

UTJECAJ ČESTOTNOSTI FONEMA/GRAFEMA Č, Ć, DŽ, Đ NA OVLADAVANJE HRVATSKOM PRAVOPISNOM NORMOM

Tamara Gazdić-Alerić¹, Marko Alerić² i Božo Bekavac²

Učiteljski fakultet, Sveučilište u Zagrebu¹
Filozofski fakultet, Sveučilište u Zagrebu²

tamara.gazdic-aleric@ufzg.hr

Sažetak - U istraživanju se navode dosadašnja utvrđivanja učestalosti pojavljivanja fonema/grafema u hrvatskome standardnom jeziku, a zatim se prikazuje postupak utvrđivanja njihove učestalosti uz pomoć suvremene računalnojezikoslovne metodologije.

Učestalost pojavljivanja fonema/grafema utvrđuje se na jednomilijunskom korpusu oblikovanom tako da strukturom odražava omjere određene u Hrvatskome nacionalnom korpusu (HNK). Obradeni korpus čine suvremeni tekstovi od kojih je 60% tekstova iz jednog dnevnika, 35% proznih tekstova i 5% tekstova iz jednog tjednika.

Nakon određivanja učestalosti pojavljivanja fonema/grafema u hrvatskome jeziku određuju se najčešća odstupanja prilikom pisanja grafema u hrvatskome standardnom jeziku, na temelju osnovnoškolskih i srednjoškolskih učeničkih školskih zadaća.

Na temelju utvrđenih odstupanja u učeničkim školskim zadaćama te podataka o učestalosti pojavljivanja fonema/grafema u hrvatskome jeziku, predlažu se nove metode i postupci u poučavanju pravopisnih pravila koji bi mogli pridonijeti boljem ovladavanju hrvatske pravopisne norme.

Ključne riječi: čestotnost fonema/grafema, odstupanja od pravopisne norme, ovladavanje pravopisnom normom

1. Cilj i zadaci istraživanja

Svaki se jezični sustav sastoji od činjenica koje u njemu imaju i određenu čestotnost (frekventnost, učestalost). Ta osobina jezičnih činjenica ne bi smjela biti zanemarena, već iskorištena, osobito prilikom njihova poučavanja. Zato je cilj ovoga istraživanja bio utvrditi koliko čestotnost pojavljivanja fonema/

grafema u hrvatskome standardnom jeziku može utjecati na njihovo ovladavanje, odnosno njihovu uspješniju primjenu u govorenoj, a osobito pisanoj jezičnoj komunikaciji.

Zadaci istraživanja bili su:

- utvrditi čestotnost fonema/grafema u hrvatskome standardnom jeziku¹
- utvrditi grafeme prilikom čijeg se pisanja javlja najveći broj odstupanja od hrvatskoga standardnog jezika
- utvrditi postupke i metode koji bi omogućili primjenu načela čestotnosti u ovladavanju pravilnim pisanjem onih grafema kod čijeg je pisanja broj odstupanja najveći.

2. Čestotnost fonema/grafema u hrvatskome standardnom jeziku

Da se u hrvatskome standardnom (književnom) jeziku neki fonem/grafem javlja rjeđe, a drugi češće, kao i da učestalost pojavljivanja fonema/grafema u jezicima nije jednaka te da ima određeno značenje, ističe Maretić (Maretić, 1931, 13, 14) koji na uzorku od 10.000 grafema u biblijskome tekstu utvrđuje čestotnost grafema i posredno fonema u hrvatskome standardnom (književnom) jeziku.²

Slično, na raznim korpusima govorenih i pisanih tekstova, čestotnost fonema (grafema) istražuju Matković i Sinković (1970), Škarić (1991) te D. Vuletić (1991).

U našem istraživanju čestotnost fonema/grafema u hrvatskome standardnom jeziku utvrđena je strojnim brojanjem grafema u korpusu koji su činile točno 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* (Silberstein, 1999) koji je već bio upotrijebljen za obradu hrvatskih tekstova (Bekavac, 2005). U korpusu su brojani samo grafemi koji pripadaju hrvatskome grafemskom sustavu. Digrami *dž*, *lj*, *nj* brojani su posebno (uz pomoć *Perlskripti*) jer ih upotrijebljeni program u statističkom izvještaju ne može prikazati.

Rezultati mjerenja čestotnosti fonema/grafema u korpusu koji nam je poslužio za istraživanje, a koji se ovdje navode abecednim slijedom, prikazani su u usporedbi s rezultatima prijašnjih mjerenja (Maretićevih, Škarićevih i

1 Može se pretpostaviti da su tekstovi u izabranom korpusu na kojemu je istraživanje provedeno bili uglavnom oblikovani na hrvatskome standardnom jeziku.

2 Rezultati Maretićeva brojanja poslije su često navođeni, usp. Antoš 1974, 30.

D. Vuletić). Također je, uz rezultate čestotnosti fonema/grafema dobivenih prijašnjim mjerenjima, izračunat postotak odstupanja u odnosu na mjerenje provedeno na korpusu koji smo odabrali za istraživanje.

Čestotnost fonema/grafema

Fonem/ grafem	Maretić	Škarić	Vuletić	Prosjek	Korpus
A	10,79% (+7,04%) ³	11,64% (-0,77%)	12,93% (-10,67%)	11,79%	11,55%
B	1,73% (-15,03%)	1,44% (+2,04%)	1,49% (-1,34%)	1,56%	1,47%
C	0,56% (+73,21%)	0,88% (+10,23%)	0,78% (+24,36%)	0,74%	0,97%
Č	1,20% (-18,33%)	0,81% (+20,99%)	0,91% (+7,69%)	0,97%	0,98%
Ć	0,91% (-21,98%)	0,75% (-5,33%)	0,65% (+9,23%)	0,77%	0,71%
D	4,28% (-20,33%)	3,56% (-11,80%)	3,81% (+10,50%)	3,89%	3,41%
DŽ	0,01% (0,00%)	0,01% (0,00%)	0,02% (-50%)	0,01%	0,01%
Đ	0,31% (-32,26%)	0,19% (+10,53%)	0,19% (+10,53%)	0,23%	0,21%
E	10,99% (-20,56%)	9,34% (-6,53%)	9,17% (-4,80%)	9,83%	8,73%
F	0,01% (+2,20%)	0,14% (+64,29%)	0,23% (0,00%)	0,13%	0,23%
G	1,92% (-11,46%)	1,60% (+6,25%)	1,42% (+19,72%)	1,65%	1,70%
H	0,73% (+9,59%)	0,71% (+12,67%)	0,54% (+48,15%)	0,66%	0,80%
I	10,35% (-2,32%)	9,61% (+5,20%)	10,53% (-3,99%)	10,16%	10,11%
J	5,03% (-21,27%)	4,16% (-4,81%)	4,11% (-3,65%)	4,43%	3,96%
K	3,18% (+14,78%)	3,44% (+6,10%)	3,45% (+5,80%)	3,35%	3,65%
L	1,94% (+41,75%)	2,89% (-4,84%)	3,00% (-8,33%)	2,61%	2,75%

3 Oznakama + i – označavaju se razlike u čestotnosti pojavljivanja navedenog fonema/grafema u korpusu koji smo odabrali za istraživanje u odnosu na njegovu čestotnost u prijašnjim mjerenjima.

LJ	0,99% (-40,00%)	0,66% (-18,18%)	0,50% (+8,00%)	0,69%	0,54%
M	3,67% (-12,26%)	4,03% (-20,01%)	4,09% (-21,27%)	3,93%	3,22%
N	4,57% (+28,88%)	5,44% (+8,27%)	5,46% (+7,88%)	5,16%	5,89%
Nj	0,63% (+28,57%)	0,70% (+15,71%)	0,48% (+68,75%)	0,60%	0,81%
O	9,60% (-1,88%)	8,98% (+4,90%)	9,72% (-3,18%)	9,43%	9,42%
P	2,26% (+28,32%)	2,69% (+7,81%)	2,74% (+5,52%)	2,56%	2,90%
R	3,88% (+31,19%)	4,79% (+6,26%)	4,15% (+22,65%)	4,27%	5,09%
S	4,64% (+5,17%)	4,97% (-1,82%)	4,62% (+5,63%)	4,74%	4,88%
Š	1,50% (-39,33%)	1,20% (-24,17%)	1,47% (-38,10%)	1,39%	0,91%
T	4,00% (+22,25%)	4,55% (+7,47%)	5,06% (-3,36%)	4,54%	4,89%
U	4,29% (+0,67%)	4,55% (-5,05%)	3,61% (+19,67%)	4,15%	4,32%
V	4,02% (-12,69%)	3,51% (0,00%)	2,96% (+18,58%)	3,50%	3,51%
Z	1,47% (+25,17%)	1,69% (+2,79%)	1,35% (+36,30%)	1,54%	1,84%
Ž	0,64% (-15,63%)	0,63% (-14,28%)	0,59% (-8,47%)	0,62%	0,54%

Redoslijed čestotnosti fonema/grafema, od onog s najvećom čestotnošću do onog s najmanjom, u Maretićevu, Škarićevu, istraživanju D. Vuletić i istraživanju provedenom u našem korpusu jest ovaj:⁴

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*⁵

Škarić: *a, i, e, o, n, s, r, t, u,*⁶ *j, m, d, v, k, l, p, z, g, c, b, š, č, ć, h, lj, nj, ž, đ, f, dž*

4 U rječniku engleskog jezika navedeno je kako se 70% svih oblika komunikacije odvija uz pomoć 1.000 različitih riječi (usp. <http://www.askoxford.com/oec/mainpage/oec02/>). Slično se u govornom korpusu hrvatskoga jezika od 114.717 riječi njih 866 javlja deset i više puta. Među tim riječima nema ni jedne riječi koja počinje fonemima/grafemima *dž, đ*, od onih koje počinju fonemom/grafemom *č* ih je 27, a od onih koje počinju fonemom/grafemom *ć* samo je jedna riječ (usp. Vuletić Istraživanje govora, 1991., str. 113)

5 Fonemi/grafemi *dž, f* u Maretićevu istraživanju imaju isti postotak pojavljivanja.

6 Fonemi/grafemi *t, u* u Škarićevu istraživanju imaju isti postotak pojavljivanja.

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 mjerjenje raste, fonema/grafema dž, đ, i, o, s, u ostaje uglavnom ista, fonema/grafema m u starijim mjerenjima raste, a u novijima se smanjuje, fonema/grafema v u starijim se mjerenjima smanjuje, a u novijima raste, a 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.

Maretićevo je mjerenje čestotnosti fonema/grafema najbliže rezultatima dobivenima u korpusu koji nam je poslužio za istraživanje za 6, D. Vuletić za 12, a Škarićevo za 14 fonema/grafema.

Razlike u mjerenjima uvjetovane su različitim brojem pojava u korpusima na kojima se istraživalo, zatim njihovom pripadnošću raznim tekstnim vrstama i funkcionalnim stilovima, ali i činjenicom da čestotnost fonema/grafema u jeziku nije zadana, već da 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,⁷ odnosno njihov konkretan broj u priručnicima hrvatskoga standardnog jezika,⁸ pokušali smo pokazati utvrđivanjem broja leksema u kojima se oni javljaju na njihovu početku, iako se, npr. fonemi/grafemi č i ć češće javljaju u morfemima na kraju leksema. Ipak, upravo je za ovladavanje pravilnim pisanjem leksema u kojima se fonemi/grafemi č, ć, dž, đ javljaju na njihovu početku, načelo čestotnosti najprimjenije 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:

1. *Hrvatski čestotni rječnik*: č (401 leksem), ć (61), dž (17), đ (24)

2. *Rječnik hrvatskoga jezika*: č (606), ć (119), dž (57), đ (50)

3. *Hrvatski pravopis*: č (851), ć (159), dž (88), đ (97)

4. *Pravopis hrvatskoga jezika*: č (1167), ć (267), dž (216), đ (182)

5. *Hrvatski pravopis*: č (719), ć (159), dž (199), đ (92)

Prosječno: č (748,8), ć (153), dž (115,4), đ (89).

7 Fonemi/grafemi č, ć, dž, đ odabrani su zato što se prilikom njihova pisanja najčešće griješi.

8 Moguš, M., Bratanić, M., Tadić, M. (1999) *Hrvatski čestotni rječnik*; Anić, V. (1991) *Rječnik hrvatskoga jezika*; Babić, S., Finka, B., Moguš, M. (2000) *Hrvatski pravopis*; Anić, V., Silić, J. (2001) *Pravopis hrvatskoga jezika*; Badurina, L., Marković, M., Mićanović, K. (2007) *Hrvatski pravopis*.

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

Bez obzira na nejednak broj leksema koji počinju fonemima/grafemima *č, ć, dž, đ* u raznim priručnicima, uvjereni smo da je učestalost leksema u kojima se ti fonemi/grafemi javljaju, a koji se u većoj mjeri podudara s njihovom ukupnom učestalošću, podatak koji treba uzeti u obzir prilikom poučavanja njihova pravilnog izgovaranja i pisanja.

3. Odstupanja prilikom upotrebe fonema/grafema *č, ć, dž, đ*

Utvrđena čestotnost fonema/grafema u hrvatskome standardnom jeziku u određenoj se mjeri podudara s čestotnošću leksema u kojima se ti fonemi/grafemi javljaju. Naime, u leksemima hrvatskoga jezika fonemi/grafemi *č, ć, dž, đ* javljaju se otprilike u onom postotku u kojem se ti fonemi/grafemi javljaju u odnosu na ostale foneme/grafeme hrvatskoga standardnog jezika. To znači da je čestotnost leksema u kojima se javlja fonem/grafem *č* otprilike 0,98%, čestotnost leksema u kojima se javlja fonem/grafem *ć* otprilike 0,71%, a čestotnost leksema u kojima se javljaju fonemi/grafemi *dž, đ* 0,01%, odnosno 0,21%. Ako pretpostavimo da se u komunikaciji na hrvatskom ili bilo kojem drugom jeziku upotrebljava manje od 2.000 različitih leksema (Šojat, 1964, 105) /4/, onda to znači da je broj leksema u kojima se javljaju 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 izgovaranju i pisanju tih fonema/grafema u tijesnoj povezanosti s njihovom čestotnošću, odnosno da joj je obrnuto razmjernan, može se i pretpostaviti da će odstupanja u izgovoru i pisanju onih fonema/grafema koji se javljaju češće biti rjeđa, a onih fonema/grafema koji se javljaju rjeđe češća.

Primjenu načela čestotnosti u ovom smo radu odlučili prikazati upravo na fonemima/grafemima *č, ć, dž, đ*:

1. zbog toga što učenici prilikom njihova pisanja čine najveći broj odstupanja u odnosu na pisanje ostalih fonema/grafema. Općenito je, naime, poznato 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 govovima, ali i u činjenici da se ti fonemi/grafemi u hrvatskome standardnom jeziku javljaju vrlo rijetko.

2. zbog toga što se rezultati čestotnosti tih fonema/grafema u korpusu koji smo odabrali gotovo potpuno podudaraju s prosječnim rezultatom svih prijašnjih mjerenja njihove čestotnosti.

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 prilikom čije upotrebe učenici čine najveći broj odstupanja. Popis leksema prilikom čijeg su pisanja češća učenička odstupanja u vezi s upotrebom fonema/grafema č, ć, dž, đ izradili smo na korpusu pisanog jezika, odnosno na korpusu školskih zadaća učenika osnovnih i srednjih škola. Korpus je činilo ukupno 370 osnovnoškolskih te 131 srednjoškolska učenička školska zadaća, napisanih u 11 osnovnih škola, 5 gimnazija i 1 strukovnoj školi iz 15 hrvatskih gradova i mjesta u školskim godinama 2000./2001. i 2001./2002.⁹

Cilj je bio dobiti relativno pouzdan popis leksema osnovnoškolskih i srednjoškolskih učenika u kojima se javljaju fonemi/grafemi č, ć, dž, đ, a prilikom čijeg je pisanja došlo do većeg broja odstupanja.¹⁰

Rezultati odstupanja prilikom pisanja grafema č, ć, dž, đ u školskim zadaćama učenika **osnovnih škola**, što je grafički prikazano na slici 1, jesu ovi:

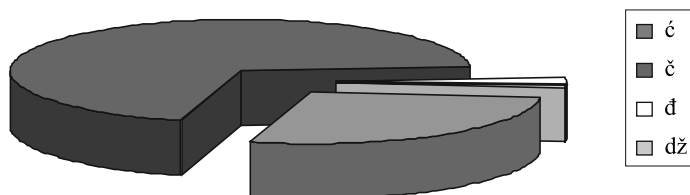
1. umjesto fonema/grafema ć trebalo je upotrijebiti fonem/grafem č (odstupanje se javlja u 46 riječi: 0,12 odstupanja po šk. zadaći)
2. umjesto fonema/grafema č trebalo je upotrijebiti fonem/grafem ć (odstupanje se javlja u 111 riječi: 0,30 odstupanja po šk. zadaći)
3. umjesto fonema/grafema đ trebalo je upotrijebiti fonem/grafem dž (odstupanje se javlja u 4 riječi: 0,01 odstupanje po šk. zadaći) (*đemper* – 4 puta)
4. umjesto fonema/grafema dž trebalo je upotrijebiti fonem/grafem č (odstupanje se javlja u 1 riječi: 0,002 odstupanja po šk. zadaći) (*kovrdžava kosa*)
5. umjesto fonema/grafema đ trebalo je upotrijebiti fonem/grafem č (odstupanje se javlja u 1 riječi: 0,002 odstupanja po šk. zadaći) (*narandasti*).

9 Osnovne škole: Čakovec, Kamešnica, Otok Dalmatinski, Ivana Kukuljevića Sakcinskog (Sisak), Vladimira Nazora (Ploče), Velika Mlaka (Podturen), Silvija Strahimira Kranjčevića (Lovreć), Luka (Sesvete), Pujanke (Split), Marija Bistrica.

Srednje škole: Gimnazija Frana Galovića (Koprivnica), Medicinska škola (Bjelovar), Gimnazija Antuna Gustava Matoša (Zabok), Gimnazija Ogulin (Ogulin), Gimnazija (Zagreb), Opća gimnazija Tituša Brezovačkog (Zagreb).

10 U vezi s tim ni razlika u broju proučenih osnovnoškolskih i srednjoškolskih školskih zadaća za istraživanje nije od važnosti.

OŠ



Slika 1. Odstupanja (pogrešno napisani grafemi) u školskim zadaćama učenika osnovnih škola

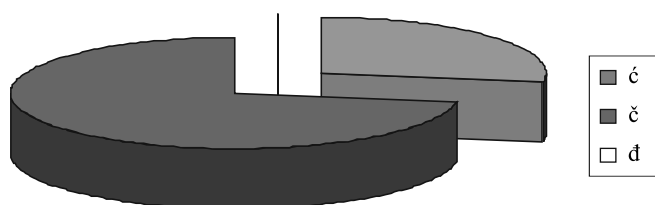
Popis leksema kod kojih se odstupanja javljaju više puta:

budučnost (2),¹¹ mačuhica (2), mogućnosti (2), lišće (od list) (2), liječnik (2), očiju (2), pomoć (2), proči (2) vračala (mi se slika) (2), doči (2), drveće (2), svečano (2), učim, naučili, učioni (3), vočki (3), več (4), vrtič (5), počeli (5), zadačnica, zadača (5), moči (6), osječa, osječala, osječaj (12), kuća, kućica, kučerak (14), ču, če, ćemo, čete, neče, nećemo (18)

Rezultati odstupanja prilikom pisanja grafema *č*, *ć*, *dž*, *đ* u školskim zadaćama učenika **srednjih škola**, što je grafički prikazano na slici 2, jesu ovi:

1. umjesto fonema/grafema *ć* trebalo je upotrijebiti fonem/grafem *č* (odstupanje se javlja u 21 riječi: 0,16 odstupanja po šk. zadaći)
2. umjesto fonema/grafema *č* trebalo je upotrijebiti fonem/grafem *ć* (odstupanje se javlja u 55 riječi: 0,42 odstupanja po šk. zadaći)

SŠ



Slika 2. Odstupanja (pogrešno napisani grafemi) u školskim zadaćama učenika srednjih škola

11 U zagradi je naveden broj odstupanja, odnosno podatak o tome koliko je puta u svim osnovnoškolskim zadaćama navedeni leksem nepravilno napisan.

3. umjesto fonema/grafema *đ* trebalo je upotrijebiti fonem/grafem *dž* (odstupanje se javlja u 1 riječi: 0,0007 odstupanja po šk. zadaći) (*tinejđerku*).

Popis leksema kod kojih se odstupanja javljaju više puta:

budućnost (2), mući (mučiti) (2), osječa (2), osvećuje (2), pomoći im (2), potjeće (2), većina (2), već (3), će ćemo, ču (4), prihvaća (5), srećom (5)

Broj odstupanja prilikom pisanja leksema *č, ć, dž, đ* pokazuje kako osnovnoškolski i srednjoškolski učenici griješe u ograničenom, relativno malom broju leksema, da se broj odstupanja u srednjoj školi, u odnosu na osnovnu školu, povećava te da i osnovnoškolski i srednjoškolski učenici otprilike dva puta više griješe prilikom pisanja leksema u kojima, u skladu s pravopisnom normom, treba napisati grafem *ć*, nego leksema u kojima treba napisati grafem *č*.

4. Primjena načela čestotnosti u ovladavanju hrvatskom pravopisnom normom

Rezultati istraživanja pokazuju kako je broj odstupanja u kojima se, u leksiku osnovnoškolskih i srednjoškolskih učenika, javljaju fonemi/grafemi *č, ć, dž, đ*, ograničen i relativno malen, što se podudara s utvrđenom čestotnošću tih fonema/grafema u hrvatskome standardnom jeziku.

Dakle, važan razlog zbog kojeg učenici čine češća odstupanja prilikom pisanja grafema *č, ć, dž, đ* u njihovoj je maloj čestotnosti jer je posve prihvatljivo pretpostaviti kako će prilikom njihove upotrebe, zbog male čestotnosti pojavljivanja, učenici činiti veći broj odstupanja, zamjenjujući ih češćim grafemima, što je potvrđeno u činjenici da učenici čine češća odstupanja prilikom pisanja grafema koji imaju manju čestotnost, odnosno prilikom pisanja 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 javljaju, vrlo malena, broj odstupanja u pisanju tih grafema u učeničkim školskim zadaćama velik,¹² osobito s obzirom na činjenicu da učenici, prilikom pisanja školske zadaće, uglavnom mogu izbjeći riječ za koju nisu sigurni kako se piše. Naime, može se pretpostaviti da bi broj odstupanja u nekom drugom, spontano napisanom

12 Sović (1967./1968., 25) upozorava na činjenicu kako je upotreba tih fonema/grafema i "danas velik problem". Nažalost, 40 godina poslije, u ak. godini 2002./2003., studenti IV. godine studija hrvatskoga jezika i književnosti na Filozofskome fakultetu u Zagrebu na 10. mjestu po brojnosti, u neobjavljenom istraživanju koje je provedeno o područjima norme u kojima studenti najčešće čine odstupanja, navode nesigurnost prilikom pisanja grafema *č* i *ć*.

tekstu, za koji učenik ne zna hoće li mu ili neće 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 kojem je kao nastavna tema navedena „Izgovor i pisanje *č, ć, dž, đ, lj, nj, ije/je/e/i*“, a kao obrazovna postignuća uz tu nastavnu temu navode se: „pravilno izgovarati i pisati *č, ć, dž, đ, lj, nj, ije/je/e/i* u češće rabljenim riječima“.

Uvjereni smo kako sustavnija i razrađenija primjena načela čestotnosti, koja podrazumijeva izradu popisa čeških riječi u kojima se navedeni fonemi/grafemi javljaju i popisa čeških riječi u kojima prilikom njihova izgovaranja i pisanja dolazi do odstupanja, može pridonijeti njihovu boljem ovladavanju i uspješnijoj primjeni.

Zato je u *Nastavnom 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 za njihovu upotrebu, ali i o tome da učenici ovladaju pravilnim pisanjem čeških riječi. Tu činjenicu moguće je iskoristiti na početku nastavnoga sata (kao motivaciju), u središnjem dijelu nastavnoga sata i/ili na kraju nastavnoga sata, kao 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 javljaju,¹³ a upravo je u toj skupini leksema, u proučenom korpusu učeničkih školskih zadaća, broj učeničkih odstupanja bio najveći.

Primjena načela čestotnosti¹⁴ u nastavi hrvatskoga jezika trebala bi se, dakle, temeljiti:

1. na isticanju male čestotnosti fonema/grafema *č, ć, dž, đ*
2. na isticanju malog i relativno ograničenog broja leksema u kojima se ti fonemi/grafemi javljaju na početku riječi
3. na isticanju malog i relativno ograničenog broja leksema prilikom čijeg pisanja dolazi do odstupanja u pisanju fonema/grafema *č, ć, dž, đ*
4. na čestom upozoravanju na lekseme za koje je utvrđeno da se prilikom njihova pisanja javlja veći broj odstupanja, kao i na provjeravanju njihova pisanja.

13 Usp. Anić i Silić (2001, 117-123); Babić-Finka-Moguš (2000, 24, 27, 30, 31); Badurina-Marković-Mičanović (2007, 9, 11, 13).

14 Načelo čestotnosti temelji se na istraživanju MacWhinneyja koji je dokazao kako su djeca uspješnija u prepoznavanju i određivanju onih gramatičkih oblika koji su često u uporabi i lako uočljivi (Ljubešić, 1997, 143).

5. Zaključak

U istraživanju je potvrđena činjenica o različitoj čestotnosti fonema/grafema u hrvatskome standardnom jeziku. Analizom suvremenog korpusu novinskih tekstova provedena je analiza čestotnosti fonema/grafema i utvrđene su razlike u odnosu na dosad provedena slična istraživanja. Manja čestotnost fonema/grafema istaknuta je kao jedan od mogućih razloga poteškoća u njihovu pravilnom izgovaranju i pisanju, odnosno kao razlog većeg broja odstupanja od standardnoga jezika, a to je potvrđeno analizom učeničkih odstupanja u školskim zadaćama provedenom u deset osnovnih i šest srednjih škola iz petnaest hrvatskih mjesta. Zbog toga je pretpostavljeno kako češća primjena načela čestotnosti u nastavi hrvatskoga jezika može pridonijeti njihovu boljem ovladavanju i točnijoj primjeni. Češća primjena načela čestotnosti u nastavi hrvatskoga jezika podrazumijevala bi njegovo naglašavanje u *Nastavnome planu i programu* i metodičku prilagodbu tom načelu udžbenika i radnih bilježnica, između ostalog i izradom popisa češćih riječi u kojima se fonemi/grafemi č, ć, dž, đ javljaju i izradu popisa češćih riječi prilikom čijeg su izgovaranja i pisanja odstupanja češća.

LITERATURA

- Anić, V. (1991) Rječnik hrvatskoga jezika. Zagreb: Novi liber.
- Anić, V., Silić, J. (2001) Pravopis hrvatskoga jezika. Zagreb: Novi liber.
- Antoš, A. (1974) Osnove lingvističke stilistike. Zagreb: Školska knjiga.
- Babić, S. i dr. (1991) Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika. Zagreb: Hrvatska akademija znanosti i umjetnosti i Globus, Nakladni zavod.
- Babić, S., Finka, B., Moguš, M. (2000) Hrvatski pravopis. Zagreb: Školska knjiga.
- Badurina, L., Marković, I., Mićanović, K. (2007) Hrvatski pravopis. Zagreb: Matica hrvatska.
- Bekavac, B. (2002) Strojno obilježavanje hrvatskih tekstova – stanje i perspektive. *Suvremena lingvistika*. 53-54. Zagreb; 173-182.
- Bekavac, B. (2005) Strojno prepoznavanje naziva u suvremenim hrvatskim tekstovima. Neobjavljena doktorska disertacija. Zagreb: Filozofski fakultet Sveučilišta u Zagrebu.
- Ljubešić, M. (1997) Jezične teškoće školske djece (oblici, uzroci, posljedice, otklanjanje). Zagreb: Školske novine.
- Maretić, T. (1931) Gramatika i stilistika hrvatskoga ili srpskoga književnog jezika. Zagreb: Naklada Jugoslavenskog nakladnog d. d. „Obnova“.
- Matković, V., Sinković, V. (1970) Teorija informacije. Zagreb: Sveučilište u Zagrebu.
- Moguš, M., Bratanić, M., Tadić, M. (1999) Hrvatski čestotni rječnik. Zagreb: Školska knjiga.
- Nastavni plan i program za osnovnu školu* (2006), Zagreb: MZOS.

- Silberztein, M. (2000) INTEX Manual. Pariz: ASSTRIL.
- Sović, I. (1967/68) Neka zapažanja o učeničkoj pismenosti. *Jezik*, 15. Zagreb; 22-29.
- Šojat, A. (1964) *Standardni jezik i standardne riječi*, *Jezik*, XXII, Zagreb; 103-110.
- Tadić, M. (1996) Računalna obradba hrvatskoga i nacionalni korpus. *Suvremena lingvistika*. 41-42. Zagreb; 603-612.
- Tadić, M. (1998) Raspon, opseg i sastav korpusa hrvatskoga suvremenog jezika. *Filologija*. 30-31. Zagreb; 337-347.
- Vuletić, D. (1991) *Istraživanje govora*. Zagreb: Fakultet za defektologiju Sveučilišta u Zagrebu.

THE INFLUENCE OF THE FREQUENCY OF PHONEMES/ GRAPHEMES ON THE MASTERING OF CROATIAN ORTHOGRAPHY

Tamara Gazdić-Alerić, Marko Alerić and Božo Bekavac

***Summary** - This research considers the methods which have so far been used to establish the frequency of the occurrence of phonemes/graphemes in standard Croatian. Further, it presents a modern computational linguistic methodology designed and used for the same purpose.*

The frequency of the occurrence of phonemes/graphemes is established on the basis of a million-item corpus designed in such a way that its structure reflects the ratios specified in the Croatian National Corpus (HNK). The analysed corpus is made up of three types of modern texts. Out of these, 60% were taken from a diary, 35% from prose texts, and 5% from a weekly newspaper.

After identifying the frequency of the occurrence of phonemes/graphemes in the Croatian language, the authors establish the most common orthographic deviations from the norm of the standard Croatian language. These deviations were identified in a corpus of primary and secondary pupils' schoolwork.

On the basis of the deviations found in the pupils' schoolwork, as well as the information on the frequency of the occurrence of phonemes/graphemes in the Croatian language, the authors propose new methods and procedures in the teaching of the rules of orthography which could contribute to the better mastering of the Croatian orthographic norm.

Key words: *frequency of phonemes/graphemes, orthographic deviations from the norm, better mastering of the orthographic norm.*