

Маријан Ј. Јелић¹
Универзитет у Новом Саду
Педагошки факултет
Катедра за језик и књижевност

ЛЕКСИЧКА ФРЕКВЕНЦИЈА У ВУКОВИМ ЛИРСКИМ НАРОДНИМ ПЕСМАМА

Апстракт: У раду смо истражили лексичку фреквенцију Вукових лирских народних песама (Српске народне пјесме, прва књига). Укупан број лексема је 5.102, док је број употребљених (текућих) речи 72.500. Овакав однос броја лексема и текућих речи указује на то да је невелика разноврсност лексике лирских народних песама, одн. да има доста понављања. Најфреквентније лексеме су: бити/јесам, он, ја, и ти, себе/се, у, на, не, да. Фреквенција према врстама речи изгледа овако: именице (2420), глаголи (1652), придеви (574), прилози (212), заменице (49), предлози (64), бројеви (40), узвици (35), везници (30) и речце (22). У овим песмама, које је Вук називао женским, најбројније су именице које означавају особе женског пола (девојка, мајка).

Кључне речи: лирске народне песме, лексика, фреквенција, врсте речи, текуће речи.

1. Уводне напомене

Иако је на српском језику претходних деценија урађено неколико десетина фреквенцијских речника (углавном књижевних дела), истраживања лексичке фреквенције Вукових умотворина тек су у зачетку. Рајна Драгићевић у раду *О човеку у Вуковом Рјечнику* наводи списак 10 најфреквентнијих речи, те наводи најфреквентније именице, глаголе и заменице у *Српском рјечнику* (Dragičević, 2019:pp.28–29). Ауторка скреће пажњу и на *Фреквенцијски речник сабраних дела Вука Ст. Караџића* Ђорђа Костића, који је као (интерни) рад објављен 2004. г. на Институту за експерименталну фонетику и патологију говора у Лабораторији за експерименталну психологију Филозофског факултета Уни-

¹jeliciso@uns.ac.rs

верзитета у Београду, који тек треба да буде објављен као део ширег пројекта *Квантитативни опис структуре српског језика*. Из истог рада Р. Драгићевић сазнајемо да је Софија Милићевић у необјављеном мастерском раду (одбрањеном 2011. године) истраживала фреквенцију именица у српским народним пословицама.

Полазећи од тога како је веома важно да се што детаљније и свестраније истраже Вукове народне умотворине, у овом раду бавили смо се лексичком фреквенцијом лирских народних песама (Српске народне пјесме², прва књига, 1841, даље у тексту СНП1).

У истраживању је коришћен софтвер AntConc ради добијања фреквенције свих (текућих) речи (у одређеном облику) које се јављају у песамама. Посао груписања различитих облика око једне лексеме морао је бити урађен ручно из два разлога – прво, софтвер који смо користили није имао ту могућност, и друго, веома је велики број варијаната³: екавске, јекавске и икавске речи (*девојка, дјевојка/ђевојка и дивојка*), испуштени слогови или делови слога (нпр. *опас'о* – укупно 2212 таквих речи), спојено и одвојено писање речи (*би ли и бил', јенђибула и јенђи-була, јели и је ли, некмо ли и некмоли, мермер камена и мермер-камену* итд.), речи са додатим партикулама или без њих (*како и каконо, који, којизи и којино* итд.) и сл., чине овај материјал веома тешким за софтверско лематизовање.

Приликом издвајања одредничке речи користили смо пре свега Речник САНУ, али и шестотомни (РМС) и једнотомни речник Матице српске (РСЈ).

Све речи које су сличног гласовног састава, односно које имају исту (или сличну) творбену основу (или корен) и притом имају потпуно исто значење наведене су као једна лексема (нпр. *аршлама* и *арашлама, баба* и *бабо, Београд* и *Бијоград* итд.).

² Користили смо издање Просвете и Нолита из 1987. год. Пун библиографски податак в. у одељку Извори.

³ За илустрацију наводимо све појавне облике/варијанте придева бео: Бел 1, бел 2, Бео 1, бео 4, беда 22, Бела 7, беле 14, Беле 2, бели 7, Белим 1, белим 4, белима 1, бело 14, Бело 2, белог 1, Белога 1, белога 1, белој 1, белом 13, Бела 2, белу 8, беља 1, Беље 1, Бељи 1, б'јел 1, б'јела 39, Б'јеле 3, б'јеле 9, б'јели 3, б'јелијех 1, Б'јелим 1, б'јелим 1, б'јело 16, Б'јело 3, б'јелој 2, б'јелом 1, Б'јелом 2, б'јеломе 1, б'јелу 16, Б'јелу 2, бијел 6, бијела 19, Бијела 2, Бијеле 1, бијеле 37, бијелим 2, бијелијем 1, Бјели 1, бијело 10, бијелога 7, бијелој 1, бијелом 4, бијеломе 24, бијелу 24, Бијо 1, бијо 4, билом 1, биломе 1, жут-бео 1.

2. Лексичке одреднице и текуће речи фреквенцијског речника, коефицијент разноврсности и средња фреквенција у СНП1

Укупан број лексема у СНП1 је 5.102, а број текућих речи је 72.500. Ако укупан број лексема (V) поделимо са бројем текућих речи (N), добијамо коефицијент разноврсности или коефицијент креативности (C).

$$C = V / N = 5.102 / 72.500 = 0,07$$

Дакле, коефицијент разноврсности износи 0,07.

Ако исте параметре доведемо у обрнут однос, тако што број текућих речи поделимо са укупним бројем лексема, добијамо средњу фреквенцију или индекс понављања (Fsr).

$$Fsr = N / V = 72.500 / 5.102 = 14,21$$

Средња фреквенција износи 14,21. Што је добијена вредност средње фреквенције или индекса понављања мања, то је речник богатији и разноврснији.

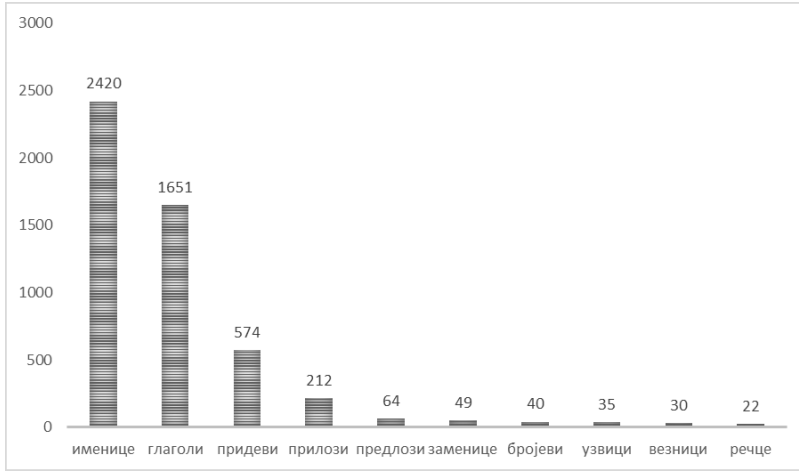
Иако однос лексичких одредница и текућих речи зависи од многих фактора (величина истраживачког корпуса, морфолошка структура језика, стилска разноврсност, функционални стилови, индивидуални стил аутора, више о овим факторима в. у Вукчевић 2005), добијени резултати – индекс понављања речи и коефицијент лексичке разноврсности довољно су добар показатељ лексичке разноврсности посматраног текста (корпуса).

Према Васић (1998) коефицијент разноврсности *Речника четири јеванђелиста* износи 0,04, а *Хазарског речника* 0,099. С друге стране, средња фреквенција (индекс понављања) у *Речнику четири јеванђелиста* износи 22,96, а у *Хазарском речнику* 10,099. Дакле, видимо да се према коефицијенту разноврсности и средњој фреквенцији лексика СНП1 налази између *Речника четири јеванђелиста* и *Хазарског речника*, разноврснија је од првог и мање разноврсна од другог речника. Упоредићемо наше резултате и са *Фреквенцијским речницима Горског вијенца* те *Луче микрокозма*. Коефицијент разноврсности *Горског вијенца* је 0,19, а *Луче микрокозма* 0,23; средња фреквенција *Горског вијенца* износи 5,252, а *Луче микрокозма* 4,355 (Вукчевић, 2005). Видимо да су ова два Његошева дела лексички разноврснија и богатија од СНП1.

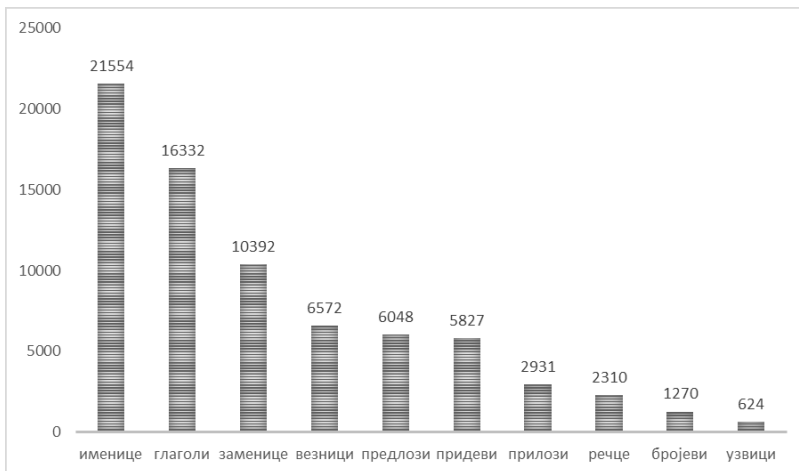
3. Дистрибуција врста речи у СНП1

Према врсти речи у СНП1 најбројније су именице (2420), затим глаголи (1652), знатно мање је придева (574) и прилога (212), док сваке од осталих врста речи има мање од 100.

У следећа два графика представимо број лексема и текућих речи према врстама речи.



Графикон 1: Број лексема у СНП1 према врсти речи



Графикон 2: Број текућих речи у СНП1

Из ова два графика видимо да су релационе речи – заменице, везници и предлози знатно фреквентније као текуће речи него придеви и прилози, којих иначе као лексема има знатно више у СНП1.

Према врсти речи најбројније су именице (2420), а међу њима најчешће се јављају: *девојка* (709), *мајка* (667), *Бог* (462), *двор* (349), *вода* (296), *коњ* (272), *брат* (231), *рука* (193), *сват* (191), *јунак* (187), *гора* (165), *злато* (150), *сунце* (141) итд. Именице са највећом фреквенцијом јесу тематске или кључне речи.

Према Драгићевић (2019) најфреквентније именице у Вуковом *Српском рјечнику* су *човјек* (406), *вода* (313), *коњ* (284), *дан* (266), *људи* (254), *дим*⁴ (219), *планина* (197), *Бог* (192), *кућа* (192), *пјесма* (189). Од ових 10 именица 3 се налазе и на списку најфреквентнијих именица у СНП1 (*Бог*, *вода*, *коњ*). Најфреквентнија именица у *Српском рјечнику* – *човек* јавља се у СНП1 само 18 пута. У *Речнику поезије Милана Ракића* најфреквентнија је именица *душа* (64), а затим следе: *ноћ* (48), *живот* (46), *дан* (38), *око* (37), *јад* (17), *суза* (17), *тело* (17), *месечина* (16), *мисао* (16) (Равковић, 1984). Видимо да међу 10 најфреквентнијих речи у СНП1 и *Речнику поезије Милана Ракића* нема ниједне заједничке речи. Списку лексема у корпусу Ракићеве поезије најближе су по фреквенцијском рангу из СНП1 речи *душа* (104), која је 93. по реду, *дан* (81), која је 121. по реду и *око* (72), која је 147. по реду.

На другом месту по бројности налазе се глаголи (1650), а најфреквентнији су: *јесам* (2591) / *бити* (769), *хтети* (726), *дати* (345), *говорити* (320), *доћи* (272), *ићи* (224), *љубити (се)* (181), *беседити* (166), *гледати (се)* (145), *носити (се)* (141).

Придева има 574: *млад* (478), *бео* (360), *леп* (261), *мио* (230), *зелен* (191), *стар* (183), *добар* (172), *златан* (120), *жут* (86), *румен* (75); а прилога 212: *када* (497), *већ* (181), *како* (88), *тако* (81), *где* (80), *ту* (68), *сада* (67), *још* (63).

Међу придевима најфреквентнији је придев *млад* (478), који је двадесета најфреквентнија реч у СНП1. Међу 10 најфреквентнијих придева готово сви су позитивног значења (осим придева *стар*). У народним песмама (лирским и епским) придеви су веома често стални епитети (*млада мома*, *в(ј)ерна љуба*, *б(и)јело лице* итд.).

Осталих врста речи има знатно мање – заменице (49): *он* (1953), *ја* (1927), *ти* (1477), *себе/се* (1314), *мој* (834), *што* (376), *тај* (335), *сав* (261), *који* (240),

⁴ Ова реч вероватно се грешком нашла међу 10 најфреквентнијих речи у *Српском рјечнику*. Наиме, завршеци многих глагола у *Српском рјечнику* наведени су као *дим* (без цртице испред), на пример *беседити*, *дим*. Таквих примера има много, а софтвер који пребројава речи није могао да препозна да је то завршетак код неких глагола у презенту, него је тај завршетак обрадио (пребројао) као посебну реч.

твој (203) итд.; бројеви (40): *два* (242), *један* (210), *три* (186), *други* (179), *трећи* (97), *девет* (82), *четири* (44), *први* (28), *двоје* (20), *оба* (18) итд.; везници (35): *и* (1657), *да* (1047), *а* (957), *али* (400), *на* (271), *ни* (247), *нак* (166), *или* (119), *нити* (119), *него* (118), *ако* (111) итд.; узвици (35): *ој* (222), *о* (95), *ајде* (62), *јао* (37), *благо* (21), *ах* (18) итд.; предлози (64): *у* (1276), *на* (1185), *од* (634), *за* (594), *са* (533), *по* (344), *до* (277), *под* (187), *из* (149) итд.; речце (22): *не* (1093), *ли* (408), *нека* (123), *ле* (105), *немој* (33), *ево* (32), *ето* (24), *еда* (14) итд.

Готово све речи које имају фреквенцију преко 100 припадају савременом српском језику. Ако их упоредимо са лексиком РСЈ (речник савременог српског језика), само лексеме *ле* и *ој* не припадају корпусу савременог српског језика (искључили смо и властита имена јер њих нема у РСЈ).

4. Најфреквентније лексеме у СНП1

Познато је да су релационе или функционалне речи најфреквентније у свим језицима света, па и у српском, тако да је резултат који смо добили када је у питању 10 најфреквентнијих лексема у СНП1 потпуно је у сагласју са тим. Од променљивих речи ту се по правилу налазе само заменице и помоћни глаголи.

У следећој табели наводимо 10 најфреквентнијих лексема у СНП1:

лексема	фреквенција
бити/јесам	3432
он	1980
ја	1928
и	1630
ти	1478
себе, се	1314
у	1276
на	1184
не	1092
да	1048

Табела 1: 10 најфреквентнијих лексема у СНП1

Из табеле се може видети и да је број текућих речи (појављивања) за свих 10 најфреквентнијих лексема преко 1.000. Ових 10 најфреквентнијих речи чини 16.370 текућих речи, а то је 22,58% свих текућих речи у СНП1, – што значи да ове речи покривају више од петине укупног текста.

Смиљка Васић за најфреквентније речи каже да су 'релационе речи – речи функтори, заједничке по својим садржајима за све фреквенцијске речнике уопште' (Vasić, 1998:р.27), речи које 'одређују односе у реченици' (р.34). Међу њима, по правилу, нема именица.

У фреквенцијским речницима на српском језику ове речи су увек у врху по фреквентности. Према подацима које наводи Рајна Драгићевић (2021:р.28), 10 најфреквентнијих лексема у Вуковом *Српском рјечнику* су: 1) *и* (7.044); 2) *на* (3.797); 3) *а* (3.350); 4) *је* (3.252); 5) *да* (2.614); 6) *од* (2.255); 7) *о* (2.252); 8) *не* (1.734); 9) *по* (1.732); 10) *што* (1.477). Као што видимо, пет је заједничких лексема које су најфреквентније у оба посматрана корпуса (*и, на, је/јесам, да, не*).

Ако списак од 10 најфреквентнијих речи проширимо са наредних пет: *а* (862), *мој* (833), *штети* (749), *девојка* (709), *мајка* (667), добијамо резултат да првих 15 речи по фреквенцији чини 20.190 текућих речи, а то је 27,87% укупног текста, што је и нешто већи проценат од оног који помиње Дејвид Кристал у *Кембричкој енциклопедији језика* – 'првих 15 речи требало би да представљају 25% текста' (Kristal, 1996:р.87). У овој скупини доминирају функцијске речи, али се међу њима налазе и две најфреквентније пунозначне лексеме – именице *девојка* (709) и *мајка* (667).

Најфреквентнијих 100 речи покрива 38.235 текућих речи, што је 52,74% текста. Након функцијских речи (углавном првих 15 по фреквенцији) долазе семантичке речи које су носиоци значења и представљају најуже лексичко језгро посматраног текста. Те речи (до ранговног нивоа 100, да поменемо само именице, глаголе и придеве) упућују на главне мотиве и ликове одн. учеснике: *девојка/дјевојка* (709), *мајка* (667), *Бог* (462), *брат* (330), *коњ* (272), *рука* (193), *сват* (191), *јунак* (187); места дешавања: *двор* (349), *вода* (289); особине: *млад* (478), *бео* (359), *леп* (264), *мио* (231), *зелен* (190), *стар* (183), *добар* (172); радње: *дати* (344), *говорити* (320), *доћи* (272), *ићи* (224), *видети* (187), *љубити (се)* (186).

Најфреквентнијих 1000 речи обухвата 61.979 текућих речи, што је 85,49% текста СНП1 (врло је сличан проценат и за 1.000 најфреквентнијих лексема *Горског вијенца* (83,66%) и *Луче микрокозма* (85,77%) (Vukčević, 2005).

Претпоставља се да ако читалац разуме 75% речи неког текста, разумјеће и осталих 25%, односно може наслутити значење непознатих речи. У СНП1 75% текста је 54.375 текућих речи, а тај број покрива 463 речи (са најмањом фреквенцијом 23 (*стојати*)), што је 9,07% укупног броја лексема у СНП1. Готово све те речи употребљавају се и данас. Ако изузмемо властита имена и списак тих речи упоредимо са РСЈ (речник савременог српског језика), можемо

рећи да само следеће речи не припадају корпусу савременог српског језика: *ле, војно, нева/неве, дода, дилбер, кавад, бријеме, миле*. Дакле, у скубини речи које чине (шире) фреквенцијско лексичко језгро СНП1 (463), које покривају 75% текста, свега је неколико речи које не припадају савременом српском језику.

Само 200 лексема покрива 75% текста четири јеванђеља, док је број речи који покрива исти проценат текста *Хазарског речника* знатно већи и износи 700 лексема (Vasić, 1998:р.32).

Ако се најфреквентније именице *девојка* и *мајка* могу довести у везу са појмом лирског, онда ова лексика (75% текста) покрива широко поље мотива различитих врста лирских народних песама (сватовске, обредне, митолошке, шаљиве, љубавне, посленичке итд.). Навешћемо све ове лексема (463) које чине шире лексичко језгро СНП1.

Именице (196) (редослед је фреквенцијски): *девојка, мајка, Бог, двор, брат, вода, коњ, рука, сват, јунак, гора, злато, девер, сунце, глава, Мара, око, сестра, драги, момче, син, душа, љуба, лице, Јово/Јова, сеја, вино, коло, град, поље, род, ружа, срце, дан, нога, сна(х)а, цар, јабука, Мујо, врата, Јела, цвеће, бисер, море, срећа, соко, земља, кита, бан, вила, драга, књига, Павле, прстен, сестрица, пут, мома, момак, Омер, трава, босиљак/босиок, дар, кум, небо, јад, војно, Маре, ноћ, удовица, Богдан, година, ливада, свила, кошуља, (п)тица, грана, нева/неве, Петар, село, братац, господин, наранџа, ћерка, коса, чедо, девојчица, сабља, суза, водица, господар, име, јела, кућа, недра, венац, вече(р), војска, месец, коњиц, дукат, крило, лепота, паун, платно, пенџер, (п)шеница/вшеница, дода, јутро, луг, глас, наша, чардак, баба/бабо, госпођа/госпоја, дуња, помоћ, фала, благо, дилбер, Јанко, овца, вој(е)вода, дика, невеста, планина, реч, Фата, вечера, грло, камен, недеља, сребро, добро, Мерима, час, вера, ветар, зора, лађа, Никола, ружица, свекар, свет, чашиа, дете, долама, Љутица, пас, скриња, бор, звезда, колено, краљ, марама, перо, цвет, баишча, кока, рана, Ана, Јован, калпак, челебија, данак, домаћин, ћердан, Имбро, Јања, појас/појаз, ру(х)о/руво, сан, свекрва, слуга, снашица, Туре, Будим, Ђураћ, Марко, господа, драго, кавад, лов, рукав, Сава, skut, страна, струк, Асан-ага, бријеме, душек, Марија, отац, пиле, првијенац, санак, уста.*

Глаголи (105): *бити, хтети, дати, говорити, доћи, ићи, видети, љубити (се), беседити, гл(ед)ати (се), носити (се), поћи, имати, отићи, звати, велити, моћи, пити, немати, пасти/панути, родити (се), казати, играти (се), молити (се), чути, знати, узети, рећи, стати, седети, водити, (х)одити, м(ис)лити, одговарати, брати (се), вести, донети, клетти (се), устати, уватити (се), проћи (се), наћи (се), оженити (се), спа(ва)ти, уд(а)рити (се), шетати (се), давати, метати (се), отворити, растити, питати (се), просити, довести, послати, по-*

гледати, долазити, чекати, вити (се), скочити (се), поиграти (се), чинити (се), метнути, вечерати, женити (се), заспати, певати, умрети, учинити (се), зачути, узимати, убити, мири(са)ти, плакати, будити, остати (се), слушати, жалити (се), даровати, појати, пољубити (се), помоћи, пустити (се), савити (се), дозивати, оставити, пасти, ронити, фалити (се), бацати (се), вратити (се), допасти и допаднути, ранити, волети, карати (се), купити, лећи, однети, писати, даривати, попити, слати, угледати, највољети, садити, стојати.

Придеви (50): млад, бео, леп, мио, зелен, стар, добар, златан, жут, румен(ит), танак, црн, бож(и)ји/божи, драг, велик/вељи, свет, мали, ситан, јадан, свилен, туђ, висок, жив, нежењен, студен, сребрн, зао, рођени, десни, (х)ладан, мушки, ведар, веран, девојачки, раван, злаћен, једини, мртав, сјајан, црвен/црљен, луд, танан, царев, весео, господски, мудар, Омеров, тежак, сив, сладак.

Заменице (23): он, ја, ти, себе/се, мој, што(но), (о)тај, сав, који, ми, твој, свој, ко, ви, онај/они, наш, сваки, сам, шта, овај/ови, ништа, ни(т)ко, њен/њезин.

Бројеви (8): два, један, три, други, трећи, девет, четири, први.

Непроменљиве врсте речи (81): и, у, на, не, да, а, од(а), за, кад(а), с(а), али, ли, по, до, па, ни ој, већ(е), под(а), пак(а)/паке, где, из, него, нех(а), ка(о)/каоно, ил(и), нити, ако, ле, како(но), а, о, те(р)/терно, тако, још/јошт(е), код, пред(а), док, сад(а)/саде, око, ту(не), добро, преко, рано, међу, тихо/тијо, кроз(а), о, мало, б(р)ез, ајде, лепо, к(а), далеко, жао, јер(а), крај, здраво, натраг, синоћ, уз(а), брзо, куд, немој, низ(а), ево, данас(ке), јао(ј), онда(ј), љуто, весело, горе, ма, но, право, тамо, туда(р)/тудијер, ето(ти), испод, много/млого, баш.

У овој скупини речи само је 18 позајмљеница: цар, бан, босиљак/босиок, наранџа, ајде, дукат, пенџер, паиша, чардак, баба/бабо, дилбер, долама, скриња, марама, калтак, челебија, ђердан, душек.

Властина имена. Међу именицама чак 364 су властите, а највише је личних имена и презимена (247). Само једна властита именица има фреквенцију већу од 100 – лично име *Мара* (128). Првих 18 места на листи најфреквентнијих властитих именица припада личним именима: *Мара* (128), *Јово/Јова* (96), *Мујо* (78), *Јела* (77), *Павле* (65), *Омер* (60), *Маре* (56), *Богдан* (55), *Петар* (49), *Јанко* (36), *Фата* (35), *Мерима* (33), *Никола* (31), *Љутица* (30), *Ана* (27), *Јован* (27), *Имбро* (26), *Јања* (26). Следећи по бројности после личних имена су етници (41): *Туре* (26), *Турчин* (21), *Скадарка* (18), *Бошњанин* (13), *Млеци* (12), *Бугарин* (8) итд. Међу топонимима најфреквентнији су хидроними *Сава* (24) и *Дунав* (22).

Варијанте и нестандардна лексика. У СНП1 наглашена је лексичка варијантност – јер су песме сакупљене из различитих крајева и отуд што је велики број варијаната употребљен 'да се испуни врста', односно из метричких разлога.

У СНП1 има доста нестандардне лексике – око 550 лексема (при том, најчешће наспрам нестандардне лексеме постоји и стандардна варијанта). Најфреквентнији нестандардни облик је *ђевојка* (221), који је чешћи од стандардног *дјевојка* (178), а најчешће се јавља екавска варијанта *девојка* (291). Нестандардна варијанта *ђевер* (71) чешћа је од обе стандардне варијанте *девер* (42) и *дјевер* (15). Глагол *хтети* (665) има и варијанте *ктети* (4), *кћети* (1), *ићети* (12) *тети* (66) и *тити* (1). Прилог *где* (31) / *гдје* (12) јавља се и у варијантним облицима *гди* (37), *ђе* (59), *зђе* (1), *ге* (4), *дјено* (5), *гдино* (5) и *ђено* (3). Прилог *одатле* (10) има пет варијаната: *отале* (2), *отлен* (2), *отоле* (3), *отолем* (1), *отолен* (1).

Када је у питању територијална распрострањеност, највише је лексема, према назнакама у СР, из Војводине (39): *птица* (5) (и *тица* (44)), *гостоја* (6) (и *гостоја* (32)), *Дунав* (14) (*Дунаво* (7) и *Дунај* (1)), *швалер* (12), *рузмарин* (6), *кафа* (4) (и *кава*) (4), *чоја* (2) (и *чо(х)а* (3)), *паргал* (4), *адвокат* (3), *официр* (3), *јурат* (2), *богословац* (2) итд.

Песничке речи. Једна од карактеристика лексике СНП1 јесте и та да у тим песмама има лексема које се јављају безмало само у лирским (народним) песмама. Вук је у *Српском рјечнику* такве речи називао стајаћим, а поред њих било је и других за које је давао напомену да се '(најчешће) јављају у песми/ песмама' (Јелић, 2022:р.58). Таквих речи у СНП1 има 188 (више о песничким речима у Вуковим лирским песмама види у Јелић 2022).

Најфреквентније стајаће (песничке) речи су (према *Српском рјечнику*) *двор* (349), *драги* (114), *љуба* (101), *драга* (67), *мома* (61), *војно* (56), *чедо* (47).

Позајмљенице. Речи страног порекла у СНП1 има 487 (9,55%). Највише је позајмљеница из турског језика (362), док је речи које воде порекло из других језика знатно мање – грч. (27), тал. (27), мађ. (18), лат. (18), нем. (8), фр. (6), перс. (5), ар. (2), рус. (1), ар.–лат. (1), ар.–тур. (1), млет. и тур. (1), перс.–тур. (1), тур. и мац. (1), фр. тал. (1), хебр. 1, тур. и тал. 1. Међу првих 130 најфреквентнијих лексема ниједна није страног порекла. У 220 најфреквентнијих речи нема ниједног турцизма. Иако су турцизми најбројнији, прве три најфреквентније позајмљенице нису турцизми, већ им је порекло из других језика: *цар* лат. (78), *бан* мађ. (68), *босиљак/босиок* лат. (58).

5. Речи са фреквенцијом 1 (хапакси).

Укупан број хапакса је 1829 (лексема које се појављују само једном) (35,86%). 'Досадашња истраживања су показала да се у сваком корпусу половина свих одредница појављује само једном ...' (Gruden, 2015:p.180). И ово одступање указује на мању лексичку разноврсност, а то је последица великог броја понављања (варијанте песама, исте или сличне теме и мотиви, стихови итд.). У ауторским књижевним делима лексика са фреквенцијом 1 указује на стилско богатство аутора. У СНП1, чији аутор није појединац него колектив, за хапаксе је карактеристично да се најчешће ради о лексемама које су ређе у употреби (и данас и у Вукове време) или се уопште више не употребљавају (а има и таквих лексема за које не знамо тачно ни које је њихово значење (као што је, на пример, реч *илинча*)).

Скоро свака десета лексема са фреквенцијом 1 страног је порекла (183): *авли-марама, агалук, адићар, ала, алет, ал-ћергелет, аманет, амбер-душа, ан, антерија, анума, апостол, аримандрит, арфа, аста, астара, ат, аџамија, баг, бадањ, безерћан, баклава, балабан, барна, баш-терзија, беглук, без, белензука, белнук, бискуп, бокор, бокчалук, бубрег, бунгур* итд. Има много властитих именица (142): *Абиба, Ајкуница, Алил-челебија, Андро, Анђелина, Анета, Анка, Анте, Арапин, Аџ-Алијћ, Балатино језеро, Бегзада, Богде, Богородица, Борић, Борковић, Брунка* итд.

Ту су и застареле речи (43): *бојник, бол* ('болест'), *даврана, дваши, додесити се, забун, заклоњати, згубити, капиција* итд.; стајаће речи и оне које се јављају у песмама (19): *армаган, батати, бисерче, боравити (сан), бућоглав, валаха, велен, вијерница, годити се, делкушица, дикица, дилберче, ћићеров, жнијевати* итд.; *покрајинизми: бата, башчованка, благословити, болта, бурунџукли, ваган, велета, веслоце, виленски, вишина, вучи-капа, газдашаг, газимир, гарофиље, гојајли* итд.; *историзми (6): беглук, вармећа, даија, делибаша, мула, сенат*, а има и речи које у нашим описним речницима (мислимо пре свега на РСАНУ) нису никако квалификоване и које се више не употребљавају, или се веома ретко употребљавају: *бајалица, банити се, бјелило, болник, гиздарев, кључаница* итд.

Наравно, у овој групи речи могу се наћи и речи које су данас сасвим обичне: *босоног, бости, бранити се, братић, братов, братство, бризнути, бунован, ваљати се, ван, варити, венчати се, веома, вијећати, висити, војнички, воће, врабац* итд.

Међу хапаксима веома су ретке релационе речи. Предлога, речци и узвика има свега 20, а везника нема.

6. Закључна разматрања

Ако бисмо од најфреквентнијих речи у СНП1 (именица, глагола, придева и заменица) саставили реченицу, она би гласила – *Ја⁵ сам млада девојка*. Ова конструисана реченица полазишна је тематска тачка из које се даље развијају теме и мотиви у Вуковим лирским народним песмама (СНП1).

Списак најфреквентнијих лексема сличан је као и у другим истраживаним делима и углавном га чине функционалне речи. Најфреквентнија реч у СП1 јесте глагол *бити/јесам* (3432). Низак коефицијент разноврсности (0,07) и висока средња фреквенција (14,21) указују на мању лексичку разноврсност, што је последица великог броја понављања (стихова, делова стихова, мотива итд.), која су карактеристична не само за лирске народне песме него и за народну књижевност уопште.

Да се подсетимо, Вук је све народне песме поделио на јуначке ('које људи певају уз гусле') и женске – 'које пјевају не само жене и ђевојке него и мушкарци' (СНП1, Предговор: XVII). Дакле, лирске народне песме су певале пре свега 'жене и ђевојке', те их је због тога Вук назвао женским. Те песме не само да су (најчешће) певале 'жене и ђевојке' него су (најчешће) и певале о женама и девојкама (најфреквентније именице у збирци су именице *девојка/дјевојка* (709) и *мајка* (667)). Овакав резултат истраживања упућује на закључак да бисмо могли и по критеријуму лексичке фреквентности ове песме (лирске) називати женским.

Извори

1. СНП1: Караџић Стефановић, В. (1987) *Српске народне пјесме I*. Београд, Просвета–Нолит.

Литература

1. Васић, С. (1998) Полазне основе новије српске прозе. Књ. 2, *Хазарски речник Милорада Павића: фреквенцијски речник* (1. изд.). Завод за уџбенике и наставна средства; Институт за педагошка истраживања.
2. Вукчевић, З. (2005) *Фреквенцијски речници Горског вијеница и Луче микрокосма*. Подгорица, Црногорска академија наука и умјетности
3. Груден, Т. (2015) *Политички дискурс у паролама, слоганима и графитима протеста НАТО бомбардовања СР Југославије 1999*. Докторска дисертација. Универзитет у Београду, Факултет политичких наука.

⁵ Међу заменицама најчешће се јавља лична заменица за треће лице, али ако се гледа сваки род засебно, онда је најчешћа у употреби лична заменица за прво лице.

4. Драгићевић, Р. (2021) О човеку у Вуковом Рјечнику. *Вуковим трагом на Косову и Метохији* (тематски зборник). Филозофски факултет Универзитета у Приштини с привременим седиштем у Косовској Митровици, Градска библиотека *Вук Караџић* у Косовској Митровици, 27–36.
5. Јелић, М. (2022) Песничке речи у Вуковим лирским народним песмама, *Узданица*. XIX/2, 55–67.
6. Караџић Стефановић, В. (2017) *Српски рјечник*: дигитално издање. Тома Тасовац (уредник). Београд, Институт за српски језик и Центар за дигиталне хуманистичке науке. <http://gaskovnik.org/recnici/VSK.SR>
7. Krištal, D. (1996) *Kembrička enciklopedija jezika*. Beograd, Nolit.
8. Павковић, В. (1984) *Речник поезије Милана Ракића*, Нови Сад, Матица српска.
9. РМС (1967–1976) *Речник српскохрватскога књижевног језика*, I–VI, Нови Сад, Матица српска.
10. РСАНУ (1959–) *Речник српскохрватског књижевног и народног језика*. Београд, Институт за српски језик, Српска академија наука и уметности.
11. РСЈ (2011) *Речник српског језика*. Друго, измењено и поправљено издање. Ур. М. Николић. Нови Сад, Матица српска.

Marijan J. Jelić
University of Novi Sad
Faculty of Education
Department of Language and Literature

LEXICAL FREQUENCY IN VUK'S LYRICAL FOLK POEMS

Summary

In this paper, we investigated the lexical frequency of Vuk's lyrical folk poems (Serbian folk poems, first volume). The total number of lexemes is 5102, while the number of used (current) words is 72500. The diversity coefficient is 0,07, and the mean frequency or repetition index is 14,21. This ratio of the number of lexemes and current words indicates that the lexicon of lyrical folk poems is not very diverse, i.e., that there is a lot of repetition, which is expected because it is one of the poems' important characteristic. The most frequent lexemes are: *biti/jesam, on, ja, i ti, sebe/se, u, na, ne, da*. The frequency, according to the part of speech, looks like this: nouns (2420), verbs (1652), adjectives (574), adverbs (212), pronouns (49), prepositions (64), numerals (40), exclamations (35), conjunctions (30) and particles (22). It is assumed that if the reader understands 75% of the words of a text, he will also understand the remaining 25%. In the observed corpus, 75% of the text covers 463 words (those with the highest frequency). Almost all of

these words belong to the corpus of the modern Serbian language. In lyrical folk poems, which Vuk called feminine, the most numerous are nouns that denote persons of the female gender (girl, mother), so even according to the criterion of the most frequent nouns these poems could be called feminine.

Sources

1. SNP1: Karadžić Stefanović, V. (1987) *Srpske narodne pjesme I*. Beograd, Prosveta–Nolit.

References

1. Dragičević, R. (2021) O čoveku u Vukovom Rječniku. *Vukovim tragom na Kosovu i Metohiji* (tematski zbornik). Filozofski fakultet Univerziteta u Prištini s privremenim sedištem u Kosovskoj Mitrovici, Gradska biblioteka *Vuk Karadžić* u Kosovskoj Mitrovici. 27–36.
2. Gruden, T. (2015) *Politički diskurs u parolama, sloganima i grafitima protesta NATO bombardovanja SR Jugoslavije 1999*. Doktorska disertacija. Univerzitet u Beogradu, Fakultet političkih nauka.
3. Jelić, M. (2022) Pesničke reči u Vukovim lirskim narodnim pesmama, *Uzdаница*. XIX/2, 55–67.
4. Karadžić Stefanović, V. (2017) *Srpski rječnik*: digitalno izdanje. Toma Tasovac (urednik). Beograd, Institut za srpski jezik i Centar za digitalne humanističke nauke. <http://raskovnik.org/recnici/VSK.SR>
5. Kristal, D. (1996) *Kembrička enciklopedija jezika*. Beograd, Nolit.
6. Pavković, V. (1984) *Rečnik poezije Milana Rakića*, Novi Sad, Matica srpska.
7. RMS (1967–1976) *Rečnik srpskohrvatskoga književnog jezika*, I–VI, Novi Sad, Matica srpska.
8. RSANU (1959–) *Rečnik srpskohrvatskog književnog i narodnog jezika*. Beograd, Institut za srpski jezik, Srpska akademija nauka i umetnosti.
9. RSJ (2011) *Rečnik srpskog jezika*. Drugo, izmenjeno i popravljeno izdanje. Ur. M. Nikolić. Novi Sad, Matica srpska.
10. Vasić, S. (1998) Polazne osnove novije srpske proze. Knj. 2, *Hazarški rečnik Milorada Pavića: frekventijski rečnik* (1. izd.). Zavod za udžbenike i nastavna sredstva, Institut za pedagoška istraživanja.
11. Vukčević, Z. (2005) *Frekventijski rečnici Gorskog vijenca i Luče mikrokozma*. Podgorica, Crnogorska akademija nauka i umjetnosti.

Преузето: 4. 10. 2023.
Корекције: 15. 10. 2024.
Прихваћено: 25. 10. 2024.