

Tomislav Stojanov  
Institut za hrvatski jezik i jezikoslovlje  
Ulica Republike Austrije 16, HR-10000 Zagreb  
*tstojan@ihjj.hr*

## JEZIČNOPOVIJESNI I RAČUNALNOJEZIKOSLOVNI ASPEKTI OPISA I NORMIRANJA PISANJA VODORAVNIH CRTA U HRVATSKOME JEZIKU

Rad opisuje jedan od dvaju interpunkcijskih znakova (središnje crte i navodnici) koji bitno odstupaju od odnosa jednoga znaka za jedno (unikodno) značenje. Dok navodnici imaju višestruke grafeme (točnije, osam) za jedno značenje, središnje crte najčešće imaju dva grafema (kraća i dulja crta) koja pokrivaju čak 11 (unikodnih i latiničnih) crtnih znakova. Dok je kriterij crtne duljine tradicionalno visoko istaknut u pravopisnim priručnicima, u predstavljenoj se kategorizaciji on nalazi tek na šestoj hijerarhijskoj razini. Osim što su u međuvremenu standardizirana dva nova unikodna crtna znaka (*two-em dash* i *three-em dash*, Unicode 6.1, siječanj 2012.), drugačija metodologija i uspoređivanje jezičnopovijesnoga i računalnojezikoslovnoga aspekta proširila je spoznaje o crtnim znakovima u hrvatskome jeziku o kojima je pisano u Portada i Stojanov (2009). Predstavlja se kategorizacija osjetljiva na dihotomiju grafičkoga prikaza i značenja koja sve crtne znakove dijeli u pet hijerarhijskih razina. Između 44 unikodna vodoravna i neprekinuta crtna znaka, podjelom prema tipu, vremenu, funkciji, smjeru i visini, došlo se do 11 latiničnih suvremenih pismovnih vodoravnih središnjih znakova među kojima svaki latinični jezik odabire svoje crtne znakove. Svim se crtnim unikodnim grafemima opisalo značenje i uporaba. S druge strane, crtni se znakovi promatraju u kroatističkoj jezičnopovijesnoj pravopisnoj perspektivi. U odnosu na bogati repozitorij unikodno standardiziranih crtnih znakova utvrdilo se da je pravopisna norma bitno redukcijaska. Pravopisno normiranje crtnih znakova podijelilo se u dva razdoblja i tri skupine, ovisno o grafemskome obliku (prva i druga generacija pravopisnih pri-

ručnika) i nazivlju (prijestandardna faza i dva standardna smjera normiranja ovisno o prihvaćanju terminoloških parova *spojnica – crtica* i *crtica – crta*). Na temelju jezičnopovijesnoga i računalnojezikoslovnoga poredbenog istraživanja te na temelju supostavljanja unikodne standardizacije crnih znakova pravopisnoj tradiciji opisa središnjih crta željelo se ukazati na (i) potrebu za širim i interdisciplinarnim pristupom opisa pisane jezične prakse, (ii) nedovoljnost opisa školske razine pravopisnih priručnika za suvremeno pisanje, kao i na (iii) nedostatnost postojeće kroatističke kodifikacije obaju terminoloških smjerova. Da bi se pravopisni priručnici mogli nazvati znanstveno utemeljenim djelima, u znatnijoj bi mjeri trebali opisati računalno pisanje i na razini interpunkcije uvesti razlikovanje znaka i grafema. Jedno od takvih mjesta opisa prema kojima bi pravopisi mogli unaprijediti svoju tehnološku suvremenost jest pitanje pisanja spojnice na početku prelomljena retka o čemu se iznijelo osam argumenata za odbacivanje aktualne tradicije. Raščlamba je pokazala da ima opravdanosti da se crna kodifikacija temelji na trima ili četirima znakovima koji se iz 11 unikodnih latiničnih znakova svode uspostavljanjem osnovnih skupina središnjih crta, radno nazvanih  $c_1$ ,  $c_2$ ,  $c_3$  i  $c_4$  kao najkraća, srednje duga, duga i jako duga središnja crta.

## Kategorizacija crnih znakova

U ovome poglavlju istražuju se crni znakovi.<sup>1</sup> Iako naslov rada upućuje na pravopisne crne znakove, u skladu s računalnojezikoslovnom metodologijom i perspektivom tema će biti znatnije proširena. Crtni se znakovi i njihova uporaba u jezikoslovnoj literaturi primarno nalaze opisani u pravopisnim priručnicima čijom su formom oni bitno ograničeni, što namjenom, a što svojom ciljanom (školskom) publikom. U odnosu na pravopisnu deskripciju i preskripciju, ova kategorizacija sa svojim pojmovnim pristupom ne može biti izravno pravopisnonormativno uspoređivana i evaluirana, ali može jezičnoj normativistici pružiti uvid u širinu interdisciplinarne problematike i možda utjecati na stvaranje drugačijih tipova jezičnih priručnika.

U opisu crnih znakova kreće se od standarda ISO/IEC 10646 ili unikoda (eng. *Unicode*, [www.unicode.org](http://www.unicode.org)) koji razvija *Unicode Consortium* od 1991., a predstavlja jedan od važnih standarda u svijetu računalstva, ali i ljudske tehnologije uopće. Popularno bi se moglo reći da je ono što je Opća deklaracija o pravima čovjeka za Ujedinjene narode, to je unikod za jezike i pisma. Temeljni mu je cilj da svaka kodna adresa ili znakovni indeks jednoznačno upućuje na određeni znak bez obzira na to o kojoj se zemlji, tipkovnici, jeziku, pismu

---

<sup>1</sup> Crtni se znakovi mogu jednostavnije nazvati crtama. U radu se gdješto navodi naziv *crtni znak* samo u kontekstu gdje se željelo jasno razjednačiti plan sadržaja od plana izraza.

ili vrsti grafema radi (slova, brojke, razgodci, matematički, glazbeni, tehnički i opći znakovi ili simboli). Unikod standardizira 123 pisma, a od inačice 4.1 (2005.) obuhvaća i glagoljično pismo (tom je prigodom standardizirano i koptsko i staroperzijsko pismo). Od inačice 5.0 (sredina 2006.) uvršteno je feničko i druga stara pisma (N'Ko, Phags-Pa...), a znatnija se nadogradnja dogodila s inačicom 6.0 (listopad 2010.), kada su uvršteni i standardizirani prometni, kartografski, alkemijski i emotikonski znakovi. Novi grafemi pojavljuju se iz godine u godinu, pa se tako standardizacijom turske lire u novčanome prometu Turske pojavila potreba za uvrštavanjem znaka **₺**, a koji je ubačen u unikodnu inačicu 6.2 (listopad 2012.). Zadnja inačica 7.0 (lipanj 2014.) standardizirala je točno 113 021 znak iz drevnih i suvremenih pisama ljudske civilizacije. Neka pisma još čekaju standardizaciju (npr. znakovno pismo, majanski, afakski), a neka se tek trebaju protumačiti i kodificirati (npr. meroitsko pismo, rongo-rongo). Umjetna pisma bez stvarne uporabe, kao što je klingonsko ili feringijsko pismo iz znanstvene fantastike, odbijena su u unikodnoj standardizaciji, ali jesu unikodno opisana unutar tzv. područja za osobnu uporabu. Za lipanj 2015. najavljena je inačica 8.0, koja će donijeti novih 7716 znakova, 6 pisama i nove simbole, među kojima se nalaze i rasno raznoliki emotikoni.

Budući da temeljna jedinica unikoda nije grafem (najmanja razlikovna jedinica nekoga slovopisa ili pisma) nego znak (eng. *character*) kao najmanja jedinica pisanoga jezika koja ima semantičku vrijednost, odatle interes jezične struke za njega. Više o određivanju jezikopisnih jedinica i pojmova vidi u Stojanov (2015).

Kategorizacija crtnih znakova iz nastavka nije unikodno zadana, ona je samo unikodno motivirana. Unikodno zadana kategorizacija dijeli svih 113 012 znakova u 30 razreda od kojih je 5 razreda slova (eng. *letters*), 3 razreda brojaka (eng. *digits*), 7 razreda razgodaka (eng. *punctuation*), 4 razreda simbola (eng. *symbols*), 3 razreda razdjelnika (eng. *separators*), 3 razreda obilježja (eng. *marks*) i 5 razreda *ostale* (eng. *others*). Brojevi se, primjerice, dijele u decimalne, slovne (npr. rimska slova koja služe za izražavanje brojaka) te *ostale* brojke. Simboli se dijele na matematičke, novčane, modifikacijske i *ostale*. Razgodci se dijele u konektorske (10 znakova), crtne (24 znaka), otvarajuće (75 znakova, npr. otvarajuće zagrade), zatvarajuće (73 znaka, npr. zatvarajuće zagrade itd.), razgotke za početak navođenja (12 znakova), razgotke za završetak navođenja (10 znakova) te *ostale* razgotke (484 znaka). Ova podjela 688 razgodaka, koliko god bila razrađena, proizišla je iz računalnoinfrastrukturnoga gledanja na jezik te ima ograničavajuće čimbenike – prvi joj nedostatak dolazi iz povijesno-tehnološkoga aspekta. Kategorizacija je nastajala u dužemu vremenu pa se radi

sukladnosti s postojećim znakovima morala nadograđivati, a ne bitnije modifikirati u neku možda bolju podjelu. Tako su se u najranijemu razdoblju unikoda standardizirala dva okomita prezentacijska crtna znaka s danas upitnom valjanošću<sup>2</sup>, ali kako se zbog potrebe tehnološke uskladivosti znakovi ne brišu, oni trajno ostaju kao spomenici vremenu. Drugi je nedostatak pragmatična okrenutost – standard nije ni zamišljen kao univerzalno teorijsko i objasnidbeno dostignuće, nego kao tehnologija i praktična računalna primjena. Tako se računalnoinfrastrukturni pristup nomenklaturi zadovoljava brojnim klasifikacijama razreda *ostali* u kojima se nalazi većina razgodačnih znakova. Također, u razgodačne crtno znakove svrstavaju se i znakovi koje se ne bi načelno kategoriziralo kao takve, npr. *hyphen with diaeresis* U+2E1A (*spojnica s vodoravnim dvotočjem*). *Minus sign* U+2212 (*znak minusa*) kao i valovita crta (za razliku od kose crte) ne smatraju se crtnim znakom.

Iz svih tih razloga pristupilo se novoj i ciljanoj klasifikaciji u kojoj su se promatrali samo vodoravni i neprekinuti crtni znakovi, bez obzira na njihov položaj, duljinu ili vrstu. Iako imaju unikodnu vrijednost, crtnim znakovima ne smatraju se znakovi iz skupine *combining marks* (*spajajuće oznake*), *modifier letters* (*slova za modifikaciju*) i *modifier symbol* (*znakovi za modifikaciju*) jer to, zapravo, nisu samostalni grafemi, nego oni koji utječu na drugi znak.<sup>3</sup>

Dva su dodatna razloga ili motiva za provedeno istraživanje: (i) ispitivanje mogućnosti i dosega računalnojezikoslovne klasifikacije jedne vrste grafemā izdvojenih unutar univerzalnoga popisa postojećih znakova, te (ii) doprinos jezičnoj normativistici na metodološkoj i teorijskoj razini u možebitnome stvaranju nove vrste jezičnih priručnika.

Prema određenim kriterijima crtnih je znakova u unikodu ukupno 44 razvrstanih u osam razreda (ukupan broj razreda u unikodu je 30)<sup>4</sup>.

---

<sup>2</sup> *Presentation form for vertical em-dash* (U+FE31) i *Presentational form for vertical em-dash* (U+FE32).

<sup>3</sup> Neki od njih su *modifier letter minus sign* [U+02D7] i *katakana-hiragana prolonged sound mark* [U+30FC] koji se rabe u fonetici te *combining minus sign below* za grafijsku modifikaciju.

<sup>4</sup> <http://www.fileformat.info/info/unicode/category/index.htm>

1. *aegean number ten* [U+10110] – No
2. *armenian hyphen* [U+058A] – Pd
3. *balinese musical symbol left-hand open pang* [U+1B78] – So
4. *box drawings heavy horizontal* [U+ 2501] – So
5. *box drawings heavy left* [U+2578] – So
6. *box drawings heavy right* [U+257A] – So
7. *box drawings light horizontal* [U+2500] – So
8. *box drawings light left* [U+2574] – So
9. *box drawings light right* [U+2576] – So
10. *canadian syllabics final short horizontal stroke* [U+1428] – Lo
11. *CJK ideograph, first* [U+4E00] – So
12. *counting rod unit digit one* [U+1D360] – No
13. *dashed low line* [U+FE4D] – Pc
14. *dashed overline* [U+FE49] – Pc
15. *em dash* [U+2014] – Pd
16. *en dash* [U+2013] – Pd
17. *figure dash* [U+2012] – Pd
18. *fullwidth hyphen-minus* [U+FF0D] – Pd
19. *hangul jungseong eu* [U+1173] – Lo
20. *heavy minus sign* [U+2796] – So
21. *hebrew punctuation maqaf* [U+05BE] – Pd
22. *horizontal bar* [U+2015] – Pd
23. *horizontal line extension* [U+23AF] – Sm
24. *hungul letter eu* [U+1173] – Lo
25. *hyphen* [U+2010] – Pd
26. *hyphen bullet* [U+2043] – Po
27. *hyphen-minus* [U+002D] – Pd
28. *low line* [U+005F] – Pc
29. *minus sign* [U+2212] – Sm
30. *mongolian todo soft hyphen* [U+1806] – Pd
31. *monogram for yang* [U+268A] – So
32. *musical symbol one-line staff* [U+1D116] – So
33. *non-breaking hyphen* [U+2011] – Pd
34. *paragraphos* [U+2E0F] – Po
35. *roman uncia sign* [U+10191] – So
36. *small em dash* [U+FE58] – Pd
37. *small equals sign* [U+FE66] – Sm
38. *small hyphen-minus* [U+FE63] – Pd
39. *soft hyphen* [U+00AD] – Cf
40. *straightness* [U+23E4] – So
41. *subscript minus* [U+208B] – Sm
42. *superscript minus* [U+207B] – Sm
43. *three-em dash* [U+2E3B] – Pd
44. *two-em dash* [U+2E3A] – Pd

Nazivi na engleskome izvorno su preuzeti, u uglatim je zagradama unikodna oznaka, a u nastavku unikodna klasifikacijska oznaka razreda. Oznake predstavljaju sljedeće razrede uz podatak o zastupljenosti među 44 znaka:

- Punctuation, Dash [Pd] – 15
- Symbol, Other [So] – 13
- Symbol, Math [Sm] – 5
- Punctuation, Connector [Pc] – 3
- Letter, Other [Lo] – 3
- Punctuation, Other [Po] – 2
- Number, Other [No] – 2
- Other, Format [Cf] – 1

Dakle, najveći broj crtnih znakova služi za crtnu interpunkciju (15), a nakon toga kao znakovi za *ostale* potrebe (13). Slijedi 5 crtnih znakova za matematičko označavanje, 3 znaka za interpunkcijsku konekciju, 3 znaka za *ostalo* slovno obilježavanje i za *ostalu* interpunkciju te 2 crtna znaka za *ostalo* brojčano obilježavanje. Jedan znak služi za potrebe formatiranja teksta.

Pokušaju li se navedena 44 znaka rekategorizirati za potrebe ovoga istraživanja, dolazi se do pet hijerarhijskih razina. Prva je podjela po **tipu** i ona dijeli latinične i nelatinične crtno znakove.

Jedanaest je nelatiničnih crtnih znakova, a dio ih se odnosi na drevna<sup>5</sup>, a dio na suvremena pisma: *aegean number ten*, *armenian hyphen*, *balinese musical symbol left-hand open pang*, *canadian syllabics final short horizontal stroke*, *CJK ideograph first*, *counting rod unit digit one*, *hangul jungseong eu*, *hebrew punctuation maqaf*, *hungul letter eu*, *mongolian todo soft hyphen* i *paragraphos*.

Druga hijerarhijska podjela je po kriteriju **vremena** i dijeli crtno znakove na drevne latinične i suvremene latinične znakove. Samo je jedan drevni latinični crtni znak – *roman uncia sign*, a predstavlja stari rimski znak za duljinu.

Treća je razina **funkcije**. Crtni znakovi mogu se rabiti kao:

- crtežni ili grafički znakovi (6 onih koji započinju u nazivu s *box drawings* i *heavy minus sign*)
- opći, strukovni i tehnički simboli (*monogram for yang*, *musical symbol one-line staff*, *straightness*, *horizontal line extension*)
- pomoćni računalni operatorski znakovi (*small hyphen-minus*, *small equals sign*, *fullwidth hyphen-minus*, *small em dash*)
- nabrajajući znakovi (*hyphen bullet*)
- znakovi za zapisivanje prirodnoga jezika ili *pismovni* znakovi<sup>6</sup>

Važna napomena tiče se znaka *minus sign*. Po svojoj naravi pripada matematici te bi trebao biti svrstan među opće, strukovne i tehničke simbole, međutim zbog potrebe ovoga istraživanja i razjednačivanja u odnosu na druge bližnje razgodačne znakove (spojnica, crtica i ini) on će se iznimno navesti pod pismovne znakove. Slično je s hibridnim znakom *hyphen-minus* koji se rabi i kao interpunkcijski i kao matematički znak.

---

<sup>5</sup> Tako je, primjerice, *aegean number ten* znak iz etruščanskoga brojevnog sustava iz 5. st. pr. n. e., *counting rod unit digit one* znak iz drevna kineskoga brojevnoga sustava koji se rabio više od 2000 godina, a *paragraphos* starogrčki tekstni simbol.

<sup>6</sup> Pismovni (crtne) znakovi mogu se nazvati i *pismeni (crtne) znakovi*, ali se prednost daje prvo rješenju u želji da se jasnije naznači korijen „pismo”.

Ova razina pismovnih crtnih znakova obuhvaća 16 znakova. Njihov hiperonimni naziv je **crta**. Razloži zašto se govori o crtama, a ne o crticama u hrvatskome jeziku jesu sljedeći: (1) zato jer je kosa crta, a ne kosa crtica; (2) zato što je okomita crta, a ne okomita crtica; (3) zato što se duga crta (eng. *horizontal bar* ili —)<sup>7</sup> teško može nazvati *crticom* (a obrnuto funkcionira); (4) zato što je crtica jedan od pojavnih oblika crta i ima zauzeti terminološki izraz.

Četvrta je razina **smjera**: vodoravni, okomiti i kosi crtni znakovi. Zbog velikoga broja okomitih i kosih znakova, u ovome se radu oni nisu uvrštavali u unikodnu klasifikaciju pa tako svih 16 preostalih znakova iz treće razine pripadaju u vodoravne crtno znakove.<sup>8</sup>

Podcrta (eng. *underline*) je (**grafički**) **prikaz** za obilježavanje teksta istaknutim. Riječi se mogu podcrtavati, precrtavati i nadcrtavati (a ne \*podvlačivati, \*prevlačivati i \*nadvlačivati) iz čega se izvode nazivi *podcrta*, *precrta* i *nadcrta*.

Peta se hijerarhijska razina temelji na kriteriju **visine**, i to na donje, središnje i gornje crtno znakove. Donji i gornji znakovi su *dashed low line*, *dashed overline*, *low line*, *subscript minus* i *superscript minus*.

Nasuprot grafičkome prikazu stoji (**semantički**) **znak**. Za razliku od grafičkoga prikaza i njemu odgovarajućih naziva, semantički terminološki ekvivalent je podvlaka (eng. *underscore*; *low line* ili  ). Ona se rabi često u računalstvu za obilježavanje bjeline (*moja\_datoteka.txt*), a u općemu jeziku ukazuje na prazno mjesto grafema (npr. *Rek\_ i 5*)<sup>9</sup>. Suprotno podvlaci je nadvlaka (eng. *macron*) kao semantički znak koji u jezikoslovlju obilježuje dijakritički znak za duljinu (̄), srednji ton i drugo, a u farmakologiji skraćivanje zapisa latinskih riječi, npr. *ā* za lat. *ante* itd. U raspravi u kojoj nije nužno potrebno razlikovati plan (grafičkoga) prikaza i (semantičkoga) znaka, nazivi podvlaka i nadvlaka mogu se zamijeniti uobičajenim *gornja crta* i *donja crta*.

Tvorbeno analogno podvlaci i nadvlaci (a dalje uvlaci i izvlaci kao dokumentnim grafičkim oblikovanjima), mogla bi stajati povlaka kao hiperonimni naziv za ukupnost semantičkih znakova vodoravnih ravnih središnjih crta ili,

---

<sup>7</sup> Ona koja se rabi za odvajanje cjelina ili pobrojavanje.

<sup>8</sup> Umjesto *okomita crta* predložen je naziv *uspravna crta*, međutim, radi terminološke analogije *okomiti* – *vodoravni*, odnosno *uspravni* – *položeni*, ostaje se pri prvotnome nazivu.

<sup>9</sup> Ovaj je primjer iz crne kronike dnevnih novina gdje je policija tragala za osobom s natpisom na majici čije jedno slovo nije bilo čitljivo pa se umjesto toga koristila podvlaka.

sukladno dosad navedenoj kategorizaciji, **latiničnih suvremenih pismovnih vodoravnih središnjih znakova**. U diskusiji sa stručnjacima taj naziv nije bio prihvaćen iako bi mogao biti terminološki zanimljiv.<sup>10</sup> U daljnjemu tekstu, dakle, rabit će se naziv *središnja crta* koji tako stoji u opreci prema *gornjoj* i *donjoj crti*. Potrebno je uspostaviti zajednički naziv jer su dosadašnji izrazi *crte* ili *crtice* višeznačni i nedovoljno precizni. Osim toga, njegovim se uvođenjem nazivima *crta* i *crtica* oslobađa semantičko mjesto čime će ih se moći upotrijebiti za imenovanje konkretnih crtnih oblika.

Nakon pet razina broj crtnih znakova sveo se na 11. Abecedno posloženi to su:

1. *em dash*
2. *en dash*
3. *figure dash*
4. *horizontal bar*
5. *hyphen*
6. *hyphen-minus*
7. *minus sign*
8. *non-breaking hyphen*
9. *soft hyphen*
10. *three-em dash*
11. *two-em dash*

Tih jedanaest središnjih crta dalje se može dijeliti u skupine ovisno o kriteriju **prijeloma**:

- znakovi koji omogućuju prijelom: *hyphen-minus*, *hyphen*, *soft hyphen*;
- znakovi koji ne omogućuju prijelom: *non-breaking hyphen*, *minus sign*, *figure dash*, *en dash*, *em dash*, *horizontal bar*, *two-em dash*, *three-em dash*.

Znakovi koji omogućuju prijelom dijele se prema kriteriju vidljivosti:

- vidljivi: *hyphen-minus*, *hyphen*;
- nevidljivi: *soft hyphen*.

---

<sup>10</sup> Povelaka je hrvatski dijalektizam i arhaizam te terminološki nezauzet izraz. U srpskome povelaka je istoznačnica sa spojnicom, ali je taj naziv napušten (u korist *crtice*). Dakle, potencijalna uporaba naziva povelake u hrvatskome obuhvaćala bi sve ravne središnje crte, a ne samo jedan pojavni oblik kao što je to bilo donedavno u srpskome nazivlju.



Osim po prijelomu, deset vidljivih latiničnih suvremenih pismovnih vodoravnih središnjih znakova (dakle, svi osim *soft hyphena*) mogu se nazvati i slovnima te dalje dijeliti i grupirati prema kriteriju **uobičajene crtnje duljine u fontovima**. Namjerno se ističe *uobičajena duljina* jer podjela prema duljini nije egzaktna i temelji se na grafičkome prikazu u raširenim fontovima. Drugim riječima, kategorija je duljine neprecizna i relativna zbog grafičke uvjetovanosti. Binarna podjela na kratku i dugu središnju crtu također nije precizna jer neki fontovi nemaju ni tu osnovnu opreku. Unatoč svojoj relativnosti, kriterij duljine snažno je prihvaćen u pravopisnim priručnicima pa je tu podjelu nužno uvesti i na ovome mjestu.

Kao što se navelo, dva su znaka vrlo nova (*two-em dash* i *three-em dash*) i u vrijeme pisanja teksta niti jedan zadani font u sustavu Windows 8.1 nije ih bio u stanju prikazati pa se na njihovu mjestu pojavljivala oznaka ☐. Razredba središnjih crta prema duljini uspostavljena je prema ključu da određeni font mora imati svih preostalih osam grafema za osam znakova čime su se isključili fontovi koji su dio grafema preuzimali iz drugoga znakovnika. Od 290-ak zadano instaliranih fontova samo su četiri imala svih osam grafema: Lucida Sans Unicode, Arial Unicode MS, MS UI Gothic i Cambria. Drugi su popularni fontovi preuzimali jedan ili dva grafema: zadani font u Wordu, Calibri, nema svoj prikaz znaka *non-breaking hyphen*, široko korišteni font na kojemu se piše ovaj tekst, Times New Roman, nema grafeme *hyphen* i *non-breaking hyphen*, a popularni i znatno obogaćeniji font za potrebe jezikoslovne struke (rabi se u Institutu za hrvatski jezik i jezikoslovlje), Zrcola, nema grafeme za *non-breaking hyphen* i *figure dash*.<sup>11</sup>

U nastavku je prikazano tih osam znakova u navedenim fontovima veličine 12 točaka i s umetnutim bjelinama. Poredak je znakova *hyphen-minus*, *hyphen*, *non-breaking hyphen*, *figure dash*, *en dash*, *minus sign*, *horizontal bar*, *em dash*:

- ----- — Lucida Sans Unicode
- ----- Arial Unicode MS
- ----- MS UI Gothic
- ----- Cambria

---

<sup>11</sup> Broj podržanih grafema mijenja se s vremenom. Tijekom 2008. i rada u sustavu Windows XP Portada i Stojanov (2009) naveli su da Times New Roman nije imao tri grafema (sada nedostaju dva, dok je u međuvremenu uveden *figure dash*), a Cambriji su nedostajala dva grafema (*non-breaking hyphen* i *figure dash*) koja su nekoliko godina kasnije uvrštena.

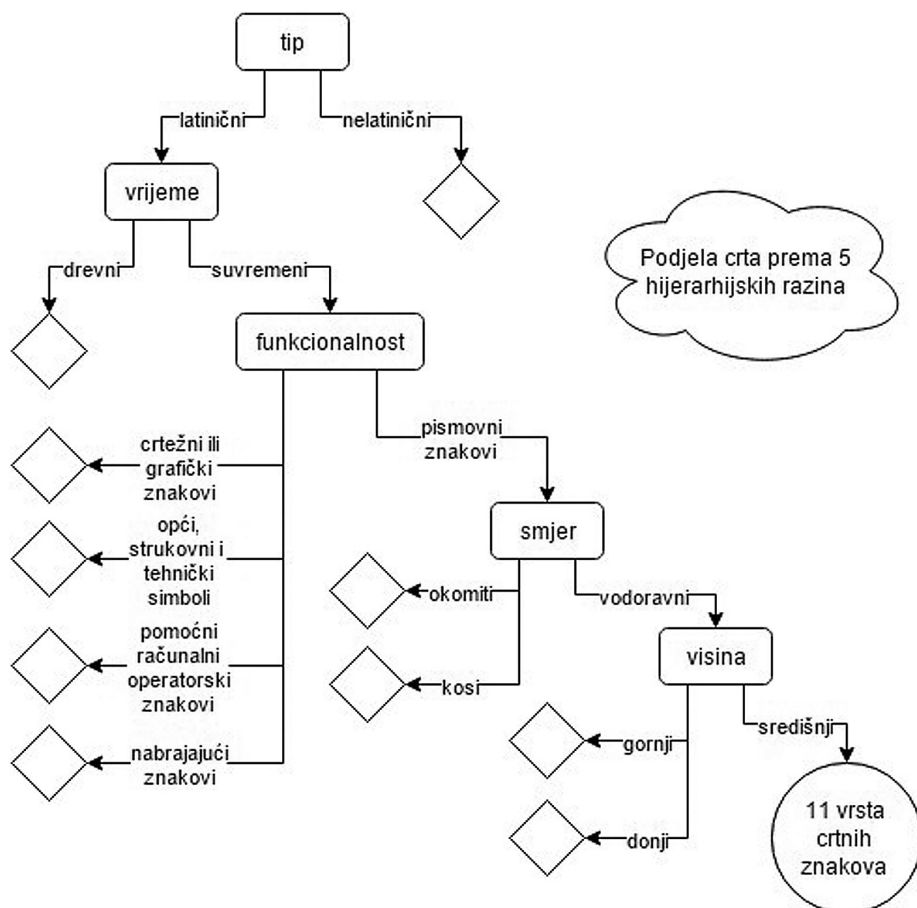
Kao što se vidi, fontovi vrlo različito prikazuju znakove, pa tako MS UI Gothic ima uočljivo kraći *em dash* od *horizontal bara*, a Arial Unicode MS obratno. Lucida Sans Unicode i MS UI Gothic imaju *hyphen-minus* koji je bitno različit od *hyphena* i duljinom bliskiji duljini *en dasha*, dok Arial Unicode MS i Cambria imaju *hyphen-minus* identičan *hyphenu*. Na osnovi takvih razlika teško je uspostaviti jednoznačnu i *zadanu* duljinu središnjih crta. Unatoč odstupanjima, može se napraviti okvirna vizualna podjela osam središnjih crta u tri skupine: (i) najkraće, (ii) srednje duge i (iii) duge. Četvrta skupina (iv) jako dugih središnjih crta (*two-em dash* i *three-em dash*) još nema grafemski prikaz, ali se i one navode u zasebnoj skupini. Na stranicama standarda Unicode može se vidjeti da predstavljaju dvije (—), odnosno tri duljine (——) jednoga *em dasha*, što ih čini uvjerljivo najduljim unikodnim grafemima.

U ovoj je podjeli *horizontal bar* svrstan među duge, a ne srednje duge središnje crte, dok je *hyphen-minus* postavljen u kratke, a ne srednje duge središnje crte.

1. <i>hyphen-minus</i> (U+002D)	}	najkraći (radni naziv $c_1$ )
2. <i>hyphen</i> (U+2010)		
3. <i>non-breaking hyphen</i> (U+2011)		
4. <i>figure dash</i> (U+2012)	}	srednje dugi (radni naziv $c_2$ )
5. <i>en dash</i> (U+2013)		
6. <i>minus sign</i> (U+2212)		
7. <i>horizontal bar</i> (U+2015)	}	dugi (radni naziv $c_3$ )
8. <i>em dash</i> (U+2014)		
9. <i>two-em dash</i> (U+2E3A)	}	jako dugi (radni naziv $c_4$ )
10. <i>three-em dash</i> (U+2E3B)		

Dakle, latinični suvremeni pismovni vodoravni središnji znakovi prema kriteriju uobičajene crne duljine u fontovima dijele se u najkraće, srednje duge, duge i jako duge. Za potrebe daljnjega istraživanja rabit će se radni nazivi  $c_1$ ,  $c_2$ ,  $c_3$  i  $c_4$  jer će se, nakon poglavlja o crnoj tipologiji i značenju, raspravljati o njihovoj uporabi u hrvatskoj pravopisnoj historiografiji i u hrvatskome jeziku danas.

Grafički prikaz hijerarhijskih razina izgledao bi ovako:



Slika 1. Podjela crta prema pet hijerarhijskih razina

## Tipološki pregled oblika središnjih crta i njihova značenja

U poglavlju se opisuje 11 središnjih crta, a na mjestima gdje je to moguće uz engleske nazive koristit će se i hrvatski nazivi.

### Hyphen-minus

Ovaj znak ima unikodnu vrijednost U+002D, a zbog tehnološki anakronih razloga ograničenosti kodnoga prikaza ASCII-ja (eng. *American Standard Code for Information Interchange*) dugo je vremena bila jedina središnja crta.

Nalazi se na svakoj tipkovnici kraj desnoga *shifta* kao i na numeričkome dijelu tipkovnice, što *hyphen-minus* čini uvjerljivo najkorištenijom vrstom središnje vodoravne crte u svim računalno stvorenim tekstovima. U svijetu kodiranja, skriptiranja i programiranja jedini je ispravni znak. Grafički se ne razlikuje od *hyphena*, a kod nas se često prevodi kao *minus-crtica*. U vrijeme nastajanja ASCII-ja 1970-ih broj mjesta predviđen za alfanumeričke znakove i razgovor bio je vrlo ograničen, pa je dugo postojala samo jedna središnja crta koja je služila i za pravopisne i za numeričke potrebe, prema čemu je nastao naziv *hyphen-minus*. Da je tada postojala svijest o potrebi kvalitativnoga pristupa u opisu prirodnih jezika ili da je tehnologija jednostavno bila razvijenija, danas bi se vrlo vjerojatno na tipkovnici nalazilo više od jedne središnje crte ili bi, u još optimističnijemu scenariju, središnja crta s numeričkoga dijela tipkovnice bila različita od one kraj desnoga *shifta*.

## Hyphen

Znak *hyphen* ima unikodnu vrijednost U+2010 i predstavlja znak za središnju crtu. Budući da je nastao nakon ASCII-ja, ne postoji na tipkovnici. Često se prevodi kao *spojnica*. Najčešće se grafički ne razlikuje od *hyphen-minusa*.

## Minus sign

Ovaj znak s vrijednošću U+2212 označuje operaciju oduzimanja ili negativna predznaka. Grafički se razlikuje od *hyphen-minusa* jer je visina i dužina ove središnje crte razmjerna visini i dužini znaka plus (samo bez okomite crte) te je korištenje pravilnoga znaka važno u udžbenicima iz prirodnih i tehničkih znanosti. Pedantniji korektori i lektori uočavaju razliku - +, odnosno – +. Prevodi se kao *znak minusa*. U svijetu programiranja ne smije se koristiti jer ga kompajleri i interpreteri ne smatraju znakom s matematičkom vrijednošću.

## Non-breaking hyphen

Znak *non-breaking hyphena* često se prevodi kao *neprelamajuća spojnica* (ili *crtica*), ima vrijednost U+2011, a temeljna je razlika u odnosu na *hyphen* to što ima funkciju da se na njegovu mjestu ne događa prijelom retka.<sup>12</sup>

---

<sup>12</sup> Naravno, aplikacije moraju biti dovoljno napredne da se znaju koristiti tom opcijom. Najčešće je to slučaj kod aplikacija za obradu teksta (Word, Writer i dr.), za razliku od običnih (eng. *plain*) uređivača teksta (Notepad, Notepad++ itd.).

Prijelom (ili kako jezikoslovci često nazivaju „prijenos dijela riječi u novi redak”) *tiče se* jezikoslovne struke i on nema isključivo tipografsku narav. Ako se slogovanjem riječi bave jezikoslovci (svi pravopisi navode pravila o slogovima), onda ih se tiču i znakovi za prijelom. Pitanje prijeloma interdisciplinarno je područje i predmet zanimanja ne samo tipografa i jezikoslovaca nego i računalaca i grafičara (Mesaroš 1971). Tipografi se bave izgledom prijeloma, grafičari tehničkom izvedbom, računalci medijem u kojem se prijelom ostvaruje, a jezikoslovci sadržajem, kao što bi se i očekivalo. Sadržaj prijeloma mjesto je na kojemu se prijelom ostvaruje. Dakle, drži se da jezikoslovna norma mora standardizirati mjesto prijeloma, čime se pokazuje potreba da se znakovi za prijelom obrade i protumače u jezikoslovnim priručnicima, osobito u onima koji razlikuju jezični sadržaj i značenja. To što se *hyphen* i *non-breaking hyphen* grafički ne razlikuju, ne znači da je riječ o istim znakovima. (I znak pravopisne točke identično izgleda na kraju rečenice i na kraju skraćenic, pa se opisuje u oba konteksta različito.)

*Non-breaking hyphen* rabi se kada se želi spriječiti prijelom riječi, npr. kod predmetka *e-* (*e-pošta*), akronimnih padežnih nastavka (INA-e), u dvojnim prezimenima (Jelić-Bosnar) itd.

Prijelom većini korisnika nije zanimljiv za potrebe pravopisa jezika koji se ostvaruje rukom ili u tisku („rukopisni ili tiskovni pravopis”), ali jest važan za računalno pisanje (tipkano pisanje ili tipkopis), koje danas prevladava, te se smatra da bi taj znak trebao biti opisan u priručniku o jezičnoj pismenosti.

## Figure dash

Znak *figure dash* s vrijednošću U+2012 ima funkciju odjelitelja brojevnoga niza radi lakšega razumijevanja i čitanja. Nad sastavljeno pisanim dužim brojevnim nizovima čitatelji se dugo zadržavaju jer ih slabije vizualno percipiraju. Razdvajanje niza ubacivanjem središnje crte bitno doprinosi ne samo čitanju nego i izgovoru. Tako je uobičajeno brojeve brzoglasa, bankovnih računa, knjižničnih signatura i drugih brojevnih oznaka vizualno dijeliti središnjom crtom. Drugim riječima, znak *figure dash* ne dijeli brojevni niz, nego signalizira računalu da je i s lijeve i s desne strane isti podatak te da je njegova svrha isključivo vizualna. Upravo se zbog toga u većini fontova grafički razlikuje od *hyphena* (tj. *hyphen-minus*), koji se često rabi na njegovu mjestu, jer ima širinu brojke čime je znatno duži i tipografski uređeniji. Pedantniji lektori i korektori uočavaju potrebu za njegovom uporabom (iako griješe ubacujući *en dash* s kojim je u većini fontova identičan, ali posve drugim značenjem).

Dakle, umjesto 3783-803 trebalo bi zapisivati 3783–803.

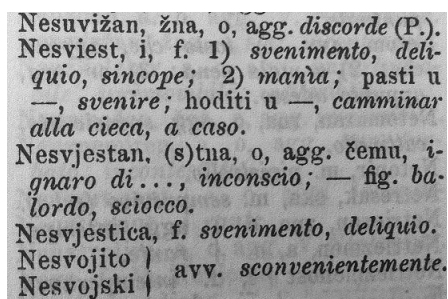
## En dash

Znak s vrijednošću U+2013 drugi je najčešći znak nakon *hyphen-minus* koji se koristi u tipkopisu. Ima brojne nazive u hrvatskome: crtica, crta, dulja crtica, srednja crtica itd., a autori pravopisa IHJJ rabe – crtica. Duljina središnje crte varira u fontovima pa je nazvana *en dash* upravo da se naglasi njezina razmjernost u odnosu na veliko latinično slovo N (kao što se *em dash* odnosi prema velikome latiničnome slovu M). U fontu Times New Roman izgleda ovako: –. Ima značenje dosega, odnosa, razmjera, a hrvatski se pravopisi razlikuju u njezinu korištenju pa neki dosljedno normiraju pisanje s bjelinama (npr. IHJJ 2013), a drugi ne (npr. BMM 2008).

## Em dash

Ova središnja crta ima vrijednost U+2014 i u većini je fontova dulja crtica od *en dash* te razmjerna širini velikoga slova M: —.<sup>13</sup>

Znak predstavlja stanku jaču od zagrada, izostavljanje, obilježavanje navoda itd. Uporaba *em dash* varira od jezika do jezika te ju neki pravopisi uopće ne spominju, a drugi ju koriste. U hrvatskim se suvremenim pravopisima uopće ne spominje, zadnji koji ju je rabio bio je BFM iz 1971. Njezina je uporaba rjeđa u suvremeno doba računalnoga tiskarstva, a prije je bila ne samo znatno češća nego i jedina zamjena za *en dash*, a o čemu će se još raspravljati. Taj su znak koristili mnogi stariji hrvatski leksikografi, npr. Parčić (1901) i JAZU (1880. – 1976.).



Slika 2. Središnja crta *em dash* u Parčićevu Rječniku iz 1901.



Slika 3. Središnja crta *em dash* u Rječniku JAZU iz 1886.

<sup>13</sup> Razmjernost je smjernica, a ne pravilo. Fontografi imaju punu slobodu umjetničkoga i tehničkoga izražavanja pri stvaranju novih fontova. U fontu Times New Roman duljina *em dash* nešto je duža od širine velikoga slova M.

## Two-em dash

Znak s vrijednošću U+2E3A unikodno je kodificiran prije tri godine – u siječnju 2012. – zbog čega ga još nema u fontovima. Kao što mu naziv govori, riječ je o središnjoj crti duljine dva *em dasha* koja se u praksi dobiva spajanjem tih dvaju znakova: —. Značenje mu je izostavljanje riječi.

## Three-em dash

Kao i prethodni znak, kodificiran je u siječnju 2012. te ga također još nema u fontovima. Vrijednost je U+2E3B, a predstavlja najdužu središnju ravnu vodoravnu crtu u latiničnome pismu. Jedini je način ostvarivanja spajanjem triju *em dashova*: ——. Zamišljen je za izostavljanje više riječi ili dijelova teksta, a njegova se primjena može naći i u bibliografiji kod uzastopno ponavljajućih nositelja bibliografskih jedinica, primjerice:

Kovačić, Ivan. O zamjenicama u hrvatskoj književnosti 19. stoljeća. // ...  
———. Pridjevne zamjenice u Matoševim pjesmama. // ...  
———. Komparativna stilaska analiza jezika proze 1880. – 1900. // ...

## Horizontal bar

Znak s vrijednošću U+2015 naziva se i *quotation dash* čime mu se naznačuje glavna primjena. U mnogim je jezicima (francuski, španjolski, ruski i dr.) standardni znak za navode i dijaloge. Engleska Wikipedia ističe da je James Joyce inzistirao kod izdavača svojih djela na stilu pisanja u kojemu je rabio *horizontal bar*. Pedantniji lektori i korektori uočavaju potrebu da rabe ovaj znak, a ne *hyphen-minus* ili *en dash* radi psiholingvističkoga čimbenika brzine čitanja i raspoznavanja zasebnih dijelova ekranskoga teksta te radi estetike. U većini je fontova duljine između *en dasha* i *em dasha*, a zna biti prikazan i nešto debljom središnjom crtom.

Na sljedećim se slikama zorno vidi prednost njegova korištenja za ekran-sku čitljivost.



Slika 4. Korištenje *hyphen-minus*a u pod-slovima



Slika 5. Korištenje *horizontal bara* u pod-slovima za početak dijaloga

## Soft hyphen

Znak *soft hyphen* ima vrijednost U+00AD, a naziva se i *discretionary hyphen*, *optional hyphen* ili SHY. Riječ je o nevidljivoj središnjoj crti (ne vidi se pri tipkanju) koja obilježuje potencijalno hifenacijsko mjesto. Pri prijelomu on će se ostvariti na postavljenome mjestu, ali kada nema prijeloma, riječ je pisana sastavljeno. Potreba za *soft hyphenom* pojavila se tek u suvremeno računalno doba za prikaz teksta u prebirnicima (eng. *browsers*) s različitim razlučivostima i na raznim platformama i medijima gdje tekst mora imati svojstvo plutanja (eng. *float*) u odnosu na druge tekstne okvire (eng. *placeholders*). Iako je znak unikodno standardiziran već duže vrijeme, proizvođači softvera različito ga interpretiraju pa s njegovom uporabom izvan prebiričkoga medija treba biti oprezan.

U sljedećem će se poglavlju opisati oblici i nazivlje središnjih crta u hrvatskim pravopisnim priručnicima koji se njima primarno bave.

## Oblici središnjih crta i njihovo nazivlje u hrvatskim pravopisnim priručnicima

Povijesni pregled središnjih crta načinjen je u Portada i Stojanov (2009), a ovdje se rezultati dopunjuju s pet novih izvora<sup>14</sup>, predočuju u tabličnome prikazu i zaključuju s novim zapažanjima i spoznajama o središnjim crtama u hrvatskim pravopisnim priručnicima.

Daktilografska i stenografska praksa posebno se ne obrađuje, a za nju se može reći da je ograničena svojim tehnološkim rješenjem postojanja samo jedne središnje crte (Mimica i Vunić 1973: 82–83).

---

<sup>14</sup> Uputyenje k slavonskomu pravopisanju (1779.), Naputchenyje za horvatzki prav chteti y pizsati (1808.), Mažuranić (1861), Tkalec (1886), IHJJ (2013).



	oblik $c_1$	oblik $c_2$	oblik $c_3$	nazivlje
Kratki navuk za pravopisznye horvatzko (1779.)	+		+	znaménye vèzanya; preztanye ili paufa
Uputjenje k' slavonskomu pravopisanju (1779.)	+		+	zlamènjè razdilènja; zlamènjè priftájánja ili pocsivánja
Naputchenye za horvatzki prav chteti y piszati (1808.)	+		+	znamenye razdelivanya i znamenye vezanya; prestaja, premishlavanya znamenka ili razmishlavanya znamenka
Kratka osnova horvatsko-slavenskoga pravopisaña (1830.)		+	+	nazivi nisu navedeni
Pravopis jezika ilirskoga (1850.)	+		+	znak dèljenja, znak sastavljenja; znak stanke i znak navoda ili pozvanja
Slovnica Hèrvatska (1861)	+		+	znak skupnoga izgovora
Hrvatski pravopisni rječnik (1886.)	+		+	/
Hrvatsko-njemačko-talijanski rječnik znanstvenoga nazivlja (1873.)	+		+	vezuljka ili sponka; stanka ili pauza
Nauka o pravopisu hrvackoga ili srpskoga (1889.)	+		+	spojnik; bez naziva za $c_3$
Hrvatski pravopis (1892.) i (1893.)	+		+	vezica; crtica
Dra Ivana Broza Hrvatski pravopis (1904.), (1906.), (1911.) i (1915.)	+		+	crtica; stanka
Pravopisni rječnik s pravopisnim pravilima (1921.)	+		+	crtica; stanka
Pravopis hrvatskoga ili srpskoga jezika (1921.), (1923.), (1926.) i (1928.)	+		+	crtica; stanka
Pravopisno uputstvo (1929.)	+		+	crtica (vezica); znak počivka
Pravopis hrvatskoga ili srpskoga jezika (1930.)	+		+	vezica; crtica
Pravopis hrvatskosrpskoga jezika (1934.) i (1937.)	+		+	vezica; crta
Hrvatski pravopis (1941.)	+		+	spojnica; crtica

Pravopis hrvatskoga ili srpskoga jezika (1941.), (1947.) i (1951.)	+		+	crtica; stanka
Koriensko pisanje (1942.)	+		+	/
Korienski pravopis (1944.)	+		+	spojnica; crtica (stanka)
Pravopis hrvatskosrpskoga književnog jezika (1960.)	+		+	crtica; crta
Pravopis hrvatskosrpskoga jezika (školsko izdanje) (1960.)	+		+	crtica; crta
Hrvatski pravopis (1971.)	+		+	crtica; crta
Pravopisni priručnik hrvatskoga ili srpskoga jezika (1986.), (1987.), (1990.)	+	+		crtica; crta
Osnova pravilnoga pisanja (1987.)	+	+		crtica; crta
Hrvatski pravopis (1994.), (1995.), (1996.), (2000.), (2002.), (2003.), (2004.), (2006.),	+	+		spojnica; crtica
Pravopis hrvatskoga jezika (2001.)	+	+		crtica; crta
Pravopisni priručnik (dodatak Velikom rječniku hrvatskoga jezika) (2003.)	+	+		spojnica; crta
Pravopisni priručnik (dodatak Enciklopedijskome rječniku hrvatskoga jezika) (2004.)	+	+		spojnica; crta
Hrvatski školski pravopis (2005.) i (2008.)	+	+		spojnica; crtica
Pravopis, pravogovor, rječnik (2010.)	+	+		spojnica; crtica
Hrvatski pravopis (2007.) i (2008.)	+	+	+	crtica kao znak rastavljanja; crtica ili spojnica; crta
Hrvatski pravopis (2010.) i (2011.)	+	+		spojnica; crtica
Hrvatski pravopis (2013.)	+	+		spojnica; crtica

Slika 6. Pregled oblika i nazivlja središnjih crta u hrvatskim pravopisnim priručnicima

Prvi pravopis koji govori o polusloženicama je Naputcheny (1808.). Kad se definira u pravilu,  $c_1$  je predug u odnosu na primjere korištenja.

Mažuranić (1861) ima 6 stranica o pravopisu (26–32). U pravilima samo spominje  $c_1$ , ali ga ne definira. U tekstu se rabi  $c_1$  i  $c_3$ .

Među svim starijim djelima, Kušar (1889) ima najbogatija pravila pisanja središnjih crta. Prvi uspostavlja pravilo o pisanju  $c_1$  kada se prvi dio ne mijenja, spominje „nedovršeno složene ili primaknute riječi” (*nazovi-čovjek, nadri-pjesnik, pra-svijet*), „dekompoziciju rečenice” (dan-i-noć, dvadeseto-ljetni), od-

nos početka i svršetka (*pruga Knin-Siverić*), prijedlog + r + suglasnik (*za-rzati, po-rđati*), samoglasno *r* koje se piše sa spojnikom da se ne bi čitalo suglasno (*umr-o, dopr-o*), sklanjanje stranih riječi (*Shakespeare-ov*), niječnicu koja daje „individualno značenje” (*ne-dobar*).

Pravopisno uputstvo za sve osnovne, srednje i stručne škole Kraljevine S. H. S iz 1929. rabi naziv *crtica* za  $c_1$  koji se dvaput pojavljuje u latiničnome tekstu; u ćirilichnoj inačici na jednome se mjestu navodi *crtica*, a na drugome *vezica*. Pravila se za  $c_3$  ne obrađuju, a usputno se na jednome mjestu navodi njegov naziv *znak počivka*.

Cipra-Guberina-Krstićev pravopis iz 1941. nije objavljen, a svjetlo je dana ugledao tek tiskom u 1998. Taj je pravopis prvi koji spominje naziv *spojnica* koji je i danas standardan. Bitno je proširio pravila o pisanju središnjih crta, opisao sklanjanje kratica te istaknuo pravilo „spojnica spaja, crtica rastavlja” koje naglašava pisanje crtica s bjelinama.

Korienski pravopis (1944.) autora F. Cipre i A. B. Klaića ima vrlo opširno poglavlje s pravilima pisanja središnjih crta, a važan je po tome što prvi defini- ra slučajeve kada se crtica piše s bjelinama, a kada bez njih.

Pravopis hrvatskosrpskoga književnog jezika (1960.) uveo je praksu koja se nije promijenila sve do današnjih dana da se središnja crta iz riječi piše i na početku prelomljenoga retka. Uspostavljeni nazivi očuvali su se do danas i normiraju se u pojedinim suvremenim hrvatskim pravopisima.

Iako je školsko izdanje Pravopisa hrvatskosrpskoga književnog jezika (1960.) napravljeno prema „velikome novosadskome pravopisu”, on ustaljenim izrazima propisuje drukčije pisanje. To je prvi pravopis koji frazemske riječi piše odvojeno i sa  $c_1$ , a  $c_3$  bez bjelina.

Primjena računala u tiskarsku tehnologiju uzrokovala je skraćivanje duljine  $c_3$ , a prvi pravopis koji je to registrirao je Anić-Silić (1986) nakon kojega postaje standard.<sup>15</sup> Ovaj je pravopis prvi i po tome što je spomenuo središnje crte kao znakove pobrojavanja te unutar broječanoga niza. Pravopis Anić-Silić uvodi i pravopisni znak *crte*, ali ne kao novi pravopisni znak, nego kao slučaj pojavljivanja dviju crta u paru. Anić-Silićev pravopis (2001) prvi uvodi mogućnost pisanja *5-dnevni*.

Iako su Pravopisni priručnik (2003.) i (2004.) dodatci Anićeve rječniku hrvatskoga jezika, odnosno Enciklopedijskome rječniku hrvatskoga jezika, koji

---

<sup>15</sup> Gubitak dulje crte neki tumače trendom opće popularizacijom tiskarske tehnologije u vidu stolnoga izdavaštva (DTP) i snažniji utjecaj nestručnjaka u grafičarskoj i tipografskoj struci („deprofesionalizacija”).

se pak temelje na Anić-Silićevu pravopisu (2001), urednik i izdavač odstupaju od njihovih naziva i proširuju pravila pisanja.

Babić, Ham i Moguš (2005) prvi spominju središnju crtu u kontekstu računalnoga pisanja.<sup>16</sup>

O pravopisu BMM (2008) već je bilo riječi u Portada i Stojanov (2009), a ovdje će se rezimirati. Pisanje je središnjih crta bitno inovativno, ali ujedno i tehnički loše izvedeno pa korisnike dovodi u zabludu o njihovu broju i obliku. Sudeći prema nazivlju (*crtica* i *crtā*), zaključilo bi se da autori predlažu uobičajeno rješenje s dvjema središnjim crtama, međutim to se čini upitnim na više mjesta. Temeljni je problem nastao iz želje da se tradicionalno isti grafem ( $c_1$ ) koji služi za rastavljanje riječi u novi redak i za spajanje riječi počne međusobno razlikovati u duljini, a što je dovelo do produljivanja znaka  $c_2$  (*crtā* u BMM). Drugim riječima, *crtica kao znak rastavljanja* ima jednu grafičku reprezentaciju, *crtica ili spojnica* drugu, a *crtā* treću. Nadalje, u poglavlju gdje se predstavlja znak dulje središnje crte (*crtā*) u istome se odlomku rabe dva bitno različita grafema što bi moglo upućivati i na  $c_4$ . S druge strane, pravopis dodatno zbunjuje korištenjem polubjelina (polurazmaka) u primjerima s polusloženicama u kojima su sastavnice trajno ili privremeno združene čime se isti znak (*crtica*) zahvaljujući polubjelinama i odabirom središnje crte veće duljine vrlo lako može prepoznati kao *crtā*.

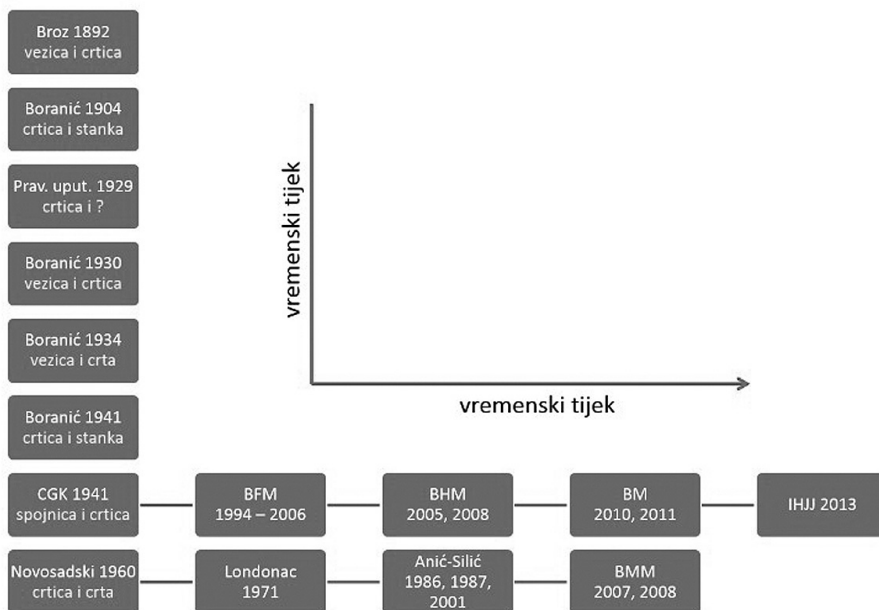
Nije netočno zaključiti da se u BMM (2008) koristi čak pet grafema za (vjerojatno) dvije ili tri vrste središnjih crta. Prvo je *crtica kao znak rastavljanja* ( $c_1$ ), drugo je *crtica ili spojnica* ( $c_2$ ) u polusloženicama (nema bjeline ispred i iza znaka), treće je *crtica ili spojnica* ( $c_3$ ) u složenicama (s polubjelinama ispred i iza znaka), četvrto je *crtā* ( $c_4$ ) i peto je neimenovani znak za najdulji grafem koji predstavlja središnju ravnu vodoravnu crtu ( $c_5$ ).

Svi hrvatski pravopisni priručnici od 1779. do zadnjega IHJJ iz 2013., izuzevši BMM, definiraju i/ili rabe isključivo dvije različite središnje crte. Svi hrvatski pravopisni priručnici od 1779. do 1987., izuzevši Gaja (1830), imaju  $c_1$  i  $c_3$ , a svi izdani poslije 1987. do danas (kao posljedica uvođenja računalne tehnologije u tiskarsku proizvodnju) skratili su  $c_3$  u  $c_2$  pa je prevladavajuće suvremeno stanje  $c_1$  i  $c_2$ .

---

<sup>16</sup> „Nema li računalo dva različita (crtā, op. T. S.) znaka, upotrijebit ćemo jedan, a tada se *crtica* i *spojnica* razlikuju samo po bjelinama.” (§87.)

Iako je pohvalno prvo pojavljivanje tipkopisnoga pravila u hrvatskim pravopisima, navođenje izraza „nemanja znakova u računalu” nije osobito precizno. Bolje je istaknuti izričaj „ako nije moguće računalno ostvariti...”.



Slika 7. Kronološki prikaz razvoja nazivlja središnjih crta u hrvatskim pravopisnim priručnicima

Pravopisi koji u znatnijoj mjeri opisuju uporabu središnjih crta jesu CGK (1941), Cipra-Klaić (1944), Anić-Silić (1986) i BMM (2007).

Što se tiče nazivlja, uočavaju se tri nazivoslovna smjera: prvi je prijestandardni u kojemu nije bilo dosljednosti u nazivlju središnjih crta i u kojem je svaki autor predlagao svoja rješenja (od početaka pa sve do 1941., odnosno 1960.). Drugi je smjer od 1941. do danas koji ima *spojnicu* i *crticu* (od CGK do IHJJ iz 2013.), a treći je smjer od Novosadskoga pravopisa 1960. do pravopisa BMM iz 2008. koji rabe *crtica* i *crta*. Najveći je terminološki nered imao Boranić koji je u svojih 10 pravopisnih izdanja čak 4 puta mijenjao nazivlje – 1904., 1930., 1934. i 1941.

## Jezikopisni opis središnje crte

Hrvatski pravopisni priručnici u opisu središnjih crta konceptualno identično pristupaju tematici, a to je promatranje pravopisnih vodoravnih znakova na razini grafema ili pojavnoga oblika. Pitanje pravopisnih središnjih crta temeljno se svodi na njihov broj (dvije ili tri) i nazivlje (novosadsko ili starije kroati-

stičko nazivlje), dok je uporaba bitno redukcijaska, a opis usmjeren na osnovnu, školsku razinu pismenosti.

Dok je pravopisna metodologija temeljena na odnosu grafema i fonema, računalna jezikopisna metodologija posve drugačije pristupa središnjim crtama i uspostavlja dihotomiju znaka i grafema. Ona ukazuje na drugu perspektivu u kojoj se pisanje središnje crte ne svodi na opreku kraće i dulje, nego, posve suprotno, utvrđuje veliki broj središnjih vodoravnih crtnih znakova (11). Visoki stil računalnojezičnoga izražavanja („tipkopisna pismenost”) izražava se u obožavanju i njegovanju semantičkoga zapisivanja teksta na računalu.

Budući da se crtni znakovi vežu za pismo, a ne za jezik, moguće je zamisliti i međujezičnu standardizaciju pisanja crtnih znakova.

Sva značenjska mjesta koja unikod kodificira postoje i u hrvatskome jeziku, ali se ona razlikovno ne ostvaruju. Jedanaest pojavnih crtnih oblika u hrvatskim se pravopisnim priručnicima prikazuju primarno s dvije središnje crte (više od dvije u slučaju pravopisa BMM). U prikazu jedanaest oblika navedeno je da se sigurno može govoriti o tri, a čak i o četiri skupine središnjih crta koje se obilježavalo oznakama  $c_1$ ,  $c_2$ ,  $c_3$  i  $c_4$ .

Za razliku od primarno kroatističkoga normativno-pravopisnog pristupa, višedisciplinarni deskriptivno-jezikopisni model polazi od znaka i njegova opisa vrijednosti i pripisane funkcije. Jezikopisni opis znakova podrazumijeva usredotočenost na odnos značenja i grafema, a ne samo na formu. Umjesto tradicionalne podjele pravopisnih poglavlja na točku, zarez i zvjezdicu, moglo bi se govoriti o tematskim cjelinama opisa kraja rečenice, pisanju rednih brojeva, nabranju i umetanju, odnosno o izostavljanju grafema ili ispravljanju zatipaka u *chatu*.<sup>17</sup> Tako bi se sada raštrkana pravila u suvremenim pravopisima mogla objediniti u tematsku cjelinu kao što je „pisanje brojeva, nadnevaka i vrijednosti”, „izražavanje omjera i dosega”, a umjesto poglavlja o pisanju skraćenica dati općenita i generativna pravila o akronimizaciji. Na isti bi način, umjesto dijela poglavlja o središnjoj crti, zasebna cjelina o prijelomu riječi opisala dotičnu problematiku i iz računalne, primijenjene korisničke perspektive.

Odvajanjem sadržaja od oblika dobila bi se preciznija argumentacija pa bi se tako govorilo da se „unikodni znak minus-crtice (eng. *hyphen-minus*, U+002D) u hrvatskome ostvaruje grafemom  $c_1$ ”, odnosno da se „svi znakovi skupine  $c_1$  nazivaju spojnicama, zapisuju najkraćim oblikom ravne središnje vodoravne crte i grafemom koji se na hrvatskoj tipkovnici ostvaruje tipkom kraj desnoga

---

<sup>17</sup> Dio pravopisnih priručnika ima tako građena poglavlja pa nailazimo na poglavlja „navođenje u tekstu”, „pisanje oblika i posvojnih pridjeva od općih riječi i imena” itd., ali bi taj funkcionalno-sadržajni pristup trebalo dosljednije provesti.

*shifta*”. Praktično je za školsku uporabu pojednostaviti pisanje središnjih crta svođenjem na opreku kraće i dulje središnje crte, ali to ujedno ne mora biti pravilo i akademskoga pisanja koje će htjeti rabiti  $c_3$  i  $c_4$  za pauze, nabranjanja, ponavljanje nositelja bibliografskih jedinica i drugo.

Ne tvrdi se da je suvremeno pisanje na hrvatskome nužno potrebno posložiti na praktičnoj ili deskriptivnoj razini, kao što se ne tvrdi da postojeći pravopisni priručnici ne zadovoljavaju većinu svojih korisnika. Kritički se promišlja da ovoj razini opisa treba dodati još jednu perspektivu – enciklopedijsko-deskriptivnu, višedisciplinarnu i višejezičnu perspektivu, a koja je obuhvaćena jezikopisnom metodologijom. Potreba višedisciplinarnoga pristupa najbolje se uviđa u problematici prijeloma kojoj, osim jezične struke, svoj doprinos daju i tipografi, računalci te grafičari.

Jezikopisni opis uporabe vodoravnih crta motiviran je prihvaćanjem snažno izražene tendencije suvremene jezične tehnologije koja nastoji obogatiti zapis prirodnoga jezika dodatnim semantičkim obilježjima i oznakama koje pridonose stvaranju naprednijega semantičkog *weba*.<sup>18</sup> Taj pristup u konačnici može posvijestiti potrebu obogaćivanja hrvatske tipkovnice novim vrijednostima tipaka kao što su različite vrste središnjih crta, a koje bi se trebale moći ostvariti i bez numeričkoga dijela tipkovnice te na dovoljno jednostavan način (npr. korištenjem pomoćnih tipaka *control* ili *alt gr*).

Da bi se pokazala različitost pravopisnoga i jezikopisnoga pristupa, slojevitost rukopisane, tiskane i računalne prakse u kontekstu odnosa plana sadržaja i izgleda, bilo je dovoljno odabrati jedno pravopisno poglavlje kao što je pisanje središnjih crta. Iz dosad navedenoga uočava se koliko nedostaje opis jezične prakse koji će uvažavati računalno pisanje. Diskrepancija računalnoga pisanja i tradicionalnoga opisa hrvatskoga jezika s vremenom postaje sve veća, a metodološka hermetičnost pravopisnih priručnika čini njihovo korištenje usmjerenim isključivo na školsku dob i bez dovoljnoga utjecaja na suvremeno pisanje i javne medije. Jezik se mijenja kroz svoj medij te je potrebno uvažiti računalno pisanje kroz kratko predstavljena jezikopisna načela. Osim što bi nova metodologija opisa jezika dodatno ukazala na jezičnu slojevitost i bogatstvo, našoj bi kroatistici mogla dati prepoznatljiv međunarodni pečat suvremene deskriptivnosti.

---

<sup>18</sup> Razvoj semantičkoga *weba* važan je dio rada organizacije koja promiče razvoj otvorenih i javno dostupnih internetskih standarda konzorcija W3 (<http://www.w3.org/>), a koji stoji iza današnjega interneta. Cilj obogaćivanja sadržaja (jezičnim i drugim) značenjem je stvaranje univerzalnoga medija za razmjenu podataka i informacija. Tako umrežene baze znanja i ekspertni sustavi bitno bi trebali pomoći razvoju inteligentnijih strojeva.

Za pravopise se često govori da su značajnije tradicionalno orijentirani. Što se događa ako računalno tipkanje bitno naruši tradiciju – treba li to norma prihvatiti ili ne? U nastavku se opisuje jedan jasan primjer. Od 1960. do 2013. svi pravopisi propisuju pravilo pisanja spojnice na početku prelomljenoga retka kada riječ sadržava spojnicu. To je pravilo došlo u koliziju sa suvremenim pisanjem te će se u nastavku navesti osam argumenata u prilog napuštanju te tradicije. Argumenti koji će se koristiti dolaze iz prakse računalnoga pisanja s ciljem dokazivanja da, za razliku od pravopisnih smjernica, jezični opisi moraju biti deskriptivniji i obzirniji na uvažavanje specifičnosti tipkopisa kao i prakse koja je u znatno većoj mjeri okružena računalno stvorenim tekstovima.

## Pravopisna tradicija prema tipkopisnoj praksi pisanja spojnice na početku prelomljena retka

### Tehnički argument nemogućnosti ostvarivanja takva pisanja pri ručnoj hifenaciji

U suvremenim aplikacijama za obradu teksta kao što su Microsoft Office, OpenOffice i LibreOffice nije moguće ručno hifenirati tekst koji bi rezultirao željenim ishodom. Postavi li se pokazivač na početak prelomljenoga retka i upiše li se spojnica, ona će se automatski premjestiti u prethodni redak. Aplikacije za obradu teksta jednostavno ne sadržavaju funkciju kojom bi se to moglo ostvariti, a ni među stotinjak tisuća unikodnih znakova ne postoji znak sa zadanim svojstvom preko kojega bi se ostvarila takva funkcija.

S pet navedenih zaobilaznih načina, doduše, moguće je ostvariti takvo pisanje, ali uz sljedeća ozbiljna ograničenja: 1. rješenje vrijedi isključivo za Word (dakle, ne vrijedi za druge aplikacije u kojima se često piše, kao što su Outlook, Thunderbird, Google Docs, Notepad itd.); 2. tekst koji se tako oblikuje više se ne smije uređivati (ne smiju se upisivati nove riječi ili brisati postojeće jer se promjenom sadržaja teksta mijenja kraj retka) niti se stranica smije preoblikovati (npr. mijenjati margina, veličina fonta itd. jer se promjenom strukture teksta također mijenja kraj retka); 3. tekst se ne smije kopirati jer će lijepljenje u drugi dokument rezultirati pravopisnom greškom.<sup>19</sup> Načini rada koji se predstavljaju odnose se, dakle, isključivo na finalnu fazu dokumentnoga uređivanja teksta koji radimo u Wordu netom prije ispisa ili pohranjivanja u druge formate.

---

<sup>19</sup> To znači da se kopiranjem i lijepljenjem riječi koja se umjetno prelomila dobije neispravna riječ, npr. umjesto jedne dobivaju se dvije središnje crte s bjelinom: *hrvatsko- -bošnjački*.



Prvi je način da se znak *soft hyphen* redovito rabi umjesto *hyphen-minus* na svim mjestima gdje želimo nadzor nad prijelomom. Tada bi se dogodilo ono što se želi: spojnica bi se vidjela kao regularni znak, a kada bi došla na kraj retka, u novome bi se retku automatski pojavila spojnica (Slike 8 i 9 prikazuju takvo korištenje spojnice tijekom pisanja u Wordu 2003 prije prijeloma i nakon nje-ga). To bi bio idealni scenarij da nije riječ o *bugu* u Wordu 2003. Takav se prikaz događa isključivo u Wordu 2003, dok isti scenarij u Wordu 2007, 2010 i 2013 ne ostvaruje identični rezultat.

Ovo je ispitni primjer za Word s korištenjem funkcije *soft hyphena* i rastavljanja „jezično-povijesni“.



Slika 8. Znak *soft hyphen* u Wordu 2003 prije prijeloma

Ovo je ispitni primjer za Word s korištenjem funkcije *soft hyphena* i rastavljanja primjera „jezično-povijesni“.



Slika 9. Znak *soft hyphen* u Wordu 2003 poslije prijeloma

Da je riječ o pogrešci, zna se po tome što je zadano svojstvo znaka *soft hyphen* da bude nevidljiv, tj. da se pojavi samo kada se prijelom ostvari (zato se alternativno i zove *optional hyphen*). Word ima problema s prikazivanjem znaka *soft hyphen* jer takve funkcije ne želi ostvariti s pomoću unikatnih znakova, nego s pomoću vlastitih izvedbenih rješenja, pa je najbolje mjesto na kojemu se može testirati svojstvo *soft hyphena* u prebircima. Tako i Felici u *The Complete Manual of Typography* (2003: 85) konstatira: „A hyphen added by a hyphenation program is called a *soft hyphen*, and it will disappear when the word in which it occurs no longer needs to be broken at line’s end.”

Drugi je način korištenje znaka prijeloma retka (eng. *line break*) koji se ostvaruje kombinacijom tipki *shift + enter*.

Na slici 10 uključen je prikaz skrivenih znakova da bi se pokazalo korištenje funkcije *line break* koji u Wordu ima prikaz s pomoću grafema ¶. Na taj su način fizički razdvojena dva retka nakon čega se može ručno upisati znak *c*, na početku novoga retka.

·Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor  
·incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud  
·exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.¶

Slijedi navođenje teksta koji želimo prelomiti koristeći *shift + enter*: hrvatsko-  
·bošnjački rječnik.¶

Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat  
·nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia  
·deserunt mollit anim id est laborum.¶

Slika 10. Korištenje *line breaka* za dobivanje spojnice na početku prelomljena retka

Treći je način na početku prelomljena retka upisati znak *non-breaking hyphen*.

Četvrti je način na mjesto  $c_1$  upisati *soft hyphen*, a na početak prelomljena retka *non-breaking hyphen*. Taj je način jednako podložan pogreškama u prikazu kao i prvi navedeni način.

Peti način pretpostavlja uključivanje opcije automatske hifenacije pri kojoj ručno upisujemo spojnicu na početku prelomljena retka. Uključivanje automatske hifenacije može biti nepoželjno ako ne želimo hifenaciju provesti na cijelom dokumentu.

Svi ti zaobilazni načini ne zadovoljavaju kriterij računalne primjenjivosti pisanja spojnice na početku prelomljena retka.

Tehnički argument nemogućnosti ostvarivanja takva pisanja pri automatskoj hifenaciji

Automatskom hifenacijom nije moguće ostvariti prijelom riječi koje sadržavaju spojnicu na način da početak prelomljenoga retka započinje spojnicom.

Jezičnotehnoški argument otežanoga strojnoga pretraživanja, čitanja i prevođenja

U suvremeno vrijeme pretraživanje tekstova vrlo je važna računalna operacija, bez obzira na to je li riječ o mrežnome s pomoću prebirknika (eng. *browsers*), lokalnome s pomoću tražilica mapa i dokumenata u operativnome sustavu ili pojedinačnome unutar dokumenta. Primjenom pravila o zapisivanju središnjih crta na početku prelomljenoga retka točnost je rezultata pretraživanja manja jer korisnici u pretraživanju, primjerice tekstnoga dokumenta u PDF-u, ne računa-

ju da računalo pojavnicu *jezično-povijesni* zapravo vidi kao *jezično--povijesni*, *jezično- -povijesni*, *jezično<sup>⊃</sup> povijesni* ili drugačije.

Uvođenjem pravopisnih pravila koja su udaljena od računalne prakse otežava se strojno čitanje (generiranje govora, eng. *text-to-speech*), odnosno tekstovi se čine manje čitljivima kada se oni strojno prevode na druge jezike.

### Argument činjenja masovne pogreške

Normiranjem ovoga pravila propisuje se nešto što velika većina korisnika hrvatskoga jezika na računalu ne može provesti, a iz čega se stvaraju masovne korektorske pogreške. Pristup propisivanju pravila udaljenoga od pravopisne prakse može se usporediti s restriktivnim prometnim pravilom ograničenja brzine vožnje kojega se nitko ne pridržava jer prometnica omogućuje bržu i jednako sigurnu vožnju. Propisivanje brzine može jedino koristiti policiji radi propisivanja prometnih kazni. Taj argument, drugim riječima, odnosi se na to da se pravila moraju prilagoditi masovno prihvaćenoj praksi kao što se prometna ograničenja moraju prilagoditi uvjetima na cesti.

Propisivanje pravila koja se ne provode (osim kod korektora u fazi prijeloma pred tisak), rezultira time da svi korisnici koji pišu hrvatskim jezikom, a osobito đaci i učenici, već u startu čine pravopisne pogreške.

### Argument dosljednosti

Ako se propisuje uporaba pisanja na početku prelomljenoga retka, zašto se onda ona ne propisuje u svim slučajevima? Primjerice, pisanje dužih prelomljenih brojevnih nizova koji sadržavaju spojnicu također bismo na početku prelomljenoga retka trebali započinjati središnjom crtom (bez obzira na to rabimo li *hyphen-minus* ili *figure dash*). Isto se odnosi i na pisanje središnje crte koja rastavlja osnovu i morfološki ili tvorbeni nastavak (npr. 80-godišnjaci).<sup>20</sup>

### Argument poništavanja dvaju identičnih znakova kada se nalaze jedan kraj drugoga

Kad rečenica završava rednim brojem, dvije se točke nalaze u kontaktu (prva točka obilježuje rednost, a druga kraj rečenice), a pravopisi propisuju da se dva grafički identična znaka međusobno poništavaju radi preglednosti. U kontaktu

---

<sup>20</sup> Slovenski pravopis iz 2001. bio je dovoljno oprezan pa je u člancima §608 i §609 proglasio pisanje središnje crte i na početku prelomljena brojevnoga niza.

dviju središnjih crta (spojnica koja rastavlja i spojnica iz riječi) argument poništavanja dvaju grafičkih identičnih znakova ne vrijedi.

Obrana koju pravopisci ističu jest da se time dokida semantičko razlikovanje polusloženice (književno-umjetnički) i složenice (književnoumjetnički). Međutim, pravopisna je praksa svjesno izgubila toliko grafijskih razlikovanja (npr. neprovođenje razlikovanja zapisivanja središnje crte iza negacije<sup>21</sup>, izostavljanje pisanja više interpunkcijskih znakova zaredom, provođenje jednačanja itd.) čime ta argumentacija ne može predstavljati odsudni kriterij.

### Argument nedovoljne međunarodne prihvaćenosti

Iako bi za konačnu potvrdu ovoga argumenta trebalo provesti obuhvatnije međujezično istraživanje, na primjeru nekoliko slavenskih jezika, i to samo onih koji značenjski razlikuju riječi kao što su *književno-umjetnički* i *književnoumjetnički*, rezultat je 3 prema 2 u korist pisanja spojnice: Poljaci i Ukrajinci ne pišu *c*, na početku prelomljena retka, a Česi, Slovaci i Slovenci pišu.<sup>22</sup>

### Argument otežana rada

Osim što je naglašeno da je svima koji tipkaju hrvatskim jezikom teško ostvarivo ili nemoguće provesti ovo pravilo u djelo, korektorima, grafičkim urednicima i drugim stručnjacima koji rade na poslovima prijeloma teksta uklanjanjem pravila olakšao bi se svakodnevni rad.

### Zaključak o pravilu pisanju središnje crte na početku prelomljena retka

U odnosu na tiskani tekst, danas se većinom pisanje ostvaruje računalno, a taj se nerazmjer sve više povećava na štetu tiskanoga medija. U eri tipkanja i čitanja teksta sa zaslona stvarno okruženje hrvatskoga jezika jest u prebirnicima, u mrežnim dokumentima, u aplikacijama za obradu teksta i drugim uređivačima te u plutajućim okvirima internetskih stranica koji se prema pravilima suvremenoga *web*-dizajna (mrežnoga dizajna) povodi načelom tzv. responzivnosti, odnosno načelom u kojemu se tekst mora kvalitetno prikazati bez obzira na korisnikovu razlučivost, medij, orijentaciju ekrana, jezik, veličinu

---

<sup>21</sup> Primjerice, razlika između *nepreporučeni* (onaj koji nije preporučeni, npr. nepreporučeni naziv) i *nepreporučeni* (sve ono što nije preporučeno – dakle, i ono što je *dopušteno*, *zastarjelo*, *žargonsko* itd.)

<sup>22</sup> Podatci o pisanju dobiveni su od lektora s matičnih odsjeka na Filozofskome fakultetu Sveučilišta u Zagrebu.

i oblik fonta, odnosno uređaj s kojega se učitava internetski sadržaj. To znači da se tekst s riječima sa središnjim crtama mora moći **dinamično** prikazati u vrlo raznolikim tehničkim kontekstima (širinama, veličini i obliku fonta, mediju itd.) i pritom biti bez pogrešaka u prikazu teksta i njegova formata, kao što su: *hrvatsko- -bošnjački*, *hrvatsko<sup>~</sup>bošnjački*), odnosno: *hrvatsko-<sup>-</sup>bošnjački*.<sup>23</sup>

Navedenih osam argumenata kritiziraju logiku propisivanja pravila koja se ne mogu provesti ili računalno ostvariti, odnosno koju mogu tehnički izvesti samo posebni zahvati u području grafičkoga oblikovanja teksta. Ističe se da pravopisna pravila za tiskani hrvatski nisu važnija od tipkopisnoga hrvatskoga, te da ne treba normirati pravila pisanja samo za korektore i grafičke obrađivače.

## Zaključak

Temeljni motiv za istraživanje vodoravnih crta kao razgodaka jest što oni, uz navodnike, bitno odstupaju od odnosa jednoga znaka za jedno (unikodno) značenje. Dok navodnici imaju višestruke grafeme (točnije, osam) za jedno značenje, središnje crte imaju najčešće dva grafema (kraća i dulja središnja crta) koja pokrivaju čak 11 (unikodnih i latiničnih) crtnih znakova.

U prvome poglavlju dao se tipološki pregled i opisana je univerzalna semantička mreža središnjih crta kroz jezike i pisma u računalnome kontekstu te su predstavljene 44 ravne središnje vodoravne vrste crta od kojih je izdvojeno njih 11 latiničnih koje predstavljaju ukupnost unikodnih semantičkih znakova središnjih crta.

U drugome poglavlju opisani su oblici i nazivlje središnjih crta u hrvatskim pravopisnim priručnicima. Povijest hrvatskih pravopisa vezano za crtno grafeme može se podijeliti u dva razdoblja i tri skupine, ovisno o grafemskome obliku, odnosno o nazivlju.

Prva podjela prema obliku središnjih crta dogodila se trenutkom ulaska računalne tehnologije u tiskarstvo i dijeli pravopisne priručnike na prvu i drugu generaciju. Međni priručnik dviju pravopisnih generacija je Anić-Silić (1986). Budući da su računala dosad utjecala na pravopisna pravila, očekuje se da će utjecati i dalje te se najavljuje pojavljivanje i treće generacije pravopisnih (tj. jezikopisnih) priručnika koji će se oslanjati na tekovine suvremenoga pisanja i tipkopisa, unikodni standard i prihvaćanje odnosa grafema i značenja.

---

<sup>23</sup> Preveliki razmak među riječima ovdje je namjerno prikazan kao posljedica korištenja funkcije *line-break*.

Druga podjela prema nazivlju središnjih crta dijeli pravopise u tri skupine. Prva je prijestandardna skupina s neuređenim nazivljem gdje je svaki autor uvodio svoja rješenja. Druge su skupine normativni smjerovi s obzirom na slijedenje pravopisa CGK iz 1941. ili Novosadskoga pravopisa iz 1960. Oba nazivoslovna smjera imaju stalnicu korištenja u hrvatskim pravopisnim priručnicima. U odnosu na terminološke parove *spojnica – crtica* i *crtica – crta* dviju tradicija, u radu se pokazala unikodna perspektiva latiničnoga pisma koja sa svojim službenim engleskim nazivima za 11 središnjih crta dosad nije bila prevedena i normirana, a u čijoj nomenklaturi ima mjesta za sva tri korištena naziva (*spojnica*, *crtica*, *crta*) kao i za nove nazive, bilo terminologizirane arhaizme (*povlaka*) ili neologizme (*neprelemajuća spojnica*).

Raščlamba je pokazala da ima opravdanosti da se crtna kodifikacija temelji na trima ili četirima znakovima koji se iz 11 unikodnih latiničnih znakova svoje uspostavljanjem osnovnih skupina središnjih crta, radno nazvanih  $c_1$ ,  $c_2$ ,  $c_3$  i  $c_4$  kao najkraća, srednje duga, duga i jako duga središnja crta. Iz toga se može zaključiti da bi svaka skupina mogla imati svoga grafemskoga predstavnika u vidu spojnice kao  $c_1$ , crtice kao  $c_2$ , crte kao  $c_3$  i duge crte kao  $c_4$ . Iako konačan broj ovisi i o vrsti pravopisnoga/jezikopisnoga priručnika, osnovna bi i školska uporaba mogla biti zadovoljena binarnom podjelom, dok bi viši, akademski stil ipak zahtijevao tri, odnosno četiri središnje crte. Semantički najtočniji pristup bio bi korištenje jedanaest unikodnih grafema za latinične suvremene pismovne vodoravne središnje crtno znakove.

Na temelju jezičnopovijesnoga i računalnojezikoslovnoga poredbenog istraživanja te na temelju supostavljanja unikodne standardizacije crtnih znakova pravopisnoj tradiciji opisa središnjih crta željelo se ukazati na (i) potrebu za širim i interdisciplinarnim pristupom opisa pisane jezične prakse, (ii) nedovoljnost opisa školske razine pravopisnih priručnika za suvremeno pisanje, kao i na (iii) nedostatnost postojeće kroatističke kodifikacije obaju terminoloških smjerala. Da bi se pravopisni priručnici mogli nazvati akademskima, navedeno je da bi u znatnijoj mjeri trebali opisati računalno pisanje i na razini interpunkcije uvesti razlikovanje znaka i grafema. Jedno od takvih mjesta opisa prema kojima bi pravopisi mogli unaprijediti svoju tehnološku suvremenost jest pitanje pisanja spojnice na početku prelomljena retka o čemu se iznijelo osam argumenata za odbacivanje aktualne tradicije.

Korištenje spojnice može olakšati i ubrzati pisanje pa se predviđa da će biti sve više očekivanja od suvremene pisane norme da dopusti i primjere koji su trenutačno izvan nje (npr. *etno-* i *bioproizvodi*). Ne ulazeći dublje u raspravu, temeljno se predlaže da se, ako nema osobita gramatička razloga da se takvo pi-

sanje ne dopusti te ako je pisanje dotičnih središnjih crta samo stvar stila, pravila koja reguliraju prošireno pisanje središnjih crta dopuste i za visoki stil.

U pravopisu Badurina-Marković-Mićanović (2007) pokazali su da se pravopisna standardizacija može povoditi i vrlo inovativnim rješenjima te da je dio javnosti spreman prihvatiti novine, ali da bi za prihvaćanje novih ideja u pravopisnu standardizaciju trebalo pristupiti s drugačijom metodologijom i tehničkim rješenjima.

Budući da pisanje središnjih crta po svome načelu nije pravopisno, nego jezikopisno pitanje (tiče se pisma, a ne jezika), standardizacija pisanja središnjih crta mogla bi biti i međunarodno opisana.

Koliko god elektroničko pisanje može biti redukcijsko prema korištenju središnjih crta, treba razlikovati i rabiti različite vrste središnjih crta jer se dodatnim semantičkim obilježavanjem teksta ukazuje na viši stil pisanja koji, među ostalim, doprinosi čitljivosti (npr. središnje crte u podslovima).

Znanstvena bi javnost trebala procijeniti opravdanost i uvjerljivost provedenoga istraživanja kroz odgovor na sljedeće temeljno pitanje: postoji li potreba u hrvatskoj akademskoj zajednici kao i u hrvatskoj javnosti općenito za jezičnim priručnicima koji će, uz postojeće školske pravopisne priručnike, nadograditi jezični opis akademskim jezikopisnim priručnicima, a koji će u normiranju crtnih znakova, osim jezičnopovijesne tradicije, obuhvatiti i računalnojezikoslovne argumente suvremenoga pisanja?

## Literatura:

- ANIĆ, VLADIMIR; SILIĆ, JOSIP. 1986. *Pravopisni priručnik hrvatskoga ili srpskoga jezika*. Sveučilišna naklada Liber – Školska knjiga. Zagreb.
- ANIĆ, VLADIMIR; SILIĆ, JOSIP. 2001. *Pravopis hrvatskoga jezika*. Novi Liber – Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 1971. Pretisak 1990. *Hrvatski pravopis*. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 1994. *Hrvatski pravopis*. Drugo dotjerano izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 1995. *Hrvatski pravopis*. Treće izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 1996. *Hrvatski pravopis*. Četvrto izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2000. *Hrvatski pravopis*. Peto prerađeno izdanje. Školska knjiga. Zagreb.

- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2002. *Hrvatski pravopis*. Šesto izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2003. *Hrvatski pravopis*. Sedmo izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2004. *Hrvatski pravopis*. Osmo izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2007. *Hrvatski pravopis*. Deveto izdanje. Školska knjiga. Zagreb.
- BFM: BABIĆ, STJEPAN; FINKA, BOŽIDAR; MOGUŠ, MILAN. 2005. *Hrvatski školski pravopis*. Školska knjiga. Zagreb.
- BMM: BADURINA, LADA; MARKOVIĆ, IVAN; MIĆANOVIĆ, KREŠIMIR. 2007. *Hrvatski pravopis*. Matica hrvatska. Zagreb.
- BMM: BADURINA, LADA; MARKOVIĆ, IVAN; MIĆANOVIĆ, KREŠIMIR. 2008. *Hrvatski pravopis*. Drugo izdanje. Matica hrvatska. Zagreb.
- BORANIĆ, DRAGUTIN. 1921. *Pravopis hrvatskoga ili srpskoga jezika*. Zagreb.
- BORANIĆ, DRAGUTIN. 1923. *Pravopis hrvatskoga ili srpskoga jezika*. Drugo nepromijenjeno izdanje. Naklada Ćirilo-Methodske nakladne knjižare d.d. Zagreb.
- BORANIĆ, DRAGUTIN. 1926. *Pravopis hrvatskoga ili srpskoga jezika*. Treće izdanje. Društvo hrvatskih srednjoškolskih profesora. Zagreb.
- BORANIĆ, DRAGUTIN. 1928. *Pravopis hrvatskoga ili srpskoga jezika*. Četvrto izdanje. Društvo hrvatskih srednjoškolskih profesora. Zagreb.
- BORANIĆ, DRAGUTIN. 1930. *Pravopis hrvatskoga ili srpskoga jezika*. Peto izdanje, prerađeno prema propisima Ministarstva prosvjete. Tisak i naklada Jugoslavenske štampe. Zagreb.
- BORANIĆ, DRAGUTIN. 1934. *Pravopis hrvatskosrpskoga jezika*. Šesto izdanje, prerađeno prema propisima Ministarstva prosvjete. Obnova. Zagreb.
- BORANIĆ, DRAGUTIN. 1937. *Pravopis hrvatskosrpskoga jezika*. Sedmo izdanje, prema propisima Ministarstva prosvjete. Zagreb.
- BORANIĆ, DRAGUTIN. 1941. *Pravopis hrvatskoga ili srpskoga jezika*. Osmo izdanje, prema propisima Ministarstva prosvjete. Nakladni zavod Banovine Hrvatske. Zagreb.
- BORANIĆ, DRAGUTIN. 1947. *Pravopis hrvatskoga ili srpskoga jezika*. Deveto, prerađeno izdanje. Nakladni zavod Hrvatske. Zagreb.
- BORANIĆ, DRAGUTIN. 1951. *Pravopis hrvatskoga ili srpskoga jezika*. Deseto izdanje, prema propisima Ministarstva prosvjete. Školska knjiga. Zagreb.
- BROZ, IVAN. 1892. *Hrvatski pravopis*. Troškom i nakladom Kr. hrv.-slav.-dalm. zemaljske vlade. Zagreb.



- BROZ, IVAN. 1893. *Hrvatski pravopis*. Troškom i nakladom Kr. hrv.-slav.-dalm. zemaljske vlade. Zagreb.
- BROZ, IVAN. 1904. *Hrvatski pravopis (priređio D. Boranić)*. Treće prerađeno izdanje. Trošak i nakl. Kr. hrv.-slav.-dalm. zem. vlade. Zagreb.
- BROZ, IVAN. 1906. *Hrvatski pravopis (priređio D. Boranić)*. Četvrto (nepromijenjeno) izdanje. Trošak i nakl. Kr. hrv.-slav.-dalm. zem. vlade. Zagreb.
- BROZ, IVAN. 1911. *Hrvatski pravopis (priređio D. Boranić)*. Peto izdanje. Trošak i naklada Kr. hrv.-slav.-dalm. zemaljske vlade. Zagreb.
- BROZ, IVAN. 1915. *Hrvatski pravopis (priređio D. Boranić)*, šesto izdanje. Naklada Kr.-Slav.-Dalm. Zemaljske vlade. Zagreb.
- CGK: CIPRA, FRANJO; GUBERINA, PETAR; KRSTIĆ, KRUNO. 1941. Pretisak 1998. *Hrvatski pravopis*. ArTresor naklada. Zagreb.
- CIPRA, FRANJO; KLAIĆ, ADOLF BRATOLJUB. 1944. Pretisak 1992. *Hrvatski pravopis*. Hrvatska sveučilišna naklada. Zagreb
- FELICI, JIM. 2003. *The Complete Manual of Typography*. Adobe Press.
- GAJ, LJUDEVIT. 1830. *Kratka osnova horvatsko-slavenskoga pravopisaña*. Budim.
- GAVAZZI, MILOVAN 1921. *Pravopisni rječnik s pravopisnim pravilima*. St. Kugli. Zagreb.
- IHJJ: *Hrvatski pravopis*. 2013. Gl.ur. Jozić, Željko. Autori: Goranka Blagus Bartolec, Lana Hudeček, Željko Jozić, Kristian Lewis, Milica Mihaljević, Ermina Ramadanović, Matea Birtić, Jurica Budja, Barbara Kovačević, Ivana Matas Ivanković, Alen Milković, Irena Miloš, Tomislav Stojanov, Kristina Štrkalj Despot). Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- JAZU: *Rječnik hrvatskoga ili srpskoga jezika* (1880. – 1976.). Jugoslavenska akademija znanosti i umjetnosti. Zagreb
- KLAIĆ, ADOLF. 1942. *Koriensko pisanje* Hrvatski državni ured za jezik. Zagreb.
- KLAIĆ, ADOLF. 1942. *Koriensko pisanje*. Drugo izdanje. Hrvatski državni ured za jezik. Zagreb *Kratki navuk za pravopiszanye horvatszko za pòtrebnost nàrodnih skol (Einleitung zur croatischen Rechtschreibung zum Gebrauche der Nationalschulen im Königreiche Kroatien)*. 1779. Ofen (Budim). Pretisak 2003. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- KUŠAR, MARČEL. 1889. *Nauka o pravopisu jezika hrvackoga ili srpskoga (fonetičkom i etimologijskom)*. Dubrovnik.
- MESAROŠ, ARPAD FRANJO. 1971. *Grafička enciklopedija*. Tehnička knjiga. Zagreb.
- MIMICA, IVO; VUNIĆ, MARKO. 1973. *Priručnik hrvatskoga ili srpskog jezika za stenografe, daktilografe i ostale birotehničare*. Mala tehnička biblioteka. IV. izdanje. Hrvatsko stenografsko društvo. Zagreb.
- PARTAŠ, JOSIP. 1850. *Pravopis Jezika ilirskoga*. Pretisak 2002. Institut za hrvatski jezik i jezikoslovlje. Zagreb.

- PARČIĆ, DRAGUTIN. 1901. *Rječnik hrvatsko-talijanski*. Tisak i naklada „Narodnoga lista”.
- PORTADA, TOMISLAV; STOJANOV, TOMISLAV. 2009. O vodoravnim crticama u hrvatskome jeziku. *Filologija* 52. 91–120.
- Pravopis hrvatskosrpskoga književnog jezika s pravopisnim rječnikom*. 1960. Izradila pravopisna komisija. Matica hrvatska, Matica srpska. Zagreb – Novi Sad.
- Pravopisno uputstvo za sve osnovne, srednje i stručne škole Kraljevine S. H. S. s kratkim tumačenjem i objašnjenjem*. Prosvetni glasnik, službeni organ Ministarstva prosvete XLV/9. 747–771.
- Slovenski pravopis*. 2001. Skupina autora. Inštitut za slovenski jezik Frana Ramovša ZRC – Slovenska akademija znanosti in umetnosti. Ljubljana.
- STOJANOV, TOMISLAV. 2015. *Jezikopisno planiranje i elektronička pismenost na hrvatskome jeziku*. Doktorski rad (u postupku). Filozofski fakultet Sveučilišta u Zagrebu.
- The Unicode Standard*. <http://unicode.org/>

## Language Historical and Computational Linguistic Aspects of the Descriptions and Norming of Dashes in the Croatian Language

### Abstract

This paper describes one of two punctuation marks (dashes and quotation marks) that deviate significantly from the relationship of one character per (Unicode) semantic value. While quotation marks have multiple graphemes (eight, specifically) for one semantic value, dashes typically have two graphemes (a short and a long dash) that cover as many as 11 (Unicode and Latin) dash characters. While the criteria for line length has typically been highly prominent in orthography manuals, it is only found in the presented categorization on the sixth hierarchical level.

Aside from two new Unicode dash characters (the *two-em dash* and *three-em dash*, Unicode 6.1, January 2012) having been standardized in the meantime, differing methodology and a comparison of the linguistic-historical and computational linguistic aspects have spread awareness of dash characters in the Croatian language as described in Portada-Stojanov (2009). A categorization is presented that is sensitive to the dichotomy of graphic representation and meaning that divides all dash characters into five hierarchical levels. Among the 44 Unicode horizontal and unbroken dash characters, a division into type, time,

functionality, direction, and line height has resulted in 11 contemporary Latin alphabetic horizontal central characters, among which each language written in the Latin alphabet chooses its own. The semantic value and usage of all Unicode dash graphemes has been described.

On the other hand, the paper also described dash characters from the perspective of Croatian historical linguistics and orthography. In comparison to the rich repository of standardized Unicode dash characters, it has been shown that orthographic standards are significantly reductive. Orthographic norming of dash characters is divided into two periods and three groups, depending on their graphemic form (the first and second generation of orthography manuals) and terminology (the pre-standard phase and the two standard norming schools, depending on the acceptance of the terminological pairs “spojnica – crtica” and “crtica – crta”).

The historical linguistic and computational linguistic comparative research and the contrastive analysis of the Unicode standardization of dash characters with traditional orthographic descriptions of dash characters was intended to highlight (i) the need for a broader, interdisciplinary approach to describing written linguistic practice, (ii) the insufficiency of descriptions in primary and secondary school orthography manuals for modern writing, and (iii) the insufficiency of the existing Croatian codification of both terminological schools. In order for orthography manuals to be called scholarly, it is claimed that computer writing should be better described, and that a differentiation between characters and graphemes should be introduced on the level of punctuation. One of the areas in which orthography manuals could bring themselves technologically up to date is the issue of the writing of compound words at the beginning of a broken line, and the paper provides eight reasons to abandon the current tradition.

Analysis has shown that it would be justified to base dash codification on three or four characters, which reduces the 11 Latin Unicode characters to basic groups of dashes – the short, medium, long, and very long dashes, referred to as  $c_1$ ,  $c_2$ ,  $c_3$  and  $c_4$ .

Ključne riječi: hrvatski jezik, pravopis, jezikopis, unikod, crta, spojnica, crtica

Key words: Croatian language, orthography, linguography, Unicode, dash, hyphen

