od pravopisnih pravila jer pravopis postaje sredstvo za oblikovanje značenja. Stoga su promatrani Šimićevi interpunkcijski znakovi raznolikiji u značenju teksta u drugoj i trećoj fazi (*Preobraženja*, *Pjesme o tijelu*). Interpunkcija pokazuje stilsku promjenu: u

⁶² K. BAGIĆ, *nav. dj.*, str. 363.

prvoj fazi prevladavaju točke i zarezi, u drugoj (*Preobraženja*) točke i dvotočke dok u ciklusu *Pjesme o tijelu*, u trećoj fazi, ponovno prevladavaju točke i zarezi, ali ne u tolikoj mjeri kao u prvoj fazi.

Učestalost se upitnika povećava, a dvotočka zadržava svoju ravnomjernost pojavljivanja u svim trima fazama. Interpunkcijska obilježja dopunjuju rezultate o leksičkoj gustoći i leksičkoj raznolikosti te zajedno pridonose opisu Šimićeva pjesničkoga razvoja. Taj odnos ne treba tumačiti kao izravnu uzročno-posljedičnu vezu, nego kao podudaranje više pokazatelja koji zajedno upućuju na prijelaz prema sažetijem, redukcijom obilježenu i ekspresivnije organiziranu pjesničkomu izrazu.

Grafikon 2. Čestotnost četiriju interpunkcijskih znakova na 1000 pojava



Analiza četiriju interpunkcijskih znakova (točke, zareza, dvotočke i upitnika) pokazuje da je u kasnijim fazama interpunkcija rjeđa, nego u prvoj fazi. U zbirci *Preobraženja* posebno je vidljivo smanjenje točaka i zareza, dok se u odnosu na druge faze povećava važnost dvotočke i upitnika. Iako tim podacima nije obuhvaćena cjelokupna Šimićeva interpunkcija, oni podupiru tvrdnju da se njegova organizacija iskaza postupno udaljava od pravilno razgraničenih rečenica i dijelova rečenica prema ekspresivnijemu pjesničkom izrazu. Takvi

su rezultati u skladu s tumačenjima da Šimić u zrelijim fazama reducira interpunkciju i napušta tradicionalno oblikovanje iskaza.⁶³

5. Zaključak

Prema kvantitativnim rezultatima ovog istraživanja, Šimićevo se stvaralaštvo može podijeliti na tri faze koje su više ili manje različite jedna od druge. Prva, impresionistička faza, ima najviše vrijednosti leksičke gustoće i leksičke raznolikosti, ujedno i daleko najveći udio točaka i zareza među interpunkcijskim znacima. U kasnijim fazama vrijednosti tih mjera opadaju, ali tako da promatrane dvije mjere leksičkoga bogatstva u zbirci *Preobraženja* imaju srednje vrijednosti, a u ciklusu *Pjesme o tijelu* imaju najmanje, što znači najmanji udio samoznačnica i najuži raspon natuknica. Takav se pad može povezati sa stilskom ekonomijom i koncentracijom odabrana leksika. Drugim riječima, Šimić sužava svoj rječnik i usmjerava prema manjemu broju različitih i značenjski nosivih riječi. Odabrani interpunkcijski znakovi (točka, zarez, dvotočka, upitnik) ujednačenije su zastupljeni, ali i oni podupiru tvrdnju o trima različitim fazama Šimićeve stvaralaštva.

Usporedba Šimićevih faza pjesništva s pjesništvom Josipa Pupačića i Vesne Parun odnosi se isključivo na leksičku gustoću. Najveća je podudarnost u leksičkoj gustoći između Pupačićeva pjesništva (0,62) i druge Šimićeve faze (0,64). Leksička gustoća pjesništva Vesne Parun (0,71) veća je od prve, leksički najgušće, Šimićeve faze (0,66). Tom se usporedbom omogućuje isključivo orijentacijski opis položaja Šimićevih faza unutar širega hrvatskog pjesničkog registra.

Kako se u radu istražio relativno mali korpus jednog autora, a razlike u opsegu i sastavu podkorpora ipak nije moguće u potpunosti isključiti iz tumačenja, rezultati se ne mogu primjenjivati izvan analiziranih faza niti tumačiti kao hijerarhijsko rangiranje hrvatskih pjesnika. Bilo bi dobro sličnim postupcima lematizacije, morfosintaktičkoga označavanja i razvrstavanja riječi na samoznačnice i suznačnice usporediti pjesme većega broja hrvatskih pjesnika.

⁶³ S. JURIĆ, "Šimićev simultanizam u teoriji i praksi", str. 415; I. LUKÁCS, "Siromašni pjesnici o siromasima (Attila József i A. B. Šimić)", str. 303; M. KUMNOVA, "Antun Branko Šimić kod Albanaca", str. 111.