

Славко Ж. СТАНОЈЧИЋ*
Универзитет у Крагујевцу
Филолошко-уметнички факултет

Оригинални научни рад
Примљен: 28. 10. 2024.
Прихваћен: 27. 2. 2025.

КАТЕГОРИЗАЦИЈА ЛЕКСИЧКИХ КОЛОКАЦИЈА И ФРЕКВЕНЦИЈСКА КАТЕГОРИЗАЦИЈА ПУНОЗНАЧНИХ РЕЧИ У ТИПИЧНОМ УЏБЕНИЧКОМ ДИСКУРСУ**

Рад доноси резултате ауторовог истраживања корпуса уџбеничког дискурса, и појединачно – по годинама учења истраживаних предмета, као и збирно – у фреквенцијском аспект. Анализа се односи на велики број лексичких (пунозначних) речи следећих категорија: именица, глагола и придева. Управо је фреквенција датих лексема омогућила увођење појма (и термина) категоризације семантички доминантне лексеме и у истраживање конкретног дискурса који чини корпус овога рада. Такође, истраживање даје увид у елементе приказивања система колокације у датом дискурсу, а у раду се предлажу могућа решења категоризације лексичких колокација на основу савремених текстлингвистичких теорија.

Кључне речи: лексичке колокације, уџбенички дискурс, фреквенцијска категоризација, пунозначне речи, анализа корпуса.

Увод

У раду се бавимо истраживањем корпуса уџбеничког дискурса, природно, у датоме корпусу, и појединачно – по годинама учења истраживаних предмета, као и збирно – у фреквенцијском аспект. У обраду се, уз анализираних функционалне (граматичке) лексеме, узима велики број *лексичких (пунозначних) речи* свих категорија (именица, глагола, придева). Због тога, као и за прве, уз статистички дати корпус у целини, потребно је приказати и један број тих индикативних јединица – фреквенцијски, такође – у поступном поретку, чија је полазна црта прва година учења датог предмета. Како се показало у истраживању, управо је *фреквенција* датих лексема омогућила

* slavko.stanojcic@filum.kg.ac.rs; <https://orcid.org/0000-0001-7979-2212>

** Истраживање у овом раду финансирано је по Уговору о преносу средстава за финансирање научноистраживачког рада запослених у настави на акредитованим високошколским установама у 2024. години 451-03-65/2024-03/200198.

увођење појма (и термина) *семантички доминантне лексеме* и у истраживање конкретног дискурса који чини корпус овога рада.

Са своје стране, она је битан елеменат у приказивању *система колокације* у датоме дискурсу. Доследно дефиницијама наше лингвистике, природно, и овде полазимо од поставке да је *омеђени дискурс* ужа категорија појма *дискурс*, и поставке да се та ужа категорија може дефинисати као *поље посебне комуникацијске активности*, за коју је везан и посебан низ лексичких јединица изабран према *предмету* датог саопштавања.¹ Зато свако такво поље обично располаже речником зависним од предмета расправе (комуникације), а лексичке јединице тога (датог) поља могу се, природно, наћи и у другим (многим) пољима и са различитим значењем, и у различитим лексичко-граматичким колокацијама.

Анализа корпуса

Овде смо за приказ стања и тенденције изградње ученичке језичке перцепције у прегледаном корпусу у класи *лексичких (пунозначних) речи*, због рационалног приступа врло обимној њиховој материји као њиховог типичног представника, дали по неколико репрезентативних лексема – *именица, придева и глагола*.

Критеријум за њихов избор, међутим, био је у домену њихове *семантике* и *везе са темом дискурса* – за дати текст дата лексема схваћена је као доминантна, она је *предмет датог учбеничког дискурса*.² Испитивање њихове фреквенције показује сада, за разлику од фреквенције функционалних (граматичких) речи, пре свега – *семантичку структуру* текста која се нуди ученику основне школе за савладавање појмова у датим (школским) предметима.

Корпусна материја *функционалних (граматичких) речи* једнака је за оба корпуса учбеничког дискурса, и за (а) део наставе који ученика уводи у *чињење језичке заједнице* којој ученик припада, као и за (б) део наставе који ученика уводи у *чињење природе и друштва* којем ученик припада (Шире о овоме у Станојчић 2024). Супротно томе, материја *лексичких (пунозначних) речи* условљена је стварно *тематиком дискурса*, па тиме располаже разноврснијом лексиком. За опис у границама рационалног, економичног (чак и када је он фреквенцијски преглед материје дакле када је у домену статистич-

¹ Р. Картер за код нас уобичајену категорију с термином *омеђени дискурс* употребљава термин *поље дискурса* (*field of discourse*, в.: Ronald Carter, *Vocabulary – Applied Linguistic Perspectives*, Second edition, First published 1998, by Routledge, London EC4P 4EE – This edition published in the Taylor & Francis e-Library, 2002, 53–55).

² “Each field usually has a specialized, topic-related vocabulary... some items of which may turn up in other fields – -je

With different meanings (and often in different grammatical collocations...): for example, *wings* in air transport and in theatrical parlance; *lock* in the field of river or canal transport and in the language of security (Ronald Carter, *Vocabulary – Applied Linguistic Perspectives*, Second edition, First published 1998, by Routledge, London EC4P 4EE – This edition published in the Taylor & Francis e-Library, 2002, 53–54). В. и: V. Polovina, *Diskurs u funkciji kognitivne analize jezičkih kategorija*, Semantika i tekstlingvistika, Beograd, 1999, str. 125–135, где се говори о *омеђеном дискурсу*, који као појам има све карактеристике појма *field of discourse*, што укључује и значење појма и термина *ситуација*, као прихваћеног у нашој лингвистици.

ке лингвистике), то значи – констатовање доминантне лексики у дискурсу анализираних уџбеника и поређење њеног кретања по разредима, поређење по уџбеницима за те разреде.

I

У наредном тексту и табелама приказаћемо удео семантички доминантних пунозначних речи у уџбеницима матерњег језика – читанкама, као и фреквенцију пунозначних речи у уџбеницима за природу и друштво (у *Изворима* на крају рада наводимо библиографске податке).

У оба случаја, и у читанкама и у уџбеницима за природу и друштво, од првог до четвртог разреда, јасно се уочава веза лексики и њеног окружења (омеђеног дискурса, ситуације, контекста – терминима опште лингвистике речено) указује и на природу читавог дискурса³: он је у оквирима *дечјег света*, пре свега везаног за *мајку*, а онда за *безбрижно (светло) окружење, радозналост, тумачење стварности кроз игру, песму и школу* као носиоца баш таквог – *дечјег окружења*.

Још јаснију слику уочавамо кад погледамо следећу табелу:

2. Табела кретања датих доминантних лексема показује следеће њихово фреквенцијско кретање:

I разред	II разред	III разред	IV разред
Укупан број речи у читанкама			
4.441	5.230	8.295	12.629
именичке лексеме			
<i>прича</i> (51)	<i>прича</i> (59)	<i>прича</i> (115)	<i>прича</i> (156)
<i>зец</i> (47)	<i>зец</i> (8)	<i>зец</i> (47)	<i>зец</i> (53)
<i>песма</i> (43)	<i>песма</i> (70)	<i>песма</i> (180)	<i>песма</i> (139)
<i>дете</i> (44)	<i>дете</i> (14)	<i>дете</i> (68)	<i>дете</i> (132)
<i>лопта</i> (41)	<i>лопта</i> (2)	<i>лопта</i> (1)	<i>лопта</i> (15)
<i>школа</i> (36)	<i>школа</i> (26)	<i>школа</i> (16)	<i>школа</i> (42)
<i>реч</i> (35)	<i>реч</i> (64)	<i>реч</i> (79)	<i>реч</i> (77)
<i>мама</i> (32)	<i>мама</i> (35)	<i>мама</i> (17)	<i>мама</i> (15)
<i>лав</i> (30)	<i>лав</i> (2)	<i>лав</i> (37)	<i>лав</i> (30)
<i>птица</i> (29)	<i>птица</i> (18)	<i>птица</i> (57)	<i>птица</i> (39)
<i>сунце</i> (21)	<i>сунце</i> (15)	<i>сунце</i> (25)	<i>сунце</i> (53)

³ Терминима корпусне лингвистике речено, „...the high frequency of given lexical items can be a prominent feature in the characterization of text types. As it is stated in Baayen (1998: 90), [w]ord frequencies constitute a rich source of information on the relation between lexis and text type. General summary measures capture global properties such as repetitiveness and vocabulary richness, but much more detailed information is available if we allow ourselves to consider individual words in specific frequency ranges. The specificity of texts and themes is reinforced by the observation of words which occur typically in homogeneous texts” (Luísa Alice Santos Pereira, *The use of concordancing in the teaching of Portuguese*, How to Use Corpora in Language Teaching, Edited by John McH. Sinclair, John Benjamin's Publishing Company, Philadelphia – Amsterdam, 2004, 115).

придевске лексеме			
леп, -а, -о (32) плав, -а, -о (30) мали, -а, -о (30) велики, -а, -о (22) народни, -а, -о (16) златан, -ни, -а (14) сунчев, -а, -о (12)	леп, -а, -о (33) плав, -а, -о (6) мали, -а, -о (41) велики, -а, -о (26) народни, -а, -о (27) златан, -ни, -а (2) сунчев, -а, -о (9)	леп, -а, -о (54) плав, -а, -о (21) мали, -а, -о (82) велики, -а, -о (35) народни, -а, -о (39) златан, -ни, -а (12) сунчев, -а, -о (4)	леп, -а, -о (68) плав, -а, -о (23) мали, -а, -о (90) велики, -а, -о (67) народни, -а, -о (63) златан, -ни, -а (30) сунчев, -а, -о (4)
глаголске лексеме			
имати (55) питати (39) причати (19)	имати (45) питати (20) причати (17)	имати (36) питати (49) причати (15)	имати (107) питати (36) причати (30)

С обзиром на чињеницу – која се јасно уочава у првој подтабели – да се укупан број реализација морфосинтаксичких речи (обеју дефинисаних класа – граматичких и пунозначних речи) у дискурсу уџбеника за III (8.295) и IV (12.629) разред у односу на број тих реализација у уџбенику за I разред (4.441) и II разред (5.230) увећао за око 50%, очигледно је да је линија кретања *доминантних лексема* – заправо уједначена, равна, без квалитативних скокова у периоду од четири године савладавања (дате) језичке материје.

Зато би дијаграм кретања употребе и датих *пунозначних лексема* у дискурсу уџбеника са ознаком *Читанка I–IV* (у фреквенцијском попису: *читанка I–4 корпус*), заправо показивао кривуљу скоро идентичну са кривуљом кретања функционалних (граматичких) речи. То је зато што, како смо видели, фреквенцијски материјал корпуса предмета *српски језик* кроз четири посматрана разреда (*читанка I–4 корпус*), када се посматра укупна маса морфосинтаксичких речи (обе класе – функционалних и пунозначних лексема), показује скоро равну црту, уз само један (квалитативни) скок после II разреда. Наиме, очигледно је да је у прва два разреда дискурс уџбеника лексички (па тиме и у погледу броја морфосинтаксичких речи) за 50% мањи (I са 4.096 речи и II са 5.230 речи) у односу на дискурс уџбеника у наредна два разреда (III са 8.295 речи и IV са 12.629 речи).

Све овде наведено нас доводи до закључка да семантички доминантне пунозначне лексеме јесу управо оне које припадају наведеним граматичким категоријама, именичким, придевским и глаголским.

У уџбеничком дискурсу *природно-друштвене области*, с обзиром на велики број употребљених лексема (па тиме и морфосинтаксичких речи), као и у поступку праћења функционалних (граматичких) речи и њихове фреквенције у дискурсу уџбеника, и овде смо, такође, изабрали по неколико лексичких јединица из класе *именица*, *придева* и *глагола* да бисмо, руководећи се њиховим местом у дискурсу предмета, пратили њихову фреквенцију у уџбеничком дискурсу од нижег разреда ка вишем. При томе, за разматрање смо узели ограничен број јединица из лексике везане за (најширу) тематику основношколског предмета „Познавање природе и друштва”, и то ону из

семантичког поља у којем су (1) опште именице: *људи, човек, жена, дете, мајка, отац, мама, тата, породица, школа; живот, природа, земља, ваздух, вода*, као и (2) придеви: *жив, потребан, различит, школски*, и (3) глаголи: *живети, радити, имати, мислити* – као лексичке речи које се датим именицама (језички) атрибуирају, одређујући их по битним карактеристикама које се ученицима тумаче кроз наставу овога предмета.

По себи се разуме да се иста методологија посматрања може применити и на сва друга семантичка поља лексема у испитиваном уџбеничком дискурсу, било да су оне у дефиницијском дискурсу, било да су у примерима језика или у методичко-дидактичком апарату уз дефиниције.

И за ову (конкретну) дискурсну материју, дакле, можемо се приказати прегледну табелу каква је дата и за материју коју обрађују уџбеници матерњег језика, са кретањем учесталости изабраних лексема у уџбеницима за прва четири разреда основне школе:

Уџбеници за природу и друштво:

I разред	II разред	III разред	IV разред
Укупан број речи у читанкама			
1.694	2.354	3.553	4.609
именичке лексеме			
<i>људи</i> (34) <i>човек</i> (6) <i>жена</i> (ø) <i>дете</i> (36) <i>мајка</i> (1) <i>отац</i> (2) <i>мама</i> (1) <i>тата</i> (1) <i>породица</i> (20) <i>школа</i> (11) <i>живот</i> (2) <i>природа</i> (12) <i>земља</i> (2) <i>ваздух</i> (14) <i>вода</i> (37)	<i>људи</i> (9) <i>човек</i> (1) <i>жена</i> (1) <i>дете</i> (47) <i>мајка</i> (15) <i>отац</i> (15) <i>мама</i> (46) <i>тата</i> (4) <i>породица</i> (8) <i>школа</i> (26) <i>живот</i> (7) <i>природа</i> (7) <i>земља</i> (12) <i>ваздух</i> (2) <i>вода</i> (11)	<i>људи</i> (55) <i>човек</i> (30) <i>жена</i> (ø) <i>дете</i> (13) <i>мајка</i> (ø) <i>отац</i> (ø) <i>мама</i> (1) <i>тата</i> (ø) <i>породица</i> (1) <i>школа</i> (9) <i>живот</i> (21) <i>природа</i> (15) <i>земља</i> (6) <i>ваздух</i> (27) <i>вода</i> (90)	<i>људи</i> (48) <i>човек</i> (37) <i>жена</i> (1) <i>дете</i> (24) <i>мајка</i> (ø) <i>отац</i> (4) <i>мама</i> (ø) <i>тата</i> (ø) <i>породица</i> (4) <i>школа</i> (13) <i>живот</i> (31) <i>природа</i> (29) <i>земља</i> (33) <i>ваздух</i> (24) <i>вода</i> (81)

придевске лексеме			
жив (7) потребан (23) различит(12) школски (3)	жив (4) потребан (4) различит (1) школски (3)	жив (13) потребан (25) различит (33) школски (2)	жив ⁴ (16) потребан ⁵ (32) различит (41) школски (2)
глаголске лексеме			
живети(27) радити (51) имати(41) мислити ⁶ (31)	живети ⁷ (16) радити (12) имати (46) мислити ⁸ (18)	живети ⁹ (28) радити (35) имати (78) мислити ¹⁰ (4)	живети ¹¹ (28) радити (35) имати (78) мислити ¹² (4)

У овим уџбеницима укупан број реализација морфосинтаксичких речи (обеју дефинисаних класа – граматичких и пунозначних речи) – за разлику од констатованог за уџбенике предмета матерњег језика – показује лаганији, постепени пораст. У уџбенику за I разред тај број је 1.694, за II разред број је 2.354, у дискурсу уџбеника за III износи 3.553 и за IV разред – 4.609.

⁴ – са његовим граматичко-творбеним дериватом *нежив* у колокацији са именицама значењске класе *природа*, *граматички род* и сл. С тим у вези, в. односе које за колокацију дефинише, нпр. Р. Картер када каже: „*Collocation* is a term used to describe a group of words which occur repeatedly in a language. These patterns of co-occurrence can be *grammatical* in that they result primarily from syntactic dependencies or they can be *lexical* in that, although syntactic relationships are involved, the patterns result from the fact that in a given linguistic environment certain lexical items will co-occur”, да би затим навео конкретну илустрацију за енглески језик – колокације лексема: „For instance, we might expect *snow* to have a high probability of co-occurrence with *block*, *road*, *fall*, *winter*, *cold*, etc. but a low probability of co-occurrence with *cider*, *apple*, *dog*, etc. Ronald Carter, *Vocabulary – Applied Linguistic Perspectives, Second edition*, First published 1998, by Routledge, London EC4P 4EE – This edition published in the Taylor & Francis e-Library, 2002, 51–52).

⁵ – и његови граматичко-творбени деривати *најпотребнији*, *непотребан*.

⁶ Од чега гл. *мислити* са фреквенцијом 6 пута, и његови граматичко-творбени деривати *измислити*, *размишљати*, *размислити* – са фреквенцијом 25 пута.

⁷ – и његови граматичко-творбеним дериватом у једном случају: *преживети*.

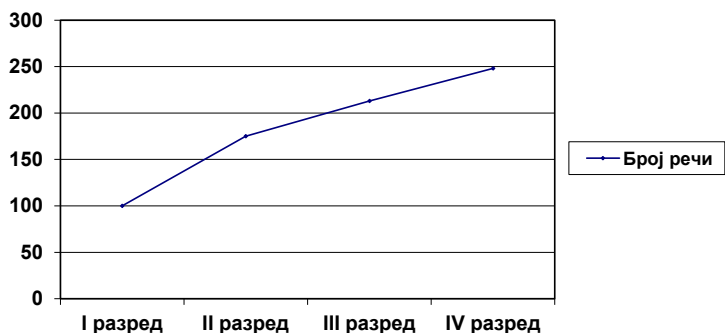
⁸ Од чега гл. *мислити* са фреквенцијом 6 пута, и његови граматичко-творбени деривати *измислити*, *размишљати*, *размислити* – са фреквенцијом 25 пута.

⁹ – и својим граматичко-творбеним дериватом у једном случају: *доживети*.

¹⁰ – и његови граматичко-творбени деривати *размислити*, *размишљати*.

¹¹ – и његов граматичко-творбеним дериват у једном случају: *доживети*.

¹² – и његови граматичко-творбени деривати *размислити*, *размишљати*.



Сл. 8: Дијаграм кретања укупног броја морфосинтаксичких речи у дискурсу уџбеника са ознаком *Природа и друштво I–IV* (у фреквенцијском попису: *свет око нас 1 корпус, свет око нас 2 корпус, природа и друштво 3 корпус, природа и друштво 4 корпус*):
 I разред – 1.694 речи; II разред – 2.354 речи; III разред – 3.553 речи;
 IV – разред 4.609 речи.

Као што се види, разлика у фреквенцији речи међу уџбеницима износи (просечно) око 1.200 статистичких јединица и равномерна је – пораст је по-степен, а не нагао. Логично је, такође, поновити да је и овде упадљива разлика у укупном броју употребљених морфосинтаксичких речи када се пореде (и са аспекта пунозначних речи) корпуси уџбеника са ознаком *Природа I–IV* и оних са ознаком *Читанка I–IV*. Овај први садржи скоро 65% мање речи, и то потврђује традиционалну констатацију о томе да су егзактне научне (и наставне) дисциплине са мање елоквенције, формализоване, друге хуманистичке у томе су слободније.

Упадљиво је да фреквенција појединих од њих од I до IV разреда и опада. Илустративан пример за то је фреквенција лексема *мајка, отац и мама, тата*. Оне се у анализираном узорку дискурса крећу тако да се од (наизменично) великог броја у I разреду (*мајка* – 1, *отац* – 15 и *мама* – 46, *тата* – 1) и у II (*мајка* – 15, *отац* – 15 и *мама* – 46, *тата* – 4), следећа два разреда њихова фреквенција спусти у III разреду *мајка* – 0, *отац* – 0; *мама* – 1, *тата* – 0, односно у IV разреду *мајка* – 0, *отац* – 4; *мама* – 0, *тата* – 0.

Друге од лексема постојано чувају линију своје фреквенције. За ово је илустративан пример употребе лексеме *дете*¹³, која има константно високу цифру фреквенције кроз уџбеничке дискурсе обе области (а) језичко-књижевне и (б) природно-друштвене:

¹³ – и њена суплетивна множина *деца*.

I разред	II разред	III разред	IV разред
(а) <i>дете</i> (44) (б) <i>дете</i> (36)	<i>дете</i> (14) <i>дете</i> (47)	<i>дете</i> (68) <i>дете</i> (13)	<i>дете</i> (132) <i>дете</i> (24)

У целини, међутим, линија фреквенције и *доминантних лексема* (и њихових морфосинтаксичких речи), као и свих које су њихов контекст – сасвим је уједначена, равна, без наглих скокова у периоду од четири године савладавања материје коју садрже ови уџбеници.

Фреквенцијски преглед речника једног дискурса може указати и на феномене који се садрже од веће или мање *способности колоцирања* појединих лексема. Под овим се подразумева чињеница, констатована у лингвистици, да дата, конкретна реч има одређену – већу или мању способност да „оперише у лексичком окружењу других речи”, док нека друга реч ту способност има у мањој мери, или је чак и нема.¹⁴

За илустрацију феномена тематске адаптације употребе доминантних пунозначних лексема у целини анализираног уџбеничког дискурса, рецимо, баш је илустративно кретање фреквенције именица из семантичког поља „родитељи : потомство” какве су именице *мајка* – *отац*, *мама* – *тата*, с једне стране, и *дете* (*деца*) с друге стране. Из датог фреквенцијског тока, наиме, видљиво је да са сваким степеном напретка школског градива у оба предмета (од I до IV разреда) лексеме *мајка*, *отац* и *мама*, *тата* имају све нижу цифру учесталости у уџбеничком дискурсу као целини, а логично – и у окружењу других лексема. При томе, у односу *мајка* – *отац* (*мама* – *тата*) несумњиво већу учесталост кроз период од четири године наставе има именица *мајка* (*мама*).

Са друге стране, фреквенцијски ток именице *дете* (*деца*) у укупном уџбеничком дискурсу оба предмета, па тиме логично и у окружењу других лексема (дакле – налазећи се у колокацији с њима) показује стабилну учесталост. При томе, у уџбеницима (а) матерњег језика, са порастом обима текста и увођењем примера књижевних текстова, фреквенција именице *дете* (*деца*), као главну тенденцију, показује бројни раст (I: 44 – II: 14 – III: 68 – IV: 132). У уџбеницима (б) из области природе и друштва, међутим, види се да именица *дете* (*деца*) углавном смањује своју учесталост са сваком годином (разредом). Основни узрок оваквог кретања сигурно је везан и за природу сваког конкретног текста, дакле – за лексичко окружење, које иде од лексике свакодневних односа деце и њихове породичне, школске, природне и друштвене средине.

Није, зато, необично што, на пример, у уџбенику матерњег језика за I разред највећу фреквенцију имају – именице *прича* (51), *зец* (47), *песма* (43),

¹⁴ „Collocability describes the company a word keeps; that is, single words operate in a lexical environment of other words. We often know a word like *lean* because it collocates with *meat*. And some words can only be differentiated by citing their normal collocability range; for example, strong tea – *powerful tea...” (Ronald Carter, *Vocabulary – Applied Linguistic Perspectives*, Second edition, First published 1998, by Routledge, London EC4P 4EE, 2002, 38).

дете (44), лопта (41), школа (36), реч (35), мама (32), лав (30), птица (29), сунце (21) – то је дискурс са лексиком односа мајке и детета, са лексиком бајке и прича за децу, са лексиком игре. Ово је сасвим у складу са у лингвистици дефинисаним чињеницама да ће „у датом језичком окружењу одређене лексичке јединице коегзистирати”, а да је та коегзистенција основана на *потенцијалу значења* те лексичке јединице, који јој то егзистирање омогућује. Наравно, последица тога односа је и чињеница да (одређена) лексема *X* лакше колоцира са лексемама *A, B...* него са лексемама *C, D...*, као и чињеница да се неке лексеме и не могу наћи једна уз другу. Све то – зависно од *значањских (семантичких) поља лексема*, које би требало да се нађу у конкретном дискурсу, са одређеним предметом језичке комуникације, укључујући и комуникацију наставник / аутор уџбеника – ученик.¹⁵

Закључак

Као закључак можемо извести да се фреквенцијске кривуље могу категоризовати истим текстлингвичким методом којим се категоризује и степен информативности неког текста, који може бити: *низак, висок* и *највиши*, тако да сваки ниво заправо одговара аудиторијуму коме је упућен, што заправо чини онај који се том аудиторијуму обраћа пажљивим одабиром адекватне лексике. Ово нам отвара могућности даљег истраживања и примене исте методологије фреквенцијске категоризације пунозначних лексема, што пружа могућност писцима уџбеника да ефикасније користе лексеме које могу колоцирати са другима, како на семантичном тако и на синтаксичком и текстлингвистичком нивоу. А сматрамо да се оваква метода категоризације може применити не само у анализи уџбеничког дискурса, већ и у анализи осталих типова дискурса и текста.

Извори

Читанка I: S. Jovanović, *Zlatna jabuka, Čitanka za 1. razred osnovne škole*, Beograd: Zavod za udžbenike, 2007.

Читанка II: S. Jovanović, *Zlatna jabuka, Čitanka za 2. razred osnovne škole*, Beograd: Zavod za udžbenike, 2006.

¹⁵ „Each field usually has a specialized, topic-related vocabulary...” – каже Р. Картер, илуструјући то дискурсом цивилног ваздухопловства: „Air transport, for example, constitutes a field with special items such as *flight, take-off, runway, check in*, etc. Each field usually has a specialized, topic-related vocabulary (for preliminary discussion see Section 2.1.9) some items of which may turn up in other fields with different meanings (and often in different grammatical collocations – see Section 3.5): for example, *wings* in air transport and in theatrical parlance; *lock* in the field of river or canal transport and in the language of security” (Ronald Carter, *Vocabulary – Applied Linguistic Perspectives, Second edition*, First published 1998, by Routledge, London EC4P 4EE, 2002, 53–54).

- Читанка III:** V. Milatović, *Čitanka za 3. razred razred osnovne škole*, Beograd: Zavod za udžbenike, 2006.
- Читанка IV:** Z. Орачић-Nikolić, D. Pantović, *Priča bez kraja, Čitanka za četvrti razred osnovne škole*, Beograd: Zavod za udžbenike, 2005.
- Природа I:** Lj. Novković, B. Trebješanin, B. Gačanović, *Svet oko nas, Udžbenik za prvi razred osnovne škole*, Beograd: Zavod za udžbenike, 2006.
- Природа II:** Lj. Novković, B. Trebješanin, B. Gačanović, *Svet oko nas, Udžbenik za drugi razred osnovne škole*, Beograd: Zavod za udžbenike, 2006.
- Природа III:** Lj. Novković, B. Trebješanin, B. Gačanović, *Priroda i društvo, Udžbenik za treći razred osnovne škole*, Beograd: Zavod za udžbenike, 2006.
- Природа IV:** Lj. Novković, B. Trebješanin, B. Gačanović, *Priroda i društvo, Udžbenik za četvrti razred osnovne škole*, Beograd: Zavod za udžbenike, 2007.

ЛИТЕРАТУРА

- Картер 2002:** R. Carter, *Vocabulary – Applied Linguistic Perspectives*, Second edition, First published 1998, by London: Routledge, This edition published in the Taylor & Francis e-Library.
- Половина 1999:** V. Polovina, *Semantika i tekstlingvistika*, Beograd: Filološki fakultet.
- Сантос Перейра 2004:** L. A. Santos Pereira, The use of concordancing in the teaching of Portuguese, in: J. McH Sinclair (ed.), *How to Use Corpora in Language Teaching*, Amsterdam – Philadelphia: John Benjamins Publishing Company, 109–119.
- Синклер 2004:** J. McH Sinclair (ed.), *How to Use Corpora in Language Teaching*, Amsterdam – Philadelphia: John Benjamins Publishing Company.
- Станојчић/Поповић 2010:** Ž. Stanojčić, Lj. Popović, *Gramatika srpskog jezika*, Beograd: Zavod za udžbenike.
- Станојчић 2024:** С. Станојчић, Методолошке импликације анализе потенцијала колокације и конкорданце речи у типичном учбеничком дискурсу, *Наслеђе*, 57, 287–300
- Требјешанин/Лазаревић 2001:** B. Trebješanin, D. Lazarević, *Savremeni osnovnoškolski udžbenik – teorijsko metodološke osnove*, Beograd: Zavod za udžbenike.
- Харисон 2001:** A. G. Harrison, How do Teachers and Textbook Writers Model Scientific Ideas for Students?, *Research in Science Education*, 31(3), Amsterdam, 401–435.

Slavko Ž. Stanojčić

CATEGORIZATION OF LEXICAL COLLOCATIONS AND FREQUENCY
CATEGORIZATION OF FULL WORDS IN TYPICAL TEXTBOOK
DISCOURSE

Summary

The paper brings the results of the author's research into the corpus of textbook discourse, both individually – by years of study of the studied subjects, as well as collectively – in terms of frequency. The analysis refers to a large number of lexical (full) words of the following categories: nouns, verbs, and adjectives. In addition to the statistically given corpus as a whole, a number of indicative units are shown – frequency-wise as well – in sequential order, the starting line of which is the first year of studying the given subject. It is precisely the frequency of the given lexemes that enabled the introduction of the concept (and term) of categorization of semantically dominant lexemes into the research of the concrete discourse that forms the corpus of this paper. Also, the research provides an insight into the elements crucial for further displaying of the collocation system in a given discourse, and the paper proposes possible solutions for the explanation of the dominance of full lexemes based on contemporary text-linguistic theories.

Keywords: Lexical collocations, textbook discourse, frequency categorization, full words, corpus analysis.