

TAMARA GAZDIĆ-ALERIĆ, MARKO ALERIĆ i BOŽO BEKAVAC\*

## Utjecaj čestotnosti fonema/grafema č, ć, dž, đ na ovladavanje hrvatskom pravopisnom normom

Na jednomilijunskome korpusu oblikovanomu tako da svojom strukturom odražava omjere određene u *Hrvatskome nacionalnom korpusu* (HNK) provedeno je istraživanje o učestalosti pojavljivanja fonema/grafema. Obradeni korpus čine suvremeni tekstovi od kojih je 60 % tekstova iz jednih dnevnih novina, 35 % prozних tekstova i 5 % tekstova iz jednoga tjednika. Nakon određivanja učestalosti pojavljivanja fonema/grafema u hrvatskome jeziku izdvojena su na temelju osnovnoškolskih i srednjoškolskih školskih zadaća najčešća odstupanja pri pisanju tih grafema. Autori na osnovi dobivenih rezultata nude nov model poučavanja pravopisne norme u osnovnoj i srednjoj školi. Šira je inačica članka objavljena u časopisu *Odgojne znanosti* Vol. 12, br. 1, 2010, str. 215. – 227 (<https://hrcak.srce.hr/59619>). Istraživanje je provedeno 2010. godine, pa u njemu nisu obuhvaćeni neki noviji izvori.

Svaki se jezični sustav sastoji od jedinica koje imaju određenu čestotnost (frekventnost, učestalost). To svojstvo jezičnih jedinica ne bi smjelo biti zanemareno, nego treba biti iskorišteno, osobito pri poučavanju. Zato je cilj ovoga istraživanja utvrditi koliko čestotnost pojavljivanja fonema/grafema u hrvatskome standardnom jeziku može utjecati na ovladavanje njima, odnosno njihovu uspješnu primjenu u govorenoj te osobito u pisanoj jezičnoj komunikaciji. Istraživanjem je utvrđena čestotnost fonema/grafema u hrvatskome standardnom jeziku. Posebna je pozornost posvećena grafemima pri čijemu se pisanju pojavljuje najveći broj odstupanja od hrvatskoga standardnog jezika. To je dovelo do definiranja postupaka i metoda koji bi omogućili primjenu načela čestotnosti u ovladavanju pravilnim pisanjem onih grafema pri čijemu je pisanju broj odstupanja najveći.

I prije je bilo sličnih istraživanja o čestotnosti nekih fonema/grafema. Maretić (1931.) na uzorku od 10 000 grafema u biblijskome tekstu utvrđuje čestotnost grafema, i posredno fonema, u hrvatskome standardnom (književnom) jeziku. Slično, na

\* Tamara Gazdić-Alerić izvanredna je profesorica na Učiteljskome fakultetu Sveučilišta u Zagrebu, Marko Alerić izvanredni je profesor na Odsjeku za kroatistiku Filozofskoga fakulteta Sveučilišta u Zagrebu, Božo Bekavac docent je na Odsjeku za lingvistiku Filozofskoga fakulteta Sveučilišta u Zagrebu.

različitim korpusima govorenih i pisanih tekstova čestotnost fonema (grafema) istražuju Matković i Sinković (1970.), Škarić (1991.) te Vuletić (1991.). U ovome istraživanju čestotnost fonema/grafema u hrvatskome standardnom jeziku utvrđena je strojnim brojenjem grafema u korpusu koji se sastojao točno od 1 000 803 pojavnice. Taj je korpus preuzet iz *Hrvatskoga nacionalnog korpusa* (Tadić, 1996. i 1998.) i njegov je podskup, a činili su ga prozni tekstovi (35 %), tekstovi iz dnevnoga tiska (65 %: *Vjesnik* iz 2000.) i tekstovi iz tjednoga tiska (5 %: *Nacional* iz 2004.). Za obradu korpusa upotrijebljen je računalnojezikoslovni alat *Intex* (Silberztein), koji je već bio upotrijebljen za obradu hrvatskih tekstova. U korpusu su prebrojeni samo grafemi koji pripadaju hrvatskomu grafemskom sustavu. Dvoslovi *dž*, *lj*, *nj* brojeni su posebno (s pomoću *Perl*-skripti) jer ih upotrijebljeni program u statističkome izvještaju ne može prikazati. Rezultati mjerenja čestotnosti fonema/grafema u korpusu koji nam je poslužio za istraživanje prikazani su u usporedbi s rezultatima prijašnjih mjerenja.

Redosljed je čestotnosti fonema/grafema, od onoga s najvećom čestotnošću do onoga s najmanjom, u istraživanjima Tome Maretića, Ive Škarića, Dušanke Vuletić i istraživanju provedenome na našem korpusu ovaj:

Maretić: *e, a, i, o, j, s, n, u, d, v, t, r, m, k, p, l, g, b, š, z, č, ć, lj, h, ž, nj, c, đ, dž, f*

Škarić: *a, i, e, o, n, s, r, t, u, j, m, d, v, k, l, p, z, g, c, b, š, č, ć, h, lj, nj, ž, đ, f, dž*

Vuletić: *a, i, e, o, n, t, s, r, j, m, d, u, k, l, v, p, b, š, g, z, č, c, ć, ž, h, lj, nj, f, đ, dž*

korpus: *a, i, o, e, n, r, t, s, u, j, k, v, d, m, p, l, z, g, b, č, c, š, ć, nj, h, lj, ž, f, đ, dž.*

Ako se usporede rezultati svih mjerenja, uočava se kako čestotnost fonema/grafema *a, c, č, f, h, k, l, n, nj, p, r, t, z* iz mjerenja u mjerenje raste, čestotnost fonema/grafema *dž, đ, i, o, s, u* ostaje uglavnom ista, čestotnost fonema/grafema *m* u starijim mjerenjima raste, a u novijima se smanjuje, čestotnost fonema/grafema *v* u starijim se mjerenjima smanjuje, a u novijima raste, a čestotnost fonema/grafema *b, ć, d, e, g, j, lj, š, ž* u novijim se mjerenjima uglavnom smanjuje. Najveća ujednačenost u svim mjerenjima čestotnosti (razlika je  $\pm 15$  % ili manja) utvrđena je za foneme/grafeme: *dž, i, k, o, s, ž*. Razlike u mjerenjima uvjetovane su različitim brojem pojava u korpusima na kojima se istraživalo, zatim njihovom pripadnošću različitim tekstnim vrstama i funkcionalnim stilovima, ali i činjenicom da čestotnost fonema/grafema u jeziku nije zadana, nego podliježe promjenama do kojih s vremenom dolazi, a koje su uvjetovane i promjenama u leksiku s obzirom na to da je leksička razina u standardnome jeziku najpodložnija promjenama. Čestotnost fonema/grafema *č, ć, dž, đ*, odnosno njihov konkretan broj u priručnicima hrvatskog standardnog jezika, pokušali smo pokazati utvrđivanjem broja leksema u kojima se oni pojavljuju na njihovu početku iako se npr. fonemi/grafemi *č* i *ć* češće pojavljuju u morfemima na kraju leksema. Ipak, upravo je za ovladavanje pravilnim pisanjem leksema u kojima se fonemi/grafemi *č, ć, dž, đ* pojavljuju na njihovu početku načelo čestotnosti najprimjerenije

jer se za te slučajeve najčešće ne može odrediti pravilo koje bi olakšalo ovladavanje pravilnim načinom njihova pisanja.

Broj leksema utvrdili smo u ovim priručnicima<sup>2</sup>:

1. *Hrvatski čestotni rječnik* (1999.): č (401 leksem), ć (61), dž (17), đ (24)
  2. *Rječnik hrvatskoga jezika* (1991.): č (606), ć (119), dž (57), đ (50)
  3. *Hrvatski pravopis* (2000.): č (851), ć (159), dž (88), đ (97)
  4. *Pravopis hrvatskoga jezika* (2001.): č (1167), ć (267), dž (216), đ (182)
  5. *Hrvatski pravopis* (2007.): č (719), ć (159), dž (199), đ (92)
- Prosječno: č (748,8), ć (153), dž (115,4), đ (89).

Iako se broj leksema koji započinju fonemima/grafemima č, ć, dž, đ u priručnicima i bitno razlikuje, to se može objasniti time da kriterij izbora leksema u svim priručnicima nije bio isti. Sigurno je da se pri izboru leksema nije u svim priručnicima vodilo računa npr. o načelu čestotnosti, odnosno o tome koliko je neki leksem prisutan u hrvatskome standardnom jeziku, pa u njima nisu nužno navedeni samo leksemi koji se češće pojavljuju.

Utvrđena čestotnost fonema/grafema u hrvatskome standardnom jeziku u određenoj se mjeri podudara s čestotnošću leksema u kojima se ti fonemi/grafemi pojavljuju. Naime, u leksemima hrvatskoga jezika fonemi/grafemi č, ć, dž, đ pojavljuju se otprilike u onome postotku u kojem se ti fonemi/grafemi pojavljuju u odnosu na ostale foneme/grafeme hrvatskoga standardnog jezika. To znači da je čestotnost leksema u kojima se pojavljuje fonem/grafem č otprilike 0,98 %, čestotnost leksema u kojima se pojavljuje fonem/grafem ć otprilike 0,71 %, čestotnost leksema u kojima se pojavljuje fonem/grafem dž 0,01 %, a čestotnost leksema u kojima se pojavljuje fonem/grafem đ 0,21 %.

Ako pretpostavimo da se u komunikaciji na hrvatskome ili bilo kojemu drugom jeziku upotrebljava manje od 2000 različitih leksema, onda to znači da je broj leksema u kojima se pojavljuju fonemi/grafemi č, ć, dž, đ vrlo malen (č = 19,6 riječi, ć = 14,2 riječi, dž = 0,2 riječi, đ = 4,2 riječi). Kako se može pretpostaviti da je broj odstupanja u izgovoru i pisanju tih fonema/grafema u tijesnoj povezanosti s njihovom čestotnošću, odnosno da joj je obrnuto razmjeran, može se i pretpostaviti da će odstupanja u izgovoru i pisanju onih fonema/grafema koji se pojavljuju češće biti rjeđa, a onih fonema/grafema koji se pojavljuju rjeđe češća.

Iskustvo u nastavi pokazuje da učenici pri pisanju čine najveći broj odstupanja u pisanju grafema č, ć, dž i đ u odnosu na pisanje ostalih grafema. Općenito je, naime, poznato

<sup>2</sup> Moguš, Milan; Bratanić, Maja; Tadić, Marko. 1999. *Hrvatski čestotni rječnik*; Anić, Vladimir. 1991. *Rječnik hrvatskoga jezika*; Babić, Stjepan; Finka, Božidar; Moguš, Milan. 2000. *Hrvatski pravopis*; Anić, Vladimir; Silić, Josip. 2001. *Pravopis hrvatskoga jezika*; Badurina, Lada; Marković, Ivan; Mićanović, Krešimir. 2007. *Hrvatski pravopis*.

kako su odstupanja u izgovoru i pisanju upravo tih fonema/grafema najbrojnija. Razlog svakako može biti u sličnosti njihova izgovora, u njihovoj različitoj artikulaciji u hrvatskome standardnom jeziku u odnosu na artikulaciju njima sličnih fonema u hrvatskim mjesnim govorima, ali i u činjenici da se ti fonemi/grafemi u hrvatskome standardnom jeziku pojavljuju vrlo rijetko.

Da bi podatak o čestotnosti fonema/grafema mogao biti iskorišten za uspješnije ovladavanje izgovorom i pisanjem fonema/grafema *č, ć, dž, đ* u osnovnoškolskoj i srednjoškolskoj nastavi hrvatskoga jezika, bilo je potrebno utvrditi konkretne lekseme pri čijoj uporabi učenici najčešće odstupaju od pravopisne norme. Popis leksema pri čijemu su pisanju češća učenička odstupanja u vezi s uporabom fonema/grafema *č, ć, dž, đ* izradili smo na korpusu pisanoga jezika, odnosno na korpusu školskih zadaća učenika osnovnih i srednjih škola. Za ovo istraživanje korpus je činilo ukupno 370 osnovnoškolskih zadaća te 131 srednjoškolska zadaća, koje su napisane u jedanaest osnovnih škola, pet gimnazija i jednoj strukovnoj školi iz petnaest hrvatskih gradova i mjesta u školskim godinama 2000./2001. i 2001./2002. Za ovaj članak donosimo skraćenu interpretaciju istraživanja.

Osnovna škola – ovo su leksemi kod kojih se odstupanja pojavljuju više puta: *budućnost* (2), *mačuhica* (2), *mogućnosti* (2), *lišće* (od *list*) (2), *liječnik* (2), *oćiju* (2), *pomoć* (2), *proči* (2), *vračala (mi se slika)* (2), *doči* (2), *drveće* (2), *svećano* (2), *učim, naučili, učioni* (3), *vočki* (3), *več* (4), *vrtič* (5), *poćeli* (5), *zadachnica, zadaća* (5), *moći* (6), *osjeća, osjećala, osjećaj* (12), *kuća, kućica, kućerak* (14), *ču, će, ćemo, čete, neće, nećemo* (18).

Srednja škola – ovo su leksemi kod kojih se odstupanja pojavljuju više puta: *budućnost* (2), *mući (mućiti)* (2), *osjeća* (2), *osvećuje* (2), *pomoći* (2), *potjeće* (2), *većina* (2), *več* (3), *če, ćemo, ču* (4), *prihvaća* (5), *srećom* (5).

Broj odstupanja pri pisanju leksema *č, ć, dž, đ* pokazuje kako osnovnoškolci i srednjoškolci griješe u ograničenome, razmjerno malome broju leksema, da se ukupan broj odstupanja u srednjoj školi u odnosu na osnovnu školu povećava te da i osnovnoškolci i srednjoškolci otprilike dvaput više griješe pri pisanju leksema u kojima u skladu s pravopisnom normom treba napisati grafem *ć*, nego leksema u kojima treba napisati grafem *č*.

Rezultati istraživanja pokazuju kako je broj odstupanja u kojima se u leksiku osnovnoškolaca i srednjoškolaca pojavljuju fonemi/grafemi *č, ć, dž, đ* ograničen i razmjerno malen, što se podudara s utvrđenom čestotnošću tih fonema/grafema u hrvatskome standardnom jeziku. Dakle, važan je razlog zbog kojega učenici griješe pri pisanju grafema *č, ć, dž, đ* u njihovoj maloj čestotnosti jer je posve prihvatljivo pretpostaviti kako će pri njihovoj uporabi, zbog male čestotnosti pojavljivanja, učenici činiti veći broj pogrešaka, zamjenjujući ih češćim grafemima, što je potvrđeno u činjenici da učenici često griješe pri pisanju grafema koji imaju manju čestotnost, odnosno pri

pisanju leksema s kojima se toliko rijetko susreću da njihovim pravilnim pisanjem ne uspiju ovladati. U skladu s tim je i činjenica da je, unatoč tome što je učestalost fonema/grafema *č, ć, dž, đ*, odnosno leksema u kojima se oni pojavljuju vrlo mala, broj odstupanja u pisanju tih grafema u učeničkim školskim zadaćama velik, osobito s obzirom na činjenicu da učenici pri pisanju školske zadaće uglavnom mogu izbjeći riječ za koju nisu sigurni kako se piše. Naime, može se pretpostaviti da bi broj odstupanja u nekome drugom spontano napisanome tekstu, za koji učenik ne zna hoće li mu biti pregledan i/ili ocijenjen, bio daleko veći.

Zbog poteškoća u izgovoru i pisanju fonema/grafema *č, ć, dž, đ* tu je temu, prema *Nastavnome planu i programu za osnovnu školu*, potrebno obraditi u 1., 3., 4., 7. i 8. razredu (usp. *Nastavni plan i program za osnovnu školu*, 2006.). Ipak, primjena načela čestotnosti pritom se može uočiti samo u 4. razredu, u kojemu je nastavna tema i *Izgovor i pisanje č, ć, dž, đ, lj, nj, ije/je/e/i*, a kao obrazovna postignuća uz tu nastavnu temu navodi se: „pravilno izgovarati i pisati *č, ć, dž, đ, lj, nj, ije/je/e/i* u češće rabljenim riječima”. Sustavnija i razrađenija primjena načela čestotnosti, koja podrazumijeva izradu popisa češćih riječi u kojima se navedeni fonemi/grafemi pojavljuju i popisa češćih riječi u kojima pri njihovu izgovoru i pisanju dolazi do odstupanja, može pridonijeti boljem ovladavanju pravilnim pisanjem tih riječi. Zato je u *Nastavnome planu i programu*, udžbenicima i općenito nastavi hrvatskoga jezika potrebno češće isticati kako je poznavanje izgovora i pisanja fonema/grafema *č, ć, dž, đ* u riječima jednim dijelom ovisno o znanju pravila o njihovoj uporabi, ali i o tome da učenici ovladaju pravilnim pisanjem češćih riječi. Tu činjenicu moguće je iskoristiti na početku nastavnoga sata (kao motivaciju), u središnjemu dijelu nastavnoga sata i/ili na kraju nastavnoga sata za provjeru usvojenosti pravilnoga pisanja grafema *č, ć, dž, đ*. Načelo čestotnosti osobito je primjenjivo za ovladavanje pravilnim pisanjem onih leksema u kojima su prisutni fonemi/grafemi *č, ć, dž, đ* za koje pravilima nije utvrđeno zašto se u njima pojavljuju, a upravo je u toj skupini leksema, u proučenome korpusu učeničkih školskih zadaća, broj učeničkih odstupanja bio najveći.

Primjena načela čestotnosti u nastavi hrvatskoga jezika trebala bi se, dakle, temeljiti:

1. na isticanju male čestotnosti fonema/grafema *č, ć, dž, đ*
2. na isticanju maloga i razmjerno ograničenoga broja leksema u kojima se ti fonemi/grafemi pojavljuju na početku riječi
3. na isticanju maloga i razmjerno ograničenoga broja leksema pri čijemu pisanju dolazi do odstupanja u pisanju fonema/grafema *č, ć, dž, đ*
4. na čestome upozoravanju na lekseme za koje je utvrđeno da pri njihovu pisanju često dolazi do odstupanja te na provjeravanju njihova pisanja.