

ISSN 0350–185x

UDK 808

СРПСКА АКАДЕМИЈА НАУКА И УМЕТНОСТИ  
И  
ИНСТИТУТ ЗА СРПСКИ ЈЕЗИК САНУ

## ЈУЖНОСЛОВЕНСКИ ФИЛОЛОГ

КЊ. LXXI, СВ. 1–2

Уређивачки одбор:

*др Јасна Влајић-Појовић, др Даринка Горџан-Премк, др Рајна Драгићевић,  
др Најда Иванова, др Александар Лома, др Алина Ј. Маслова,  
др Софија Милорадовић, др Мирослав Николић, др Слободан Павловић,  
др Предраг Пипер, др Слободан Реметић, др Живојин Сјанојчић,  
др Срејто Танасић, др Анађолиј Турилов, др Зузана Тојолињска,  
др Викџор Фридман*

Главни уредник

Предраг Пипер

БЕОГРАД

2015

ISSN 0350-185x, LXXI, св. 1 (2015)

UDK: 811.163.41'367

ID: 214484492

DOI: 10.2298/JF11470009T

Примљено: 7. јула 2014.

Прихваћено: 26. јануара 2015.

Оригинални научни рад

РАДОСЛАВА М. ТРНАВАЦ\*

Универзитет у Ванкуверу „Сајмон Фрејзер“

Програм за когнитивну науку

## ТЕОРИЈА ЦЕНТРИРАЊА (CENTERING THEORY), ТИПОВИ ПРЕЛАЗА И МАРКИРАЊЕ РЕФЕРЕНТА У ПИСАНОМ КОРПУСУ СРПСКОГ ЈЕЗИКА

У раду је примењена теорија центрирања (WALKER, JOSHI и др. 1998) на писани корпус српског језика. Циљ рада је да се одреди коре-лација између типова прелаза и анафорских облика који обележавају кохерентни текст, што може да послужи као почетак формализације њиховог односа у апликацијама из рачунарске лингвистике. Анализа корпуса потврђује хипотезу да постоји паралела између минималних анафорских облика и типа прелаза НАСТАВАК (најкохерентнији прелаз у тексту), али правило центрирања које одређује распоред и фреквентност прелаза (правило „уређења“) није нашло потврду у нашој анализи.

*Кључне речи:* теорија центрирања; типови прелаза; Сb; Ср; кохерентност текста; сегментација текста; клауза.

### I. Увод

Током последњих неколико деценија обележавање референта у дискурсу је предмет бројних истраживања како у лингвистици, тако и у другим, сродним, научним дисциплинама. Велики број радова из те области је настао у оквиру теорије центрирања (Centering Theory) (GROSZ & SIDNER 1986; GORDON, GROSZ и др. 1993; GROSZ, JOSHI и др. 1983, 1995; BRENNAN, FRIEDMAN и др. 1987; WALKER, JOSHI и др. 1998). Ова теорија, чије је порекло у рачунарској лингвистици, хеуристички одређује најактивиранији референт у датом исказу<sup>1</sup>, те

---

\* rtmavac@sfu.ca

<sup>1</sup> Исказ се у овом раду поистовећује са финитном клаузом, о чему ће бити више речи у одељку о сегментацији дискурса у теорији центрирања.

вероватноћу с којом ће се исти референт поновити у наредном исказу. Центрирање указује на кохерентност дискурса, која се, према датог теорији, заснива на начину на који се ентитети уводе и „опстају“ у одређеном тексту, као и на значај тих ентитета на локалном нивоу, у оквиру клаузе или неколико суседних клауза (видети POESIO, STEVENSON и др. 2004).

Главна теза теорије центрирања је да су сегменти дискурса који се састоје од два или више исказа са истим ентитетима „кохерентнији“ од два или више исказа у којима се помињу различити ентитети (CHAFFE 1976; KINTSCH AND VAN DIJK 1978; POESIO, STEVENSON и др. 2004, између осталих). Веза између два узастопна, кохерентна исказа се најчешће остварује кроз заједнички топик<sup>2</sup>, тзв. „центар пажње“ (center of attention) који се у терминологији енглеског језика назива *Cb*, или „центар који гледа уназад“ (backward-looking center). Следећи одломак илуструје један од начина на који се одржава континуитет *Cb*-а у дискурсу:

- (1) а. *Јуче сам разговарао са Маријом.*  
 б. *Била је у кувини. (∅=Марија)*  
 в. *Срела је Сању у књижари. (∅=Марија)*  
 г. *Добила је од ње њоклон. (∅=Марија)*

Према теорији центрирања (WALKER, JOSHI и др. 1998), као и другим сродним теоријама<sup>3</sup>, нулти облик у 1(г) се односи на референт *Марија*, који је субјекат у клаузи 1 (в), те је истакнутији члан те клаузе него *Сања*, референт који заузима објекатску позицију. Из овога следи да би *Марија* требало да представља „центар пажње“ (*Cb*), односно топик исказа у 1 (г), пошто је облик поновљен у тој клаузи из претходне клаузе. На основу другог правила центрирања, о коме ће бити речи у следећем одељку, најредукованији анафорски елемент, било да је то нулти облик<sup>4</sup> или заменица, мора предста-

<sup>2</sup> У овом раду прихваћена је дефиниција топика коју је понудила Гундел (GUNDEL 1988:210): „Ентитет Е је топик у реченици S ако, користећи S, говорник намерава да повећа ниво знања слушаоца о Е, или да затражи информацију о Е, или да на неки други начин примора слушаоца да делује у вези са Е. Предикација Р је коментар у реченици S ако, изразивши реченицу S, говорник намерава да окарактерише Р у односу на топик Е.“

<sup>3</sup> Видети, на пример, теорију доступности (ARIEL 1990) и теорију датости (GUNDEL, HEDBERG и др. 1993).

<sup>4</sup> Нулти облик се употребљава у тзв. *pro-drop* језицима, у којима субјекат не мора да буде исказан заменицом, уколико је већ познат из контекста. Српски језик припада групи *pro-drop* језика.

вљати *Cb*, односно бити топик датог исказа (HEDBERG 2010). У случају (г) исказа то правило је испоштовано, те је *Марија* представљена нултим обликом и кохерентност је очувана.

У овом раду је теорија центрирања примењена на писани корпус српског језика. Циљ рада је да се установи корелација између анафорских облика топика (*Cb*) и типова различитих прелаза који постоје међу суседним исказима/клаузама у дискурсу.

У наредном одељку ће детаљније бити представљена теорија центрирања, као и типови прелаза који одређују степен кохерентности текста. Одељак 3 описује начин на који је вршена сегментација дискурса у постојећем корпусу у овом раду. Одељак 4 се бави описом корпуса и резултатима анализе корпуса. Одељак 5 даје смернице за будућа истраживања у домену односа између типова прелаза у теорији центрирања и облика *Cb*-а који их прате у српском језику.

## II. Теорија центрирања и типови прелаза у дискурсу

Теорија центрирања је до сада тестирана на материјалу бројних језика, укључујући енглески, немачки, холандски, јидиш, италијански, шпански, француски, грчки и пољски<sup>5</sup> (HEDBERG 2010). Према теорији центрирања, сваки ентитет у дискурсу је члан групе, односно листе „центара који гледају напред“ (forward-looking centers)<sup>6</sup>. Чланови ове листе су ранжирани према степену важности, што је условљено њиховим граматичким статусом у реченици у великом броју језика<sup>7</sup>. „Ранг-листа“ чланова се разликује од језика до језика, али у великом броју случајева, субјекатска позиција претходи објекатској по степену важности. С обзиром на то да кохерентни текстови имају тенденцију да уводе ентитете према одређеном редоследу, најважнији члан листе је „центар пажње“, односно *Cb* (видети одељак I). Он је „најбоље ранжирани“ члан из претходног исказа који је реализован у текућем исказу. Он је такође и топик претходног исказа, те се веома често поклапа са субјектом тог исказа. Према првом правилу теорије центрирања, сваки исказ има максимално један *Cb*, односно један топик. Најактуелнији члан текућег исказа

<sup>5</sup> Комплетан број језика на које је примењена теорија центрирања је наведен у ХЕДБЕРГ (2010).

<sup>6</sup> У енглеској терминологији ова листа се назива *Cf list*.

<sup>7</sup> У неким језицима се граматички статус комбинује са редом речи у клаузи (видети POESIO, STEVENSON и др. 2004).

се назива „преференцијални“ центар или *Cp* (Preferred Center). *Cp* у актуелном исказу може али и не мора да се поклапа са *Cb*-ом, топиком из претходног исказа.

У оквиру теорије центрирања постоје четири основна типа прелаза међу клаузама (исказима), који су условљени понашањем два претходно поменута центра. Типови прелаза су приказани у табели 1.

Табела 1. Типови прелаза међу исказима према теорији центрирања

	$Cb(U_i)=Cb(U_{i-1})$ или $Cb(U_{i-1})=0$	$Cb(U_i)\neq Cb(U_{i-1})$
$Cb(U_i)=$ <i>Cp</i> ( <i>U<sub>i</sub></i> )	НАСТАВАК	ЛАГАН ПРЕЛАЗ
$Cb(U_i)\neq$ <i>Cp</i> ( <i>U<sub>i</sub></i> )	ЗАДРЖАВАЊЕ	ТЕЖАК ПРЕЛАЗ

Прелази се анализирају на материјалу две суседне клаузе, у зависности да ли оне деле исти *Cb* и да ли се *Cb* у другом исказу поклапа са *Cp*-ом у том исказу. Наредни одломак из Политике онлајн за 5. октобар 2012. године илуструје тип прелаза под називом НАСТАВАК (CONTINUE):

(2) (а) *Према судским документима, Фишченко се родио у Казахстану, (б) завршио Лењинградски физичко-технолошки институт, (в) 1994. године се преселио у САД, (г) а 2003. добио америчко држављанство, (д) али је задржао и руско (е) и чистио је његово преко океана.*

Наведена реченица је подељена на пет финитних клауза које имају исти топик, исти субјекат, те самим тим *Cb* из сваке претходне клаузе представља *Cp* и *Cb* у свакој следећој клаузи. Дакле, долази до поклапања центара *Cp* и *Cb*<sup>8</sup>, као и до понављања истог топика (*Cb*) у свим клаузама. Ова врста прелаза илуструје највиши степен локалне<sup>9</sup> кохерентности дискурса.

<sup>8</sup> Прва клауза у тексту или новом сегменту припада нултом прелазу, о коме ће бити више речи на крају овог одељка.

<sup>9</sup> За разлику од релацијске кохерентности која обухвата веће сегменте (чак и по неколико реченица) (видети MANN & THOMPSON 1988), сегменти референтне кохерентности најчешће делују на локалном нивоу, у домену двеју суседних клауза.

Следећи прелаз је ЗАДРЖАВАЊЕ (RETAIN). Он одражава нешто мањи степен кохерентности од првог прелаза. По дефиницији, у оквиру тог прелаза С<sub>b</sub> у актуелном исказу се поклапа са С<sub>b</sub>-ом у претходном исказу. Међутим, за разлику од претходног примера, не постоји поклапање између С<sub>p</sub>-а и С<sub>b</sub>-а у актуелној клаузи:

- (3) **(a)** У документима америчког министарства правосуђа (...) <sup>10</sup>, наводи се **(б)** да је забрањена микроелектроника испоручивана обавештајним и војним усљановама у Русији **(в)** и да је 11 осумњичених испоручивало робу не само у Русију, него и команијама из Тексаса и неким другим сируктурама. (Политика онлајн, 5. октобар 2012)

Цитирани одломак је подељен на три клаузе. У (в) клаузи се референт *роба* анализира као С<sub>b</sub>, односно као топик, који се поклапа са С<sub>b</sub>-ом из претходне клаузе (*микроелектроника*)<sup>11</sup> док 11 осумњичених представља С<sub>p</sub> у (в) клаузи. Дакле, не постоји поклапање између С<sub>p</sub>-а и С<sub>b</sub>-а у клаузи (в) док се С<sub>b</sub> из (б) и (в) клаузе поклапају.

ЛАГАН ПРЕЛАЗ (SMOOTH SHIFT) илуструје ситуацију по којој се С<sub>b</sub> и С<sub>p</sub> у другом исказу поклапају док не постоји подударност између С<sub>b</sub>-а из првог и другог исказа.

- (4) **(a)** Америчке власници испичу **(б)** да је оишужница подигнута у контексту кампање за модернизацију руске војске **(в)** и да је фирма „Арк електроникс“ нейосредно зависила од финансирања руске војске последњих неколико година. (Политика онлајн, 5. октобар 2012)

У примеру (4) оишужница представља С<sub>b</sub> у (б) клаузи док је Арк електроникс С<sub>b</sub> у клаузи (в) и поклапа се са С<sub>p</sub>-ом у истој клаузи.

Последњи традиционални прелаз у теорији центрирања је ТЕЖАК ПРЕЛАЗ. У оквиру овог прелаза не постоји поклапање између С<sub>b</sub>-а у две клаузе, а исто тако С<sub>b</sub> у другој клаузи није идентичан са С<sub>p</sub>-ом у истој клаузи.

<sup>10</sup> У овом делу реченице изостављена је односна клауза. У следећем одељку следи објашњење како се анализирају односне и друге зависне реченице у смислу успостављања новог или задржавања старог топика (С<sub>b</sub>) у контексту независне реченице.

<sup>11</sup> Топици *микроелектроника* и *роба* указују на исти референт и повезани су импликатуром (видети MATSUI 1993).

(5) **(а)** *Забележени су и мањи инцидентии* **(б)** *иако да инсџекџоре*  
*уводе кроз џајне улазе.* (Политика онлајн, 5. октобар 2012)

У овом примеру у другој клаузи, Ср је изражен нултим елементом, Сб је представљен речју *инсџекџор*, која је била топик у претходној реченици, док су Сб и Ср реализовани кроз именицу *инцидентии* у првој клаузи.

Анализа овога рада такође укључује НУЛТИ прелаз<sup>12</sup> који се налази на почетку текста или новог сегмента када долази до промене претходног топика, што може да се поклопи са започињањем новог параграфа.

У оквиру теорије центрирања постоје три основна правила. Прво правило одређује да сваки исказ, осим првог у тексту, мора имати један Сб<sup>13</sup>. Друго правило је „заменичко правило“<sup>14</sup>, по коме у сваком исказу у коме постоји најкраћи анафорски елемент (било да је то нулти облик<sup>15</sup> или заменица<sup>16</sup>), он мора бити у функцији Сб-а (видети TURAN 1998). Треће правило теорије центрирања (правило „уређења“<sup>17</sup>) односи се на хијерархију односа међу прелазима. Да би се обезбедила што већа кохерентност текста тип прелаза НАСТАВАК биће у предности у односу на ЗАДРЖАВАЊЕ, који је опет у предности у односу на ЛАГАН прелаз, док је ГРУБИ прелаз прелаз који указује на најмањи степен кохерентности. Дакле, према овом правилу, задржавање истог топика у тексту представља израз веће кохерентности него промена топика. Такође, већа кохерентност се постиже уколико се топик понавља у субјекатској позицији а не у објекатској позицији. Уколико говорник/писац покушава да

<sup>12</sup> У Poesio, Stevenson et al. (2004) НУЛТИ прелаз (ZERO transition) се користи у случајевима када први исказ садржи Сб, а други га нема. Такође, ти аутори разликују још две врсте прелаза: када ни у првом ни у другом сегменту не постоји Сб (NULL transition), као и случај када први сегмент нема Сб, а други га има (Establishment). У овом раду постоји само једна врста НУЛТОГ прелаза када први исказ текста или параграфа нема Сб, а други га има. За све остале случајеве користи се један од четири традиционална прелаза наведена горе у тексту.

<sup>13</sup> Ово је „јака“ верзија тога правила. Према неким другим верзијама, поједини искази не морају имати Сб нити топик (видети Poesio, Stevenson et al. 2004).

<sup>14</sup> Pronoun Rule (енгл.).

<sup>15</sup> У *pro-drop* језицима, као што је то српски, уколико у изразу постоји нулти елемент, онда он мора бити Сб у поменутом сегменту.

<sup>16</sup> На пример, у енглеском језику.

<sup>17</sup> Ordering Rule (енгл.).

максимизира кохерентност дискурса, онда ће у том дискурсу нај-фреквентнији бити прелаз НАСТАВАК док ће нешто мање фреквентан бити прелаз ЗАДРЖАВАЊЕ. ЛАГАН и ГРУБИ прелаз према овом правилу биће најмање фреквентни прелаз.

Следеће поглавље описује сегментацију текста у теорији центрирања.

### III. Сегментација текста у теорији центрирања

У анализи корпуса у овом раду сегментација текста је вршена на основу хипотезе Камејама (КАМЕЈАМА 1998) по коме је основна јединица у теорији центрирања финитна клауза. Камејама предлаже два модела центрирања у оквиру реченице: (1) секвентно центрирање, према коме се у сваком тренутку процесира само једна клауза, те постоји само један центар у датом тренутку и (2) хијерархијско центрирање, према коме се центрирање одвија паралелно и по коме је могуће имати више центара (у главној и у зависној реченици). Секвентно центрирање се примењује у овом раду на свим главним и зависним клаузама док се хијерархијско примењује само у случајевима управног говора, изричних клауза, као и на примеру односних клауза. Хијерархијско центрирање подразумева да клауза у којој се оно одвија има један центар, али се тај центар не узима у обзир када се одређује прелаз између две клаузе (нпр. односне и наредне клаузе), већ се за наредну клаузу узима у обзир центар из прве претходне главне клаузе. Ова појава је илустрована у следећем примеру из Политике онлајн:

- (6) **(а)** Фишенка сумњиче **(б)** да је „деловао на територији САД као нерегистровани агенциј руске владе, незаконито кулјујући високојтехнолошку микроелектронику у име руске владе“, **(в)** и да је радио за руске власници, **(г)** о чему сведоче подаци прислушкивањем телефона, као и његова електронска пошта. **(д)** Према информацијама ТВ канала АБЦ, крајем 2011. године осумњичени су покушали да замејну праг, уништавајући документе и друге материјале **(ђ)** који су сведочили о њиховој вези са руским власницима. (Политика онлајн, 5. октобар 2012)

У наведеном примеру прелаз који обележава сегмент (д) се везује за главну клаузу у (а) а не за односну клаузу у (г) која не може да утиче на промену центра у наредној клаузи. У случају (д) клаузе



ради се о прелазу НАСТАВАК јер су Ср и Сб осумњичени у (д) клаузи идентични док се Сб у (а) клаузи *Фишенко* налази у односу хипонимије према Сб-у осумњичени<sup>18</sup> у (д) клаузи.

Слично односним реченицама, у реченицама са управним говором центар који је део управног говора се не узима као релевантан фактор за утврђивање центра у наредној клаузи, већ је то центар који долази из претходне главне клаузе:

(7) (а) *Турска је поштом пошврдила (б) да је преко УН добила извињење од Сирије због прекограничног бомбардовања турског града на југоистоку земље (в) у коме је у ушорак сипрадало иеи цивила, међу којима мајка и троје деце. (г) „Они су саоштили (д) да се нишиа слично неће поновити, (ђ) ишо је добро“, (е) изјавио је заменик турског премијера Бешир Ајдалај прецизирајући (ж) да је извињење сиигло после посредовања УН.* (Политика онлајн, 5. октобар 2012)

За прелаз у клаузи (е) је важан Сб из главне клаузе (а) након које постоје изрична и односна клауза, као и управни говор који не могу да ажурирају центре у клаузама након себе, те не могу ни да утичу на врсту прелазу у тим клаузама. Када је реч о изричним клаузама са глаголима који имају епистемично и евалуативно значење, оне припадају истој јединици центрирања као и њихова главна клауза (видети THOMPSON 2002: 136, NADIC-ZABALA & TABOADA 2006). Такви глаголи („верујем“, „мислим“, „претпостављам“, итд.) обично се налазе уз субјекат, заменицу/нулти облик у првом лицу, и у појединим су случајевима део парентезе. У корпусу овог рада нису нађени овакви примери јер се ради о корпусу базираном на спољнополитичким вестима где је елемент језичке субјективности минимално заступљен.

Један од најважнијих аспеката теорије центрирања је рангирање чланова по степену важности у датом исказу што се врши да би се одредили Ср и Сб за дати исказ. У анализи корпуса у овом раду је испоштован принцип важности граматичких односа, по коме су

<sup>18</sup> У (д) клаузи се ради о индиректној реализацији Сб-а из клаузе у (а). Између два Сб-а постоји однос укључивања – осумњичени укључује *Фишенка* (видети о Сб-у као о индиректној реализацији ентитета у дискурсу у TABOADA 2008, што је базирано на теорији лексичке кохезије у теорији HALLIDAY AND HASAN 1976). Индиректне реализације ентитета формиране између два Сб-а у суседним клаузама у дискурсу подразумевају односе типа „скуп – члан скупа“, „целина – део целине“, као и однос синонимије између два Сб-а.

субјекти у хијерархијској предности у односу на објекте, без обзира на то да ли су исказани пуном заменицом или нултим обликом. Највишу позицију у тој хијерархији представљају агенти тзв. емпатичних или „психолошких“ глагола (HADIC-ZABALA & TABOADA 2006) у реченицама типа „чини ми се“, „мислим“, и сл. који не морају имати позицију правог субјекта. У корпусу овог рада примера са „емпатичним“ глаголима нема због објективности текста. Међутим, с друге стране, јављају се у већем броју примери са безличном реченицама<sup>19</sup>. Следећи Туранову (TURAN 1995) и ДиЕугенијеву (DIEUGENIO 1998), који су предложили да се заменице у безличној употреби у турском и немачком језику налазе на дну листе по степену важности или да се чак сматрају неререференцијалним елементима, те да се искључе из листе кандидата за С<sub>b</sub> или С<sub>r</sub>, аутор овог рада се одлучио за другу варијанту, и није их анализирао као потенцијалне кандидате на место С<sub>b</sub>-а или С<sub>r</sub>-а<sup>20</sup>.

Такође, следећи ДиЕугенијеву (1998), присвојне заменице и именице које су у склопу исте именичке фразе у овом раду су анализиране на следећи начин: уколико именица означава нешто неживо, онда је референт на који се односи присвојна заменица рангиран на вишој позицији него именица док, уколико именица означава нешто живо, она носи хијерархијску предност у односу на присвојну заменицу:

- (8) *(а) Парламент̄ у Анкари аӯторизовао је даље војне акције у случају да се сличне ситуације понове, (б) док је помоћник турског премијера изјавио (в) да његова земља не намерава (г) да објави рат̄ у Сирији.*

У случају реченице у (8) реч „његова“ и њен референт „помоћник турског премијера“ су рангирани на вишој позицији него „земља“ која представља неживо, те је С<sub>b</sub> у (в) клаузи исти као и у (б) клаузи – „помоћник турског премијера“.

Као последње обележје сегментације која је спроведена у корпусу овог рада навешћу пример односних заменица. Следећи Забалу & Табоаду (2006), односне заменице не потпадају увек под „заменичко

<sup>19</sup> У примеру (6) у клаузи (а) нулти облик безличне заменице трећег лица множине није представљао кандидата на место С<sub>r</sub>-а дате клаузе.

<sup>20</sup> Безличне заменице у оквиру парентеза типа „како се наводи“ такође нису разматране као потенцијални кандидати на место С<sub>r</sub>-а и С<sub>b</sub>-а.

правило“ према коме најкраћи анафорски облик у сваком изразу мора да буде Св. На пример, илустрација тога би била ситуација када је односна заменица у позицији директног објекта, те се субјекат односне клаузе рангира на вишој позицији, чак и да је питању именица, иако се у истој клаузи јавља односна заменица као редукованији анафорски облик.

(9) **(a)** У документима америчког министарства правосуђа, **(б)** које преносе руски медији... (Политика онлајн, 5. октобар 2012)

У клаузи (б) Св је именичка фраза „руски медији“, а не односна заменица „које“, која представља референт именичке фразе „документа америчког министарства правосуђа“.

Следеће поглавље представља резултате корпусне анализе у оквиру које је примењена теорија центрирања.

#### IV. Анализа корпуса

Корпус овог рада представља 278 клауза (јединица центрирања) сачињених од спољнополитичких вести Политике онлајн за 5. октобар 2012. године. Аутор рада је извршио три пута анотацију датог корпуса у посебним временским интервалима. Поред сегментације текста на основне јединице центрирања, анализа сваке клаузе се базира на следећих пет параметара:

- листа ранжираних ентитета у датој клаузи
- „преференцијални“ центар у постојећој клаузи (Ср)
- центар „пажње“ у датој клаузи који је поновљен из претходне клаузе (Сб)
- тип прелаза
- облик реализованог Св-а.

У следећој табели су приказани резултати анализе:

Табела 2. Облици реализације Сб-а и типови прелаза између две клаузе (изражени кроз број примера)

<i>Тип прелаза / Облик Сб-а</i>	<b>Наставак</b>	<b>Задржавање</b>	<b>Лаган прелаз</b>	<b>Тежак прелаз</b>
Нулти облик	53		10	
Лична заменица	2	1	3	
Присвојна заменица	1			
Именица	19	4	127	2
Односна заменица	11	1	13	
Заменички прилог			1	
Одрична заменица			1	
Показна заменица	3		1	
<b>Укупан број облика</b>	<b>89</b>	<b>6</b>	<b>156</b>	<b>2</b>

Табела 2 приказује различите облике Сб-а у односу на различите облике прелаза између две клаузе. Резултати анализе показују да постоји нарушавање трећег правила центрирања, према коме се у хијерархији односа између прелаза предност даје НАСТАВКУ у односу на ЗАДРЖАВАЊЕ, као и у односу на ЛАГАН И ТЕЖАК прелаз (BRENNAN, FRIEDMAN et al. 1987). Што је кохерентност текста већа, постоји јача тенденција да се чува исти топик кроз клаузе, као и да се слушацац не оптерећује додатним закључивањем у односу на постојећи текст<sup>21</sup>. Табела 2 показује да у овом корпусу има највише ЛАГАНОГ ПРЕЛАЗА<sup>22</sup> – више од НАСТАВКА – за шта, по нашем мишљењу, постоји неколико разлога<sup>23</sup>:

1. У анализи корпуса је коришћен НУЛТИ ПРЕЛАЗ за сегменте којима нови текст (вест у овом случају) или параграф започи-

<sup>21</sup> Према алтернативној верзији трећег правила центрирања кохерентност у тексту је већа уколико постоји већа фреквентност низова прелаза НАСТАВАК–НАСТАВАК у односу на низове ЛАГАН ИЛИ ТЕЖАК прелаз. У корпусу овог рада ова верзија трећег правила такође није испоштована (постоји 27 низова ЛАГАНОГ прелаза и 15 низова који маркују НАСТАВАК) (видети GROSZ, JOSHI и др.).

<sup>22</sup> Погледати о сличним резултатима корпусне анализе у POESIO (2008).

<sup>23</sup> При овоме треба имати у виду да теорија центрирања не „пропагира апсолутну истину о језику“ (POESIO 2008) већ да хеуристички указује на правила која омогућују лакше процесирање дискурса.

- ње. Волкер, Ида и др. (WALKER, IDA ET AL. 1994) анализирају ове сегменте у оквиру прелаза НАСТАВАК, тврдећи да први сегмент у тексту има Сб, али да тај Сб првобитно није специфичан и да се он одређује тек након што се процесира други сегмент. Дакле, број НУЛТИХ ПРЕЛАЗА (25<sup>24</sup> случајева у нашем корпусу или 8,99%) одвојен је од прелаза НАСТАВАК.
2. У корпусу овог рада постоји велики број изричних (24,5%) и односних клауза (18,35%) за које се одређује Сб, али њихов Сб нема утицаја на врсту прелаза који се формира при процесирању следеће клаузе. Следећа клауза заснива прелаз са претходном главном/зависном клаузом (која није изрична, односна или није управни говор). У таквој констелацији текста постоји услов за велики број промена топика, односно увођења нових топика<sup>25</sup>.
  3. С обзиром на то да се ради о писаном корпусу начињеном од спољнополитичких вести, природан је велики број разноврсних информација и постојања великог броја ентитета и нових топика који може бити разноврснији него што је то случај у усменом говору, односно дијалогу (видети ТАВОАДА 2002, рад у коме је теорија центрирања примењена на корпус телефонских конверзација на шпанском језику).

У корпусу такође постоји јако мали број прелаза који означавају ЗАДРЖАВАЊЕ (2,16%, видети Табелу 3)<sup>26</sup> што се исто тако не уклапа у схему трећег правила центрирања. Овај прелаз се најчешће јавља када топик из претходне клаузе у којој је био субјекат долази на нижу граматичку позицију у актуелној клаузи (нпр. на позицију директног или индиректног објекта), док је Ср неки други референт (најчешће у субјекатској позицији), као у примеру (10):

- (10) *(а) Он (Малики) се нагиње ка Ирану ради њодршке код куће, али и у њрејешно сунийском региону (б) у коме има мало њријашеља. (в) Исџовремено, њему је и даље њоџребна Америка – делом због војне њомоћи...* (Политика онлајн, 5. октобар 2012)

<sup>24</sup> Број НУЛТИХ ПРЕЛАЗА није наведен у Табелама 2 и 3.

<sup>25</sup> У појединим анализама у оквиру теорије центрирања увођење нових топика се подводи под НУЛТИ ПРЕЛАЗ, уколико не постоји ни један ентитет који се из првог сегмента преноси у други сегмент (видети ТАВОАДА 2002).

<sup>26</sup> У великом броју радова резултати показују да ЛАГАНИ ПРЕЛАЗ превазилази број ЗАДРЖАВАЊА у тексту (референце погледати у KIBBLE 2001).

„Малики“ (Сб, субјекат и топик у клаузи (а)) у клаузи (в) постаје индиректни објекат, али задржава статус Сб-а док Америка има позицију Ср-а и субјекта у датој клаузи<sup>27</sup>.

Корпус овог рада садржи најмањи број ТЕШКИХ прелаза (0,72%, видети Табелу 3), што је сагласно трећем правилу центрирања.

У табели 3 су приказани облици кроз које се Сб реализује у четири традиционална прелаза:

Табела 3. Облици реализације Сб-а кроз именице, заменице и нулте облике (изражени у процентима) и типови прелаза између две клаузе

<i>Тип прелаза / Облик Сб-а</i>	<b>Наставак</b>	<b>Задржавање</b>	<b>Лаган прелаз</b>	<b>Тежак прелаз</b>
Нулти облик	19,00%	0,00%	3,59%	
Заменица	6,12%	0,72%	6,83%	
Именица	6,83%	1,44%	46,00%	0,72%
<b>Укупан број</b>	<b>31,95%</b>	<b>2,16%</b>	<b>56,42%</b>	<b>0,72%</b>

Као што је било наглашено у првом одељку, Сб у свакој клаузи представља топик те клаузе. У научној литератури су успостављене многе паралеле између топика и субјекта (нпр. LI, 1976; GIVON 1983). У радовима из области теорије центрирања и генерисања природног језика, оваква корелација је била проучавана, на пример, у јапанском (КАМЕЈАМА 1986), италијанском (DI EUGENIO 1998) и турском (TURAN 1998)<sup>28</sup> језику. Аутори ових текстова су изнели хипотезу да се у језицима у којима постоји подела на „слабе“ и „јаке“ заменичке облике на форму субјекта може утицати кроз тип прелаза (НАСТАВАК, ЛАГАН ПРЕЛАЗ, итд.) који постоји између две клаузе. Главна хипотеза у наведеним радовима је да „слабе“ заменичке форме углавном представљају субјекте у оквиру прелаза НАСТАВАК док се јаки заменички облици или именичке фразе обично налазе у прелазима који означавају померање центра

<sup>27</sup> Односна клауза у (б) не може да ажурира Сб у (в). Видети одељак о сегментацији дискурса у овом раду.

<sup>28</sup> Видети Поезио (2008).

(ЛАГАН или ТЕЖАК прелаз) или задржавање центра (прелаз ЗАДРЖАВАЊЕ)<sup>29</sup>.

На основу табеле 3, сагласно изложеној хипотези, нулти облик у српском језику чешће изражава С<sub>б</sub> приликом прелаза НАСТАВАК док су именице у субјекатској позицији најчешће уз ЛАГАНИ прелаз. Заменице су у лакој предности у случају ЛАГАНОГ прелаза. Уколико се погледа табела 2, запажа се да односне заменице преовлађују међу заменицама и у оквиру прелаза НАСТАВАК и у оквиру ЛАГАНОГ прелаза, што је илустровано у наредном примеру:

НАСТАВАК и ЛАГАНИ прелаз

(11) *(а) Главни осумњичени је 46-годишњи држављанин САД Александар Фишенко (б)<sup>30</sup> који је, (в) како се наводи, основао компанију „Арт електроникс“ са седиштем у Хјустону, у држави Тексас (г) и преко ње наводно вршио незаконите испоруке високих технолошких микросистема (д) који се могу користити у различитим системима наоружања.*

У примеру (11) клаузе (а) и (б) формирају прелаз НАСТАВАК док клаузе (г) и (д) формирају ЛАГАНИ прелаз. С<sub>б</sub> је у оба случаја изражен односном заменицом.

Остали типови заменица, као што су то показне и одричне заменице, јављају се у веома малом броју у овом корпусу да би се са сигурношћу могли сврстати уз било који прелаз. Такође, у корпусу су нађени само пуни облици личних заменица док енклитика нема, што би се донекле могло објаснити жанром самог корпуса. Уколико би се радило о корпусу говорног језика, енклитике би вероватно биле заступљеније.

Прелаз НАСТАВАК у мањем броју случајева изражава С<sub>б</sub> кроз именицу. То се дешава у случајевима када је С<sub>б</sub> из претходне клаузе поновљен у актуелној клаузи, те говорник жели да постави у фокус<sup>31</sup> информацију изражену С<sub>б</sub>-ом, али и да јасно диференцира С<sub>б</sub>

<sup>29</sup> Резултати у енглеском језику указују да постоји корелација између прелаза НАСТАВАК и личних заменица док су други прелази у корелацији са показним заменицама Поезио (2008).

<sup>30</sup> Клауза (б) гласи: „који је основао компанију ’Арт електроникс’ са седиштем у Хјустону, у држави Тексас“.

<sup>31</sup> Ово се такође поклапа са теоријом хијерархије датости (GUNDEL, HEDBERG et al 1993), која предвиђа да, уколико је неки ентитет у фокусу, онда се он обично

од осталих ентитета у првој клаузи, као што је то илустровано у примеру (12):

- (12) **(a)** *Колико нам је познато, (б)*<sup>32</sup> *оптужбе (в) које су изнете тичу се неколицине грађана Руске Федерације. (г)*<sup>33</sup> *Те оптужбе, (д) према информацијама које нам је пренела америчка страна, имају кривични карактер.*

У реченици (12г) „те оптужбе“ се стављају у фокус и диференцирају у односу на „грађане Руске Федерације“.

ЛАГАНИ прелаз, као што је наглашено у претходном тексту, изражава Сб кроз именицу, док у ретким случајевима то бива и нулта форма. Она се у српском језику најчешће користи уз ЛАГАНИ прелаз када објекат из претходне клаузе постаје субјекат у актуелној клаузи:

- (13) **(a)** *„Тај инцидент јасно илустрира трагичне последице преливања сиријске кризе у суседне земље“, (б) уочирила је она (в) и позвала сиријске власти (г) да „одмах окончају насиље...“*

„Сиријске власти“ у клаузи (в) су у објекатској позицији док у клаузи (г) постају Сб и налазе се у позицији субјекта клаузе.

Један од циљева представљене анализе у овом одељку би могао да буде почетак формализације односа између анафорских појмова и типова прелаза, што може бити коришћено у апликацијама из рачунарске лингвистике.

## V. Закључак

У овом раду је примењена теорија центрирања на писани корпус српског језика. Корпусна анализа је показала да постоји корелација између типова прелаза и избора облика који се користе за изражавање „центра пажње“ у финитним клаузама. Та корелација

изражава најминималнијом формом, што је у *pro-drop* језицима нулти облик, али исто тако има могућност да буде изражен и свим другим облицима којима су исказани мање истакнути ентитети, што су именице или заменице у случају српског језика.

<sup>32</sup> Клауза (б) гласи: „Оптужбе се тичу неколицине грађана“.

<sup>33</sup> Клауза (г) гласи: „Те оптужбе имају кривични карактер“.



одговара постојећој хипотези утврђеној на примеру других језика (шпански, италијански, енглески, јапански, итд.) да се Сб исказан минималном анафорском формом везује за прелаз НАСТАВАК док је именичка форма, као најпотпунија анафорска форма, знатно фреквентнија уз ЛАГАНИ прелаз. Анализа је такође показала да су ЗАДРЖАВАЊЕ И ТЕШКИ прелаз ретки у корпусу новинских вести. То се не уклапа у потпуности у треће правило центрирања, које претпоставља да би ЗАДРЖАВАЊЕ требало да буде други по реду облик прелаза по својој фреквентности. Такође, корпус овог рада обилује ЛАГАНИМ прелазима у много већем броју од НАСТАВКА, што се такође не уклапа у концепцију трећег правила центрирања. У будућој анализи би требало утврдити у којој мери жанровски елемент утиче на фреквентност одређеног прелаза, какву улогу има синтакса (у смислу комбинације различитих финитних клауза) на промену и актуализацију топика (Сб-а), и који би био статус енклитика у односу на различите прелазе. У циљу извођења правилног и комплетног закључка потребно је упоредити писани и говорни корпус српског језика.

*Цитирана литература*

- ARIEL, Mira. 1990. *Accessing Noun Phrase Antecedents*. London/New York: Routledge, 1990.
- BRENNAN, Susan, Marilyn Walker Friedman, and Carl Pollard. "A Centering Approach to Pronouns," *Proceedings of the 25<sup>th</sup> Annual Meeting of the Association for Computational Linguistics*, Stanford, California, 1987, 155- 162.
- CHAFE, Wallace. "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Points of View." *Subject and Topic*. Charles N. Li (ed.), New York: Academic Press, 1976: 27–55.
- DI EUGENIO, Barbara. "Centering in Italian." *Centering Theory in Discourse*. Marilyn Walker, Aravind Joshi and Ellen Prince (eds.), Oxford University Press, 115–137: 1998.
- GIVÓN, Talmy. *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam/Philadelphia: John Benjamins, 1983.
- GORDON, Peter, Barbara J. Grosz, and Laura A. Gilliom. "Pronouns, Names and the Centering of Attention in Discourse." *Cognitive Science*, knj. 17, sv. 3 (1993): 311–347.
- GROSZ, Barbara, Aravind Joshi, and Scott Weinstein. "Providing a unified account of definite noun phrases in discourse," In *21<sup>st</sup> Annual Meeting of the Association for Computational Linguistics: proceedings of the conference: 15-17 June 1983*, Massachusetts Institute of Technology, Cambridge, Massachusetts, ed. Association

- for Computational Linguistics 44-50. Morristown, N.J.: Association for Computational Linguistics.
- GROSZ, Barbara, Aravind Joshi and Scott Weinstein. "Centering: A framework for modeling the local coherence of discourse." *Computational Linguistics*, knj. 2, sv. 12 (1995): 203–226.
- GROSZ, Barbara and Sidner, Candace. "Attention, intention, and the structure of discourse." *Computational Linguistics*, knj. 12, sv. 3, (1986): 175–204 .
- GUNDEL, Jeanette. 1988. "Universals of topic–comment structure." *Studies in Syntactic Typology*. Michael Hammond, Edith Moravcsik and Jessica Wirth (eds.), Amsterdam and Philadelphia: John Benjamins Publishing Company, 1988: 209–239.
- GUNDEL, Jeanette, Nancy Hedberg and Ron Zacharski. "Cognitive status and the form of referring expressions in discourse." *Language* 69, 2 (June 1993): 274–307.
- HADIC-ZABALA, Loreley & Maite Taboada. *Centering Theory in Spanish: Coding Manual. Manuscript*. Simon Fraser University. June 2006.
- HALLIDAY, Michael and Ruqayia, Hasan. *Cohesion in English*. London: Longman, 1976.
- HEDBERG, Nancy. "Centering and Noun Phrase Realization in Kaqchikel Mayan." *Journal of Pragmatic*, knj. 42, sv. 7 (2010): 1829–181.
- KAMEYAMA, Megumi. "A property-sharing constraint in Centering." In *Proceedings, 24th Annual Meeting of the Association for Computational Linguistics*, 200–206, 1986.
- KAMEYAMA, Megumi. "Intrasentential Centering." *Centering Theory in Discourse*. Marilyn Walker, Aravind Joshi, and Ellen Prince (eds.) Oxford University Press, 1998: 89–115.
- KIBBLE, Rodger. "A reformulation of Rule 2 of Centering Theory." *Computational Linguistics*, knj. 27, sv. 4, (2001): 579–587.
- KINTSCH, Walter and van Dijk, Teun "Toward a model of text comprehension and production." *Psychological review*, knj. 85, sv. 5 (1978): 363–394.
- LI, Charles. *Subject and Topic*. New York: Academic Press, 1976.
- MANN, William C. And Sandra A. Thompson. "Rhetorical Structure Theory: Toward a functional theory of text organization." *Text*, knj. 8, sv. 3 (1988): 243–281.
- MATSUI, Tomoko. "Bridging reference and the notions of 'topic' and 'focus'." *Lingua*, knj. 90 (1993): 49–63.
- POESIO, Massimo, Rosemary Stevenson, Barbara Di Eugenio and Janet Hitzeman. Centering: A parametric theory and its instantiations. *Computational Linguistics*, knj. 30, sv. 3 (2004): 309–363.
- TABOADA, Maite. Centering and Pronominal Reference: "In Dialogue, In Spanish." *Proceedings, 6th Workshop on the Semantics and Pragmatics of Dialogue, EDILOG*. Edinburgh. September 2002: 177–184.
- TABOADA, Maite. "Reference, Centers and Transitions in spoken Spanish." *Reference: Interdisciplinary Perspectives*. Oxford: Oxford University Press, 2008: 176–215.
- THOMPSON, Sandra. A. "Object complements" and conversation. Towards a realistic account." *Studies in Language*, knj. 26, sv.1 (2002): 125–164.

- TURAN, Umit Deniz. *Null vs. Overt subjects in Turkish discourse: A centering analysis*. PhD thesis, University of Pennsylvania, 1995.
- TURAN, Umit Deniz. "Ranking forward-looking centers in Turkish. Centering Theory in Discourse." Marilyn Walker & Ellen Prince (eds.), Oxford University Press, 1998: 139-160.
- WALKER, Marilyn, A., Masayo, Iida and Sharon Cote. Japanese discourse and the process of Centering. *Computational Linguistics*, knj. 20, sv. 2 (1994).
- WALKER, Marilyn, Aravind Joshi and Ellen Prince. *Centering Theory in Discourse*. Oxford: Clarendon Press, 1998.

Радослава Трнавац

ЦЕНТРИСТСКАЈА ТЕОРИЈА, ТИПИ ПЕРЕХОДОВА  
И ОБОЗНАЧЕНИЕ РЕФЕРЕНТА В СЕРБСКОЈ ПИСЬМЕНОЈ РЕЧИ

Резюме

Цель этой работы – сравнить употребление различных видов анафоры в так называемых переходах центристской теории (WALKER, JOSHI ET AL. 1998) в корпусе сербских газетных текстов. Центристская теория – это теория локальной согласованности (когерентности) текста, в которой четыре типа переходов рассматриваются как параметры этой согласованности. Согласно гипотезам, высказанным в исследованиях, базирующихся на материале разных языков, *последовательный* переход характеризуется минимальной формой топики (нулевая форма), а *плавный* и *резкий* переходы сочетаются с полной именной фразой топики. Анализ обнаруживает, что 'правило последовательности' центристской теории не соблюдается полностью в корпусе письменной речи сербского языка, поскольку *плавный* переход является самым частотным типом перехода в корпусе. В качестве доказательства данного утверждения в работе приводятся (а) жанр газетной статьи и (б) способ сочетания простых предложений в сложном предложении сербской письменной речи.

*Ключевые слова:* центристская теория; типы переходов; Cb; Cp; согласованность текста; сегментация текста; предложение.

---

Radoslava Trnavac

CENTERING THEORY, TRANSITIONS AND REFERENT MARKING  
IN THE CORPUS OF THE WRITTEN SERBIAN LANGUAGE

Summary

The aim of the paper is to map various types of anaphor to the centering transitions in the corpus of Serbian newspaper articles. Centering Theory (WALKER, JOSHI ET AL. 1998) is a theory of local coherence in which four types of transitions are used as parameters of coherence. As is hypothesized in the previous literature based on various languages, the CONTINUE transition is mostly characterized by a minimal form (a zero form of the topic), while the SMOOTH and ROUGH shifts are found with a full noun phrase topic. The analysis shows that the "Ordering Rule" of Centering Theory is not fully followed in the written corpus of the Serbian language since SMOOTH SHIFT has been identified as a prevalent type of transition in the corpus. The following two reasons were identified for that: (1) the genre of the newspaper articles, and (2) the way clauses are combined within a complex sentence in the Serbian written corpus.

*Keywords:* Centering Theory; transition types; Cb; Cp; textual coherence; segmentation of texts; clause.