

# ZRCola: vnašalni sistem za jezikoslovno rabo v programu word

Peter Weiss

*IZVLEČEK: Vnašalni sistem ZRCola je bil na Znanstvenoraziskovalnem centru SAZU v Ljubljani razvit za jezikoslovne, predvsem dialektološke potrebe in deluje v Microsoftovem programu word v operacijskem sistemu windows. Temelji na standardu unikod, okoli 1400 dodatnih znakov – med njimi so tudi slovenski zgodovinski znaki iz prve polovice 19. stoletja – pa je v pisavi 00 ZRCola razporejenih na področju zasebne rabe tega standarda. Zapleteni znaki se vnašajo tako, da se osnovnemu znaku v obveznem zaporedju dodajo zaslonski diakritični znaki, nazadnje pa program nize s temi znaki na uporabnikovo zahtevo sestavi v samostojne znake oz. pismenke. Ta način omogoča s prirejeno tipkovnico, ki je del vnašalnega sistema, preprosto vnašanje latiničnih in slovanskih ciriličnih besedil in iskanje po njih, pa tudi lažjo izmenjavo datotek in zanesljivo organiziranje podatkovnih zbirk.*

## *The ZRCola Input System for Linguistic Usage in Program Word*

*ABSTRACT: The ZRCola input system was developed at the Scientific Research Centre SASA in Ljubljana and is designed to meet the needs of linguists, especially of dialectologists. It is used with Microsoft Word under the Windows operating system. Based on the Unicode Standard, the input system includes approximately 1,400 additional characters – among those are historical Slovene characters from the first half of the 19<sup>th</sup> century – in the 00 ZRCola font within the private use area of this standard. The input of complex characters is done in canonical stack: first the user selects the base character, then adds screen diacritical marks, and finally the application combines these strings of various marks into individual characters or glyphs on user's command. With the use of a customized keyboard, which is part of the input system, this method enables simple input of texts in Latin and Slavic Cyrillic alphabets as well as it simplifies computerized search within these texts. In this way the data exchange is also much easier and the organization of data files is more reliable.*

1 Vnašalni sistem ZRCola je skupek pisave, temelječe na standardu unikod in razširjene z latiničnimi in ciriličnimi črkovnimi znaki, makrov in dokumentacije,

kar vse v Microsoftovem programu word v operacijskem sistemu windows omogoča preprosto vnašanje, urejanje in tiskanje različnih posebnih znakov, ki se uporabljajo v jezikoslovju, predvsem v dialektologiji, pa tudi pisanje v različnih jezikih, ki uporabljajo predvsem latinico in cirilico (vnašalni sistem ne zajema vseh znakov v neslovanskih jezikih, ki se pišejo v cirilici) in v njih ločevalne (diakritične) znake (nadpisane, podpisane in prepisane). Ime ZRCola je bilo izbrano zato, ker je bil sistem razvit v dialektološki sekciji Inštituta za slovenski jezik Frana Ramovša na Znanstvenoraziskovalnem centru (ZRC) Slovenske akademije znanosti in umetnosti v Ljubljani predvsem za potrebe Slovanskega lingvističnega atlasa (z uveljavljeno rusko kratico *OLA* za *Obščeslavjanskij lingvističeskij atlas*) ob upoštevanju dosežkov slovenske dialektologije in v njej pripravljajočih se del, kot je npr. Slovenski lingvistični atlas.

Sistem združuje prednosti vnašanja znakov z nekdanjimi velikimi uporabljanimi in zelo uporabnim programom *STeve* Primoža Jakopina (ta program je deloval na računalnikih atari – nasledil ga je program *eva*, ki deluje na osebni računalnikih v okolju windows ali DOS, istega avtorja) in zanemarja varljivo sestavljanje znakov v programu word, ki za resno delo ni primerno. Tako recimo lahko v besedilu v programu word znak *í* vnesemo sorazmerno elegantno kot kombinacijo naglasnega znaka (dosegljiv je na *AltGr+9*) in črke *i*, vendar pa vnaprej ne moremo z gotovostjo vedeti, ali ta sistem deluje tudi za *í*, *ś*, *Í*, *Ń* in *â* (deluje) in za *û* (deluje, vendar v primerjavi z *â* s težjo kombinacijo tipk), za spodnjelužiški *ś*, za malteški *ġ* in za *ќ* za makedonsko prečrkovanje (ti znaki so s standardiziranega unikatnega področja) ter za *ě* in *š*, ki sta z unikatnega področja zasebne rabe (za vse te bližnjični sistem v wordu ne deluje). Precej drugačno je v wordu npr. vnašanje znakov z dvojnimi krativcem, čeprav jih najdemo sestavljene v standardiziranem delu unikatnega, povsem drugje – v pisavi *00 ZRCola* na področju zasebne rabe – pa bi našli npr. polglasnik z dvojnimi krativcem, ki ga ni v standardiziranem delu unikatnega. Za vsakega od teh znakov je potrebna posebna bližnjica – v *ZRColi* je posebno le mesto polglasnika (*Alt+D*) in dvojnega krativca (*Alt+I*), osnovni samoglasniški znaki pa so seveda že na tipkovnici. – Nekatere znake z ločevalnimi znamenji lahko v wordu fingiramo tako, da najprej vnesemo osnovni znak, zatem pa še tistega, ki naj stoji nad njim ali pod njim in ga program zamakne na levo, vendar pa je potem le izjemoma prikazan ustrezn rezultat, npr. pri *đ*; sploh podpisana pika pod osnovnim znakom le redko stoji na sredini, zato je npr. v različici 2.97 pisave *times new roman* na mestih *F01E* do *F02C* cel niz, namreč 15 podpisanih pik, izmed katerih lahko ustrezno izberemo glede na širino osnovnega znaka. Takšno besedilo je grafično morda celo neoporečno, vendar ga ne moremo obvladovati, saj je po tako različnih podatkih komajda mogoče iskati. Druga nerodnost je ta, da je besedilo s fingiranimi znaki prikazano prav edino v programu word in se skvari (»podre«), ko ga prenesemo v kak drug urejevalnik besedil ali nekompatibilen namizno-založniški program, kot je npr. *pagemaker* podjetja Adobe.

2 Standard unikat (angl. *Unicode Standard*; gl. The Unicode Consortium 2003; [www.unicode.org](http://www.unicode.org)) je osnova za kodiranje znakov in na osnovni ravni vsebuje nad 65.500 (256 × 256) mest. Z znaki, od katerih se nekateri na pogled niti ne

razlikujejo, imajo pa vsak svojo kodo v različnih skupinah pisav, kot npr. latinični veliki A s šestnajstiško kodo 41, cirilični veliki A s kodo 410 in grška velika alfa (A) s kodo 391, in s katerimi se da pisati v mnogih jezikih in pisavah (npr. tudi Braillovo pisavo), zadovoljimo veliko jezikoslovnih potreb (npr. tudi zapis mednarodne fonetične abecede), mnogi posebni znaki pa še vedno ostanejo zunaj standardiziranega kodiranega nabora znakov. Zato se je za rabo v krogu sodelavcev na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU v Ljubljani in Slovenskega lingvističnega atlasa izoblikovala zamisel o oblikovanju ločenega standarda (»podstandarda«) na področju zasebne rabe. Ni namreč mogoče računati na to, da bodo znaki, ki jih uporabljamo, v doglednem času postali sestavni del unikodnega standarda, če sploh kdaj bodo: potrebujemo jih zdaj, in ker jih moramo uporabljati, z uvrstitvijo na unikodno področje zasebne rabe rešimo problem vsaj začasno. Vnašalni sistem s pisavo vred je zastoj in odprt za nove predloge, prilagoditve (za tipkovnico in za del makrov) bodo narejene za predvidoma vse slovanske jezike in vsak novi uporabnik bo dobrodošel, saj bo skupina uporabnikov tako večja, sodelovanje in posredovanje rokopisov lažje in sistem večkrat preizkušen.

V sistemu ZRCola trenutno deluje vnašanje skoraj vseh sestavljivih unikodnih in vseh posebnih latiničnih in ciriličnih znakov in nizov, ki so v pisavi 00 ZRCola, v njem so dosegljivi mnogi simboli, predvidena je razširitev na vse slovanske zgodovinske pisave, npr. dajnica in metelčica (kar je kot želja izraženo v Peterlin – Košir – Erjavec 1998) ter pisava prekmurskih piscev so že upošteevane, Popovičev poskus še bo, zajeto je tudi zapisovanje metričnih shem ipd. – vsega tega doslej v enem naboru še ni bilo. V pripravi je dokumentacija za celoten sistem, ki bo tako kot vnašalni sistem s pisavo vred dosegljiva na internetu (<http://ZRCola.zrc-sazu.si>).

Vnašalni sistem ZRCola omogoča:

- vnašanje besedil z vsemi latiničnimi in ciriličnimi, pa tudi z grškimi, hebrejskimi in arabskimi pismenkami ter veliko simboli in ločili, upoštevanimi v unikodnem standardiziranem naboru znakov;
- z dodatnimi latiničnimi in ciriličnimi pismenkami vnašanje jezikoslovnih (etimoloških, slovenskih zgodovinskih ...) in predvsem narečnih besedil;
- preprosto vnašanje posameznih sestavin s kombinacijami tipk na tipkovnici (najpogostejše bližnjice so že izdelane in dokumentirane);
- samodejno sestavitve v znake za natis, npr.  $\text{á} \rightarrow \text{á}$ ;
- možno (prav tako samodejno) razstavitev sestavljenih znakov;
- preprosto iskanje po razstavljenih znakih v besedilih;
- veliko izbiro črk, števk, ločil in simbolov: ob unikodnem standardiziranem naboru je na področju zasebne rabe upoštevanih nad 1400 dodatnih znakov, njihov obseg pa se še širi;
- nove sestavljene znake iz latinice in cirilice lahko predlaga vsak uporabnik.

3 Pisava 00 ZRCola (dve ničli na začetku imena zagotavljata, da ime stoji vedno na vrhu seznama pisav) temelji na pisavi times new roman, razširjeni z znaki v pisavah brane, ki jih je za potrebe Založbe ZRC in inštituta za slovenski jezik v letih 1998–2002 izdelal Brane Vidmar, in z latiničnimi in ciriličnimi pismenkami,

ki jih je bilo mogoče najti v pregledanih jezikoslovnih delih, ki obravnavajo slovanske jezike. Znaki v pisavi 00 ZRCola so razporejeni v unikodnem naboru, standardu, ki sicer določa kode pismenk za veliko večino pisav: na unikodnem področju so upošteevane vse standardizirane latinične in cirilične pismenke, primerno zastopane so tudi grščina, hebrejščina in arabščina, kar nekaj pa je v tej pisavi tudi simbolov, ločil in drugega. Vnašalni sistem ZRCola izrablja področje zasebne rabe s 6400 prostimi mesti in jih namenja znakom za jezikoslovno rabo, ki jih na standardiziranem unikodnem področju ni. Nabor znakov na področju zasebne rabe se bo dopolnjeval, čeprav seveda nikoli ne bo popoln, pomembno je, da se kode posameznih znakov ne bodo spreminjale, lahko pa se zgodi, da bo kateri od znakov kdaj postal del unikodnega standarda. Podstandard na področju zasebne rabe je smiseln seveda samo, če ostane centraliziran, zato morebitne ločene izvedbe (z znaki na mestih, ki jih »centrala« na ZRC SAZU v Ljubljani ne bi odobrila) niso smiselne, ker bodo zavajale in povzročale zmedo.

Pisavo 00 ZRCola je mogoče uporabljati v različnih računalniških programih, celoten vnašalni sistem pa za zdaj le v programu word v operacijskem sistemu windows, čeprav je prilagoditev za druge programe, npr. za openoffice, možna in seveda zaželeno.

4 Znaki z ločevalnimi znamenji v sistemu ZRCola se vnašajo nesestavljeni oz. razstavljeni (npr.  $i\grave{a}$  za  $i$ ,  $i\grave{u}$  za  $i$  in  $i\grave{u}$  za  $\grave{i}$  in  $i\grave{u}$  za  $\grave{i}$ ,  $u\grave{u}$  za  $u$ ), računalnik pa jih na uporabnikov ukaz npr. za natis korektur in za oddajo v nadaljnji postopek z vgrajenim makrom sestavi. Sistem ZRCola omogoča tudi obratno pot – razstavljanje znakov, vnesenih v standardiziranem unikodnem naboru in tistih na področju zasebne rabe, ki so bili vneseni v sistemu ZRCola, kar pride prav tudi za znake, ki bi jih po nesreči sestavili ali ki so bili vneseni mešano, torej kot nesestavljeni in sestavljeni. (Program za razstavljanje seveda ni nič drugega kot v drugo smer obrnjeni program za sestavljanje znakov.)

Za osnovnim znakom se vnesejo sestavljivi znaki v obveznem (kanoničnem) zaporedju, tako da se najprej napišejo zgornji, potem spodnji in nazadnje prepisani in pripisani:

- (1) nadpisane (zgornje) kopičimo navzgor (kot si dejansko sledijo), npr.  $a\bar{u}\grave{u}$  →  $\grave{a}$ ,
  - (2) podpisane (spodnje) kopičimo navzdol (kot si dejansko sledijo), npr.  $e\grave{u}\grave{u}$  →  $\grave{e}$ ,
  - (3) nazadnje pa zapišemo pripisane oz. prepisane, npr.  $Z\grave{u}$  →  $\grave{Z}$  ali  $l\grave{u}$  →  $\grave{l}$ .
- V tem smislu vnesemo npr.  $o\acute{u}\grave{u}\grave{u}$  →  $\acute{o}$  ali  $l\acute{u}\grave{u}\grave{u}$  →  $\acute{l}$ .

V vsakem črnem pravokotniku, ki nadomešča osnovni znak, mali trikotnik ob zgornji, spodnji ali desni stranici kaže, v katero od teh treh vrst spada kak znak. Zaporedje *zgoraj – spodaj – čez/poleg* je obvezno zaradi iskanja in zato, da so zamenjalne preglednice obvladljive. Dejansko sestavljajni program naredi znak  $\grave{\text{f}}$  iz niza  $l\acute{u}\grave{u}\grave{u}$  (ta je zaradi zaporedja sestavljivih sestavin *zgoraj*, potem *spodaj* in nazadnje *čez* oz. *poleg* obvezen), čeprav bi ga lahko tudi iz  $l\acute{u}\grave{u}\grave{u}$ ,  $l\acute{u}\grave{u}\grave{u}$ ,  $l\acute{u}\grave{u}\grave{u}$ ,  $l\acute{u}\grave{u}\grave{u}$  ali

lúŕŕ, kar vse pa je za resno delo in iskanje preveč razpuščeno in neobvladljivo, razen če bi ustrezní program omogočal samodejno popravo napačnega zaporedja vnesenih znakov, kar pa za zdaj ni predvideno.

Znake, ki so v ustreznem jeziku dosegljivi že na tipkovnici, lahko vnašamo sestavljene, na slovenski (jugoslovanski latinični) tipkovnici npr. č (ki ga bodo tisti, ki ga nimajo na tipkovnici, vnašali kot cč). Vendar pa je pri pisanju sestavljivih nizov treba upoštevati obvezno zaporedje, saj – kot rečeno – npr. znaka ĵ ne moremo dobiti s sestavitvijo znakov tako, da bi bil osnovni znak ĵ. (Nasploh lahko neposredno vnašamo sestavljene znake, kar je smiselno, če jih je malo, le da bližnjične tipke zanje v sistemu ZRCola niso predvidene.)

Zaradi razumljive praktičnosti je kot osnovni znak pred nadpisanimi znaki vedno treba vnesti i (ne ı) in j (ne ĵ), ne glede na to, kaj stoji nad njima; edinole ĵ je treba vnesti kot kombinacijo ĵ (in turški İ kot Iı); obrnjeni i (ı) je seveda poseben znak. (Mesta in kode vseh teh znakov so razvidni iz dodane dokumentacije in se na unikodnem področju ujemajo z znaki v drugih pisavah, ki temeljijo na unikodu.)

Ligature se vnašajo z znakom ſ med sestavinama, npr. OſE za Œ, in če je nad ligaturo še kaj dopisano, je nesestavljenemu nizu preprosto dodano: oſeſſ → čé. Edino v skupini samoglasnikov s podpisanim znakom ſ, ſ ali ſ in temu vodoravno podpisanim na desni strani stoječim znakom ſ se podpisani znaki vnašajo v tem zaporedju (»repek je na koncu«), torej se znaki pod repkom ne nalagajo: eſſſ → ê.

Posebno skupino tvorijo nizi dveh ali treh osnovnih znakov z morebitnimi ločevalnimi znaki, ki jih povezujejo zgornji ali spodnji lok, črta ali vijuga, t. i. povezaji. Če je v takem nizu kak del nadpisan, ga pri vnosu postavimo med dve polovični puščici, npr. eſſ[a] → e<sup>a</sup>. V takih nizih dva črna pravokotnika (npr. ■■) – ali tri (npr. ■■■), če naj bodo povezani trije osnovni znaki – z lokom, črto ali vijugo zgoraj ali spodaj postavimo takoj za prvi, osnovni znak, ki mu lahko sledi ločevalni znak in tudi dvopičje, ki pomeni dolžino glasu, npr. oſ:ſſuſ → oſſu. V nizih, ki – tako kot nizi v ZRColi nasploh – niso poljubni, vendar pa so razširljivi, je dolžino glasu mogoče označevati ali z dvopičjem, ki ima obliko dveh majhnih trikotnikov (torej :), kar ima v jezikoslovju prednost in je priporočeno, ali pa z navadnim dvopičjem (:).

Iskanje tako vnesenih nizov sicer ni čisto preprosto (v iskalna okenca znakov za zdaj ne moremo vtiskavati neposredno, ampak jih moramo vanj skopirati iz osnovnega besedila, v programu word xp pa jih lahko vnesemo z ustrežno kodo, sestavljeno iz štirih črk oz. števk), vendar pa je mogoče v enem koraku vedno poiskati vse bodoče sestavljene pismenke z dvojnim krativcem ali pa s cirkumfleksom in s piko spodaj: iščemo pač samo ſ ali ſſ, medtem ko je to pri sestavljenih znakih veliko bolj zapleteno, kak znak (npr. polglasnik z dvojnim krativcem) pa bi se lahko celo izmuznil, če ga ne bi imeli v zavesti.

Sestavljive znake je mogoče vnašati razstavljene tudi v drugih pisavah, kot sta npr. times new roman ali tahoma, in v njih bo delovalo tudi sestavljanje oz. razstavljanje, vendar pa je prikaz sestavljenih znakov omejen na standardizirano področje unikoda in odvisen od obsega nabora na tem področju v določeni pisavi. Nasploh torej velja: kar je na standardiziranem unikodnem področju, bo v drugi pisavi prikazano pravilno, če ta vsebuje tisti znak, sicer pa se utegne zgoditi, da se

bo namesto kake sestavljene pismenke tudi s standardiziranega področja pisave 00 ZRCola v drugi pisavi pojavil kvadratega, ki naznanja, da v drugi pisavi ustreznega znaka manjka; to se bo zagotovo zgodilo z znaki s področja zasebne rabe v pisavi 00 ZRCola v drugih pisavah.

Trenutno pisava 00 ZRCola deluje le v navadnem slogu, torej navadno pokončno, saj ležeča, polkrepka in morda ležeča polkrepka različica še čakajo na izdelavo, čeprav seveda lahko vnašamo in tiskamo elektronsko ležeče, polkrepke in ležeče polkrepke dele besedil. Celotna pisava bo izdelana na novo, saj so pismenke trenutno preveč raznorodne. Izdelati bo treba tudi navodila in priporočila za vnašanje, recimo kateri znak ustreza malemu polglasniku, Ɔ ali Ɔ̄, kako se vnašajo narekovaji v besedilih v različnih jezikih, kako stopinje Celzija, rimske številke, srbsko-hrvaške ligature lj, Lj, LJ, dž ... itd. Poimenovanje posameznih znakov v slovenščini bi omogočalo lažjo komunikacijo, zato bi bilo primerno, da bi dobili slovenski prevod standarda unikod. Tako recimo na zelo različne načine opisujemo imena pismenk ŋ, m̄, j̄ in ŋ̄ (zadnje tri so iz mednarodne fonetične abecede).

5 Področje zasebne rabe v vnašalnem sistemu ZRCola in v pisavi 00 ZRCola je sestavljeno takole (navedene šestnajstistiške kode označujejo začetke nizov, ki nikjer niso izpolnjeni do konca):

- E000 znaki, ki se pišejo nad osnovnimi, pod njimi, čeznje ali ob njih ipd.
- E100 slovenski zgodovinski znaki, posebni osnovni (samostojni) dialektološki iz latinice in cirilice
- E200 metrični znaki, simboli, ločila
- E2E0 znaki za vrste tiska iz programa eva
- E300 sestavljeni latinični z osnovnim znakom a ipd.
- E3E0 sestavljeni latinični z osnovnim znakom A
- E400 sestavljeni latinični z osnovnim znakom e ipd.
- E4E0 sestavljeni latinični z osnovnim znakom E
- E500 sestavljeni latinični z osnovnim znakom i ipd.
- E3E0 sestavljeni latinični z osnovnim znakom I
- E600 sestavljeni latinični z osnovnim znakom o ipd.
- E3E0 sestavljeni latinični z osnovnim znakom O
- E700 sestavljeni latinični z osnovnim znakom u ipd.
- E3E0 sestavljeni latinični z osnovnim znakom U
- E800 sestavljeni latinični z osnovnim znakom y
- E840 sestavljeni drugi latinični oz. cirilični samoglasniški
- E900 sestavljeni latinični z osnovnim znakom b–j ipd.
- EA00 sestavljeni latinični z osnovnim znakom k–q ipd.
- EB00 sestavljeni latinični z osnovnim znakom r–z ipd.
- EC00 povezajni nizi na latinični a (tj., ki se začenjajo z a)
- EC80 povezajni nizi na latinični e
- ED00 povezajni nizi na latinični i
- ED80 povezajni nizi na latinični o
- EE00 povezajni nizi na latinični u

- EE80 povezajni nizi na druge latinične oz. cirilične samoglasnike
- EEC0 povezajni nizi na latinične soglasnike
- EF00 rezervirano za dodatke
- F0FF rezervirano
- F100 sestavljeni cirilični z osnovnim znakom a
- F140 sestavljeni cirilični z osnovnim znakom A
- F150 sestavljeni cirilični z osnovnim znakom e
- F1A0 sestavljeni cirilični z osnovnim znakom E
- F1B0 sestavljeni cirilični z osnovnim znakom i ipd.
- F1F0 sestavljeni cirilični z osnovnim znakom I
- F200 sestavljeni cirilični z osnovnim znakom o
- F240 sestavljeni cirilični z osnovnim znakom O
- F250 sestavljeni cirilični z osnovnim znakom u ipd.
- F290 sestavljeni cirilični z osnovnim znakom U ipd.
- F2A0 sestavljeni cirilični z osnovnim znakom v–z
- F300 sestavljeni cirilični z osnovnim znakom k–š
- F400 povezajni nizi na cirilične samoglasnike
- F460 povezajni nizi na cirilične soglasnike
- F480 rezervirano (do F8FF)

6 Vnašalni sistem se bo razvijal. Prve predloge za izboljšanje preizkusnih različic so doslej prispevale moje sodelavke iz dialektološke sekcije Inštituta za slovenski jezik Frana Ramovša ZRC SAZU v Ljubljani Karmen Kenda-Jež (že pred časom pripravila nadvse uporaben pregled znakov v pisavah brane, ki je bil upoštevan pri razvrščanju znakov v pisavi 00 ZRCola), Jožica Škofic in Vera Smole, ki so verjele v nastajajoči sistem, se mi zdi, marsikdaj tudi samo na mojo besedo in celo takrat, ko sem sam obupal. Samonamestitveni računalniški program, ki omogoča zelo preprosto namestitvev vnašalnega sistema, in dokumentacijo zanj je pripravil Robert Valentak iz Računalniškega centra ZRC SAZU.

7 Vnašanje jezikoslovnih in sploh narečnih besedil je bilo zaradi mnogih posebnih znakov doslej tako naporno in zapleteno, da se je, mislim, splačalo potruditi in žrtvovati čas, ki bo s tem vnašalcem nekoliko prihranjen. Vesel bom, če bo kak jezikoslovec opazil, da mu je vnašalni sistem ZRCola olajšal poklicno življenje, in če bo vnašalni sistem ZRCola poživil sodelovanje v stroki. To bo tudi spodbuda za njegovo načrtovano dopolnjevanje.

### Navedenke

- The Unicode Consortium 2003 = The Unicode Consortium, *The Unicode Standard, Version 4.0*, Boston itd., Addison-Wesley, 2003.
- Peterlin – Košir – Erjavec 1998 = Primož Peterlin – Aleš Košir – Tomaž Erjavec, *Digitalni zapis slovenskih znakov*, v: Tomaž Erjavec – Jerneja Gros (ur.), *Jezikovne tehnologije za slovenski jezik – Language Technologies for the*

*Slovene Language, Zbornik konference – Proceedings of the Conference*,  
Ljubljana, Institut Jožef Stefan, 1998, 128–132.