

# Slavona românească în Unicode: probleme și soluții<sup>†</sup>

Ion-Mihai Felea\*

*Facultatea de Litere, Universitatea „Alexandru Ioan Cuza”, Bd. Carol I 11, 700506 Iași, România*

## Despre articol

### Istoric:

Primit 25 septembrie 2021

Acceptat 13 octombrie 2021

Publicat 12 decembrie 2021

### Cuvinte-cheie:

editarea textelor

literatură veche

## Rezumat

Editorii de texte slavone și slavo-române au la dispoziție numeroase instrumente (fonturi, aranjamente de tastaturi fizice și virtuale, procesoare de text, sisteme de operare) pentru transcrierea și digitalizarea acestora într-o manieră uniformă. Uniformitatea transcrierilor are la bază standardizarea Unicode. În articol ne propunem să explicăm locul slavonei în Unicode și să descriem succint instrumentele cele mai accesibile. În acest scop vom descrie instrumentele de lucru din perspectivă istorică și funcțională, apoi vom furniza exemple în care acele instrumente pot fi sau au fost deja utilizate pentru a obține o transcriere cât mai fidelă. Utilizatorul poate alege dintre metodele și instrumentele existente în funcție de scopuri, nevoi și mijloacele de care dispune. O înțelegere mai bună a datelor tehnice poate scurta timpul de lucru, poate îmbunătăți transcrierea, poate micșora timpii de învățare și, per ansamblu, ușurează munca editorului.

## 1. Slavona în Unicode. Scurt istoric

Standardizarea alfabetelor de telecomunicații își are originea în telegraful inventat de Samuel Morse. Nevoia de uniformitate a crescut continuu, astfel încât în anii 1960 American Standards Association a creat un sistem numit ASCII (American Standard Code for Information Exchange), în care caracterele uzuale ale limbii engleze au primit câte un număr de la 1 la 127 în sistem binar. În SUA standardul ASCII a fost implementat imediat pe calculatoare.

Sistemul și-a arătat limitele odată cu apariția internetului. Pe de o parte, codificarea ASCII era aproape universal utilizată în SUA și în unele state europene. Pe de alta, unele țări, precum cele scandinave, foloseau variante ale codului ASCII care se deosebeau de cel american, iar altele foloseau alfabet distincte de cel latin. Japonia, de pildă, folosea chiar patru sisteme de codificare complet diferite atât de ASCII, cât și între ele. Era nevoie de o standardizare la nivel global.

Această sarcină a fost încredințată în 1990 consorțiului Unicode, o organizație non-profit înființată pentru a crea, dezvolta și întreține un standard universal al caracterelor. Principiul standardului Unicode era același. Caracterele ASCII și-au păstrat numărul de ordine, iar celorlalte caractere le-a fost atribuit treptat câte un număr unic și invariabil de la 256 la cca. 150 000 (câte sînt astăzi). Armonizarea ASCII cu un sistem unic de localizare a caracterelor viitoare a reprezentat o provocare intelectuală care a beneficiat de o rezolvare elegantă explicată într-un documentar *Computerphile*<sup>1</sup>.

Consortiul s-a ridicat la înălțimea așteptărilor. Astăzi UTF8 – Unicode Transformation Format 8-biți, unul dintre sistemele de codificare dezvoltate de Unicode, este folosit de 97% dintre paginile de internet.

<sup>†</sup> Articol finanțat printr-un grant al Ministerului Educației și Cercetării din România, CNCS – UEFISDCI, proiect nr. PN III-P1.1-TE-2019-0517, în cadrul programului PNCD III și ținut sub formă de comunicare la workshopul „Medieval and Early Modern Lexicography in the Digital Age” din cadrul simpozionului „Explorări în tradiția biblică românească și europeană”, Iași, 3–5 iunie 2021.

\* Adresă de corespondență: [imfelea@gmail.com](mailto:imfelea@gmail.com).

<sup>1</sup> [youtube.com](https://www.youtube.com/watch?v=...).

## 2. Standardizarea caracterelor chirilice și slavone

Pentru alfabetul chirilic standardizarea a avut loc în etape începând cu 1990. În primă instanță au fost codificate alfabetele chirilice moderne într-un bloc compact, de bază, repartizat numeric între 1024 la 1279. Caracterul A chirilic a primit numărul 1040, Б chirilic – 1041 etc. În 2002 au fost codificate caractere specifice din alfabetele chirilice folosite de mordvină, azeră, ciuvașă ș.a. Acest prim bloc suplimentar, localizat numeric între 1280 și 1327, a fost numit Cyrillic Supplement. Următoarele blocuri, numite Cyrillic Extended A, B și C, Phonetic Extensions, Combining Half Marks, Glagolitic și Glagolitic Supplement, au venit să servească în special cercetătorilor și au codificat caractere precum literele suprascrise, iusurile iotate, Înea și Ge pentru textele românești. Inventarul incomplet de caractere și repartizarea pot fi consultate pe pagina consorțiului<sup>2</sup>. De notat că lista nu a fost actualizată de Unicode după versiunea 9.0, însă blocurile numite mai sus pot fi găsite cu ușurință pe internet. Pentru simplitate ne vom referi la numerele caracterelor în sistem zecimal, în timp ce numerotarea Unicode în interfețe se face în hexazecimal.

Merită menționat că, de regulă, oamenii care au documentat și instrumentat demersurile pentru introducerea caracterelor slavone suplimentare în Unicode au contribuit și la alte aspecte legate de digitalizare și tehnoredactare. Michael Everson și Ralph Cleminson au creat fonturi polialfabetice care acoperă inclusiv scrierea slavonă (Everson Mono și Diliana)<sup>3</sup>, Vladislav Doroș îngrijește site-ul Irmologion, unde putem găsi o serie de fonturi și literatură utilă<sup>4</sup>, Sorin Paliga a conceput mai multe *layout*-uri<sup>5</sup> de tastatură, inclusiv românești și slavone<sup>6</sup>, iar Alexandr Andreev, Nikita Simmons și Yuri Shardt sînt în prezent administratorii proiectului Ponomar<sup>7</sup>, unde găsim fonturi, literatură de specialitate și suport tehnic excelent.

Aceluiași S. Paliga (2009, p. 179) îi aparține o trecere în revistă a soluțiilor existente la acea vreme pentru transcrierea textelor vechi slave în format electronic. Pe lângă conținutul informativ, articolul citat este util pentru a înțelege cum a evoluat instrumentarul disponibil astăzi față de cel existent acum mai mult de un deceniu.

## 3. Instrumente

### 3.1. *Layout*-uri de tastatură slavonă

Una dintre problemele care trebuie rezolvate în vederea culegerii unui text slavon este cea a tastaturii. Utilizatorii români folosesc cel mai des o tastatură cu *layout* QWERTY sau una dintre variantele ei pentru română (QWERTY cu diacritice), germană (QWERTZ), franceză (AZERTY). Pentru limba română există cîteva *layout*-uri, dintre care doar două—Romanian primary și Romanian secondary—sînt utilizate în mod curent<sup>8</sup>. În spațiul slav se folosește o variantă de IŢUKEN, însemnînd că la tastarea primelor șase caractere de pe rîndul de sus al tastaturii sînt afișate literele chirilice respective. Pentru introducerea textelor în slavonă sînt, așadar, cîteva soluții.

#### 3.1.1. *Windows*

1. Utilizarea tastaturilor existente în sistemul de operare. Utilizatorii de Windows au la dispoziție, începînd cu versiunea 10, două tastaturi slavone. Aceasta are la bază tastatura rusă de tip IŢUKEN (ИЦУКЕН) și nu necesită instalare.

<sup>2</sup>[unicode.org](http://unicode.org).

<sup>3</sup>[everytype.com](http://everytype.com) și [localfonts.eu](http://localfonts.eu). Ralph Cleminson este și curatorul proiectului Obshtezhitie, un portal prețios pentru studierea textelor chirilice și glagolitice: [obshtezhitie.net](http://obshtezhitie.net).

<sup>4</sup>[irmologion.ru](http://irmologion.ru), site-ul este în limba rusă.

<sup>5</sup>Aici și mai jos am preferat pentru simplitate termenul englezesc utilizat pe larg în informatică. În traducere românească se folosește uneori termenul „aranjament”.

<sup>6</sup>[unibuc.ro](http://unibuc.ro).

<sup>7</sup>[ponomar.net](http://ponomar.net), informații disponibile atît în rusă, cît și în engleză.

<sup>8</sup>Mai multe informații și resurse putem găsi pe blogul informaticianului Cristian Secară, cel care a elaborat modelul de tastatură românească conform standardului național SR 13392:2004: [secarica.ro](http://secarica.ro).

2. O tastatură virtuală precum cea propusă de Ponomar și de unele dicționare slavone on-line<sup>9</sup>.
3. O tastatură de tip QWERTY adaptată pentru slavonă. Acestea înlocuiesc pe cât posibil caracterele latine cu corespondentele lor slavone. Asemenea tastaturi pot fi găsite pe pagina profesională a lui Daniel Bunčić, profesor la Institutul de Slavistică al Universității din Köln<sup>10</sup>. Pagina oferă o gamă largă de *layout*-uri pentru chirilica modernă și veche: rus (ЙЦУКЕН), bulgar (УЕИШЩЦ), sârb (ЋАЕПТЗ), QWERTY slavon (АБЕРТЫ) și altele (polonez, glagolitic, multilingv, AZERTY).
4. O tastatură personalizată. Există multiple instrumente care ne permit crearea propriului *layout* de tastatură. Procesul se relevă accesibil și celor fără competențe informatice, însă este laborios și, în opinia noastră, nejustificat, avînd în vedere alternativele. Există și situații în care această soluție ar putea fi îmbrățișată, de exemplu în cazul celor care folosesc un *layout* DVORAK<sup>11</sup> și ar dori să mapeze caracterele slavone în funcție de propriile deprinderi de tastare.

Pentru utilizarea simplă și îndelungată avem două opțiuni: o tastatură IŢUKEN sau una QWERTY slavonă, fiecare cu avantaje și dezavantaje. Pentru utilizatorii români avantajul major al tastaturii QWERTY este că nu trebuie să învețe o nouă dispunere a caracterelor, iar asta este important mai ales dacă nu trebuie să lucreze frecvent cu corpusuri mari de texte slave. *Layout*-ul IŢUKEN are însă și avantaje. În primul rînd, ne va permite să îl folosim pentru a transcrie texte în limbi slave moderne, precum rusa. În al doilea rînd, aprecierea noastră este că celelalte niveluri de tastatură sînt dispuse mai eficient decît în QWERTY. Nivel de tastatură înseamnă setul de caractere pe care-l obținem în combinație cu tastele Shift, AltRight, ambele sau un deadkey. Cel mai familiar ne este Nivelul 2, obținut cu tastele Shift (sau cu CapsLock), prin care accesăm majusculele. Spre exemplificare, *layout*-urile de la Köln (inclusiv QWERTY slavon) conțin majusculele pe Nivelul 2 (adică se tastează cu Shift), iar multe dintre caracterelor speciale, în special literele suprascrise, se tastează cu AltRight sau cu un deadkey. IŢUKEN însă plasează literele suprascrise pe Nivelul 2 (adică se tastează cu Shift), iar majusculele pe Nivelul 3 (cu AltRight). Deși este adevărat că această repartizare pe niveluri este mai familiară utilizatorului român, în practică suprascrisele se întîlnesc mult mai des decît majusculele în textele slavone, de aceea posibilitatea de a le introduce cu Shift ni se pare mai ergonomică. Tot mai eficient ni se par aranjate diacriticele pe rîndul numerelor. La final trebuie menționat că tastatura QWERTY slavon de la Köln este de tip ISO, adică include un deadkey între Shift și Z. Dacă este folosită o tastatură de tip ANSI sau JIS<sup>12</sup>, atunci această tastă lipsește și asta îngreunează lucrurile.

### 3.1.2. Apple, Linux, Android

Utilizatorii care lucrează pe alte sisteme de operare (Apple OS X, GNU/Linux, Android) pot găsi pe site-ul Ponomar *layout*-uri gratuite însoțite de o descriere detaliată și de instrucțiuni limpezi de instalare și utilizare<sup>13</sup>. Pe pagina profesională a lui S. Paliga putem descărca gratuit mai multe tipuri de tastaturi destinate alfabetelor slave: glagolitic, slavon, ceh ș.a. Acestea sînt destinate utilizatorilor de sisteme Apple și conțin instrucțiuni de utilizare și răspunsuri la întrebări frecvente<sup>14</sup>.

Atunci cînd utilizatorul alege un *layout* de tastatură slavonă, poate ține cont de sistemul de operare (Windows, Mac, Linux sau Android), de tipul de *layout* (IŢUKEN, DVORAK sau QWERTY), de tipul de tastatură fizică pe care intenționează să lucreze (ISO, ANSI sau JIS) și de tipul de texte care trebuie editate. Unui utilizator român de Windows pe care nu-l interesează transcrierea într-o limbă slavă modernă i s-ar potrivi cel mai bine un *layout* QWERTY slavon (АБЕРТЫ) de felul celui de la Köln și o tastatură

<sup>9</sup> [gorazd.org](http://gorazd.org), de exemplu.

<sup>10</sup> [slavistik.phil-fak.uni-koeln.de](http://slavistik.phil-fak.uni-koeln.de).

<sup>11</sup> Deși reprezintă standardul de facto al tastaturilor actuale, *layout*-ul QWERTY nu este și cel mai eficient. Unii utilizatori preferă *layout*-ul ergonomic DVORAK, care permite culegerea mai rapidă a textului. Există și o variantă DVORAK pentru limba română, numită POPAK, după creatorul ei, Nicolae Popa. Mai multe aici: [invatasingur.ro](http://invatasingur.ro).

<sup>12</sup> Puteți determina ce tip de tastatură fizică aveți urmînd acest link: [wikipedia.org](http://wikipedia.org).

<sup>13</sup> [sci.ponomar.net](http://sci.ponomar.net).

<sup>14</sup> [dropbox.com](http://dropbox.com).

ISO. Se poate cumpăra separat o tastatură ISO cu taste goale pentru a se folosi în paralel cu tastatura obișnuită, iar pe tastele ei se pot desena manual caracterele slavone pentru o învățare mai ușoară.

### 3.2. Fonturi Unicode

Fonturile sînt în esență un inventar de glife/desene comun acceptate pentru litere și înmagazinează o corespondență între o glifă și codul Unicode al unui caracter, alături de instrucțiuni despre interacțiunea dintre caractere, de exemplu așezarea semnelor diacritice pe litere, kerning (spațierea estetică), ligaturile. Calculatorul primește informația despre tasta apăsată pe tastatură, fontul verifică codul caracterului în inventarul propriu (afișat de regulă utilizatorilor sub forma pangramei *The quick brown fox jumps over the lazy dog*) și solicită plăcii video să „deseneze” caracterul pe ecran. Dacă apăsăm, de exemplu, tasta Q în timp ce avem activat un *layout* QWERTY slavon, tastatura va trimite spre calculator informația că a fost apăsat caracterul cu codul Unicode 467 (în hexazecimal), iar fontul Garamond (folosit la scrierea acestui articol) îi va cere plăcii video să-l „deseneze” așa: А. Codul Unicode al unui caracter se află plasînd cursorul imediat după caracterul respectiv și apăsînd combinația Alt+x. Procedura funcționează și în sens invers. Astfel, cînd discutăm despre un caracter slavon mai rar, cum este iusul mic închis (А), ne putem referi, pentru claritate, și la codul său Unicode (A659 în acest caz).

Majoritatea covârșitoare a fonturilor folosesc astăzi coduri conforme Unicode, acum 25 de ani însă un creator de font slavon putea trasa arbitrar corespondențe între glifă și caracter. Utilizatorii mai experimentați își aduc aminte probabil de fontul Method, care desena pe ecran litere slavone, înmagazinînd, de fapt, codurile Unicode ale literelor alfabetului latin. Utilizînd doar fonturi conforme Unicode ne asigurăm că textul nostru înmagazinează exact caracterele pe care am vrut să le transcriem de o manieră standardizată, independentă de fontul, dispozitivul, softul sau sistemul de operare pe care le folosesc potențialii noștri cititori.

### 3.3. Ce fonturi se pot folosi

Depinde de scopul propus. Pentru o analiză strict textuală a unui verset biblic unde ne interesează puțin spre deloc dimensiunea grafică a textului, un font precum Calibri sau FreeSerif este suficient.

În versiune normalizată, acest pasaj din Tetraevanghelul slavonesc, Mt 4, 3, este perfect inteligibil: *Послъди възалка и пристѣплъ къ немѣ искѣсителъ, рече аще сѣи еси бѣи, рѣци да камене се хлѣби бѣдѣтъ.*

Dacă urmărim însă o transcriere mai exactă, tinzînd spre diplomatică, este nevoie să folosim un font slavon dedicat, care să conțină litere suprascrise, diacritice, litere iotate, consoane palatalizate, litere glagolitice, dublete grafice, semne speciale ș.a. Din fericire, în comparație cu acum 20 de ani, soluțiile sînt multiple, satisfăcătoare și adesea gratuite:

1. Dilyana: fontul conceput de R. Cleminson, destul de vechi, dar încă de utilitate<sup>15</sup>.
2. BukyVede: Creat de Sebastian Kempgen. Trebuie spus că fontul Method, de la același autor, a beneficiat de o versiune nouă, compliantă Unicode.
3. Familia CyrillicaBulgarian: fonturi concepute pentru proiectul CyrilloMethodiana a Universității „Sfîntul Climent Ohridski” din Sofia. A fost adoptat de alte proiecte precum Gorazd și eRomLex.
4. Fonturile proiectului Ponomar.org. Au destinație variată. Ponomar ne propune fonturi pentru slavona sinodală (Ponomar Unicode), pre-Nikoniană (Feodorovsk Unicode, care imită litera tipar-niței lui Ivan Feodorov), manuscrise unciale (Menaion Unicode), fonturi cu destinație academică (Monomakh Unicode), decorative (Indiction Unicode). Întrucît Monomakh este destinat cercetătorilor, conține și diacriticele limbii române, lucru care scutește pe utilizator de efortul schimbării tastaturilor la transcrierea concomitentă a unor texte bilingve.

Oricare dintre fonturile menționate reprezintă o soluție validă, cu mențiunea că preferința noastră personală se îndreaptă spre fonturile Ponomar, în special spre cel eponim și spre Monomakh. Acestea conțin setul Cyrillic Extended C, tratează ceva mai bine unele ligaturi și afișează mai armonios semnele diacritice.

<sup>15</sup>[fonts2u.com](http://fonts2u.com).

## 4. Probleme

### 4.1. Caracter necodificat în Unicode

Deși inventarul de caractere codificate Unicode este enorm și se îmbogățește constant, se poate întâmpla ca un caracter de care utilizatorul (cercetător sau amator) crede că are nevoie, să lipsească din inventarul Unicode, iar pe ecran apare o căsuță precum aceasta: ☒. Există două soluții:

1. Puteți cere consorțiului Unicode introducerea caracterului în standard și atribuirea unui număr de ordine. Trebuie să pregătiți documentația necesară, să aduceți argumente, treceți niște etape și dacă factorii de decizie consideră demersul oportun, se poate face codificarea. Un model de propunere acceptată puteți găsi pe pagina consorțiului<sup>16</sup>.
2. Între timp puteți folosi Private Use Area (PUA). Unicode a lăsat intenționat anumite poziții libere pentru ca utilizatorii să poată introduce în fonturile proprii simboluri și glife necodificate oficial. De exemplu, dacă vrem să redăm ca atare un text românesc cu multe ligaturi, cum sînt cele de cancelarie, în special din perioada târzie, putem crea un font propriu care să conțină glife pentru ligaturi și să le atribuim coduri din PUA. Toate fonturile slavone pe care le cunoaștem au codificate caractere sau ligaturi în PUA, chiar și cele expres Unicode (CyrillicaBulgarian10U, Monomakh, Dilyana), pentru că toate încearcă să suplinească astfel niște curențe. În general utilizarea caracterelor din PUA scade din uniformizare și standardizare, dar pe de altă parte, poate spori precizia transcrierii.

Un posibil candidat pentru punctul 1 de mai sus ne furnizează câteva dintre textele slavo-române timpurii, unde întâlnim o formă de rîță culcat. Unii cercetători, începînd cu Rosetti (1956, p. 116), au opinat că acest grafem face parte dintr-o serie de semne care notează un *r* apical vibrant. Nu intenționăm să ne expunem pe marginea cadrului teoretic. Dorim doar să evidențiem că cercetătorul care dorește să îl includă în transcrierile sale poate înainta consorțiului Unicode o propunere argumentată științific care, odată acceptată, îi va permite să transcrie pasaje precum cel din PsVor (Fig. 1) cu mai multă acuratețe. De altfel și Giuglea (1910, p. 445) l-a transcris printr-o variantă de sigma pentru a-l scoate în evidență.

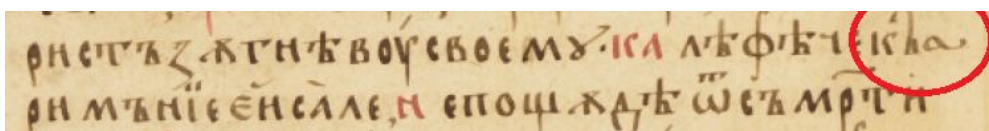


Figura 1: Psaltirea Voronețeană (PsVor, Ps. 77, 54)

Soluția alternativă este includerea grafemului în PUA, dar acest lucru trebuie făcut în faza de proiectare a fontului.

### 4.2. Dezambiguizarea morfologico-semantică

Ortografia slavonă folosește o serie de mijloace pentru antistih și pentru dezambiguizarea mesajului: alografe, diacritice, orientarea accentelor, alternanța vocalică. Vom da doar câteva exemple. Varianta lată de *slovo* (1C83, vezi Fig. 2) se folosește adesea în cuvinte care se referă la divinitate: adjective, pronume teoforice, numele lui Hristos. Pronumele demonstrativ *céŭ* se scrie frecvent cu litera de bază cînd se referă la obiecte și cu varianta lată cînd se referă la ființe. Tot varianta lată se folosește uneori pentru a deosebi omografele în cazuri oblice de cele în cazuri directe. Un alt exemplu de antistih este perechea de slove est normal (U+0435) și est lat (U+0454). Această alternanță poate fi întâlnită în texte tipărite care vor să facă distincția între *женѣ* = 'a(l) femeii' (gen. sg.) și *женѣ* = '(pe) femei' (nom. pl.). Textele slavo-române folosesc adesea slova *e* lat pentru poziția inițială a formelor verbului 'a fi', după cum arată orice analiză sumară a Ps1570.

<sup>16</sup>[unicode.org](http://unicode.org).

1C80	Cyrillic Extended-C	1C8F
0	Ѧ 1C80	<b>Historic letter variants</b> 1C80 Ѧ CYRILLIC SMALL LETTER ROUNDED VE → 0432 Ѧ cyrillic small letter ve
1	Ѧ 1C81	1C81 Ѧ CYRILLIC SMALL LETTER LONG-LEGGED DE → 0434 Ѧ cyrillic small letter de
2	Ѧ 1C82	1C82 Ѧ CYRILLIC SMALL LETTER NARROW O → 043E Ѧ cyrillic small letter o 1C83 Ѧ CYRILLIC SMALL LETTER WIDE ES → 0441 Ѧ cyrillic small letter es 1C84 Ѧ CYRILLIC SMALL LETTER TALL TE 1C85 Ѧ CYRILLIC SMALL LETTER THREE-LEGGED TE → 0442 Ѧ cyrillic small letter te 1C86 Ѧ CYRILLIC SMALL LETTER TALL HARD SIGN → 044A Ѧ cyrillic small letter hard sign

Figura 2: Blocul Cyrillic Extended-C

Mărcile de antistih și dezambiguizare pot fi folosite și concomitent: *нѣмоуѣ* = ‘a(l) slăbiciunii’ (gen. sg.) față de *нѣмоуѣы* = ‘slăbiciuni’ (nom-acc. pl.). Unele mărci și alternanțe sînt reproduse în mod curent de toate transcrierile științifice, cum este perechea *o/ω* în *ѣѡ* = ‘(pe) el’ (acc. sg.) și *ѣѡѡ* = ‘a(l) lui’ (gen. sg.). Altele sînt adesea trecute cu vederea. În această categorie cad, de regulă, caracterele din Cyrillic Extended-C, blocul slavon de data cea mai recentă în Unicode. Fonturile create înainte de 2016 nu sînt cele mai potrivite pentru afișarea acestor caractere; fie n-o să le conțină deloc, fie o să le aibă codificate doar în PUA. Fonturi precum Monomakh se dovedesc utile în acest caz. Unele dintre aceste variante sînt considerate formale, iar utilizarea lor nu poate fi prezisă algoritmic, cum este cazul perechii *з/ѣ*. În multe cazuri transcrierea lor nu îmbunătățește cu nimic transcrierea sau digitalizarea, însă există ediții și piste de cercetare care pot beneficia de pe urma transcrierii cît mai fidele a textului. Cîteva dintre caracterele acestui bloc împreună cu localizarea lor Unicode pot fi examinate în Fig. 2.

## 5. Ligaturi

Pentru ca ligaturile să fie afișate corect, designerul fontului trebuie să specifice comportamentul grafic al caracterelor care participă într-o ligatură. În cazul în care codificăm ligaturile în PUA, e nevoie să introducem între caracterele pe care le dorim ligate un tip special de caracter numit Zero Width Joiner (ZWJ, U+200D). Dacă fontul a fost programat corespunzător, va afișa o ligatură predesenată. Această strategie este folosită pe larg de alte scripturi care folosesc ligaturi: devangari, arab, dar pentru scrisul slavon lucrurile sînt încă în stadiu incipient.

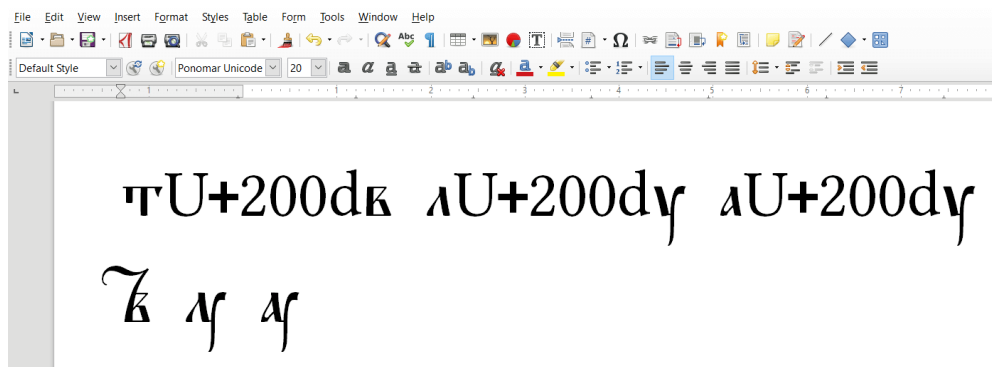


Figura 3: Ligaturile tverdi-vede, liudi-uc și az-uc în LibreOffice Writer

La momentul de față ligaturile sînt dificil de afișat în Microsoft Word. În procesorul de text al pachetului gratuit și open-source LibreOffice, ligaturile predesinate (precum *Ѧ Ѧ Ѧ* din fontul Ponomar) sînt afișate adesea corect. Dacă examinăm cu comanda Alt+x șirul de caractere slavone de mai sus, vom observa că între slove a fost inserat un ZWJ. În Fig. 3 vedem cum LibreOffice Word transformă automat secvențele slovă-ZWJ-slovă din rîndul de sus în ligaturile perfecte din rîndul de jos.

### 5.1. Spații

În tipografia slavonă întâlnim două tipuri de spații: normal (U+0020), îngust (pentru clitice, după numerale, unele particule, prepoziții). Acestea pot fi de tip break (spații obișnuite) și non-break (înainte de virgulă sau punct). Spațiile non-break formează o unitate tipografică cu caracterul pe care îl precedă. Acest lucru înseamnă că, într-un procesor de text, ele vor face trecerea în rînd împreună cu caracterul sau cuvîntul anterior. În Fig. 4 putem vedea mai multe tipuri de spații folosite la tipărirea *Apostolului* lui Coresi, cca. 1566.

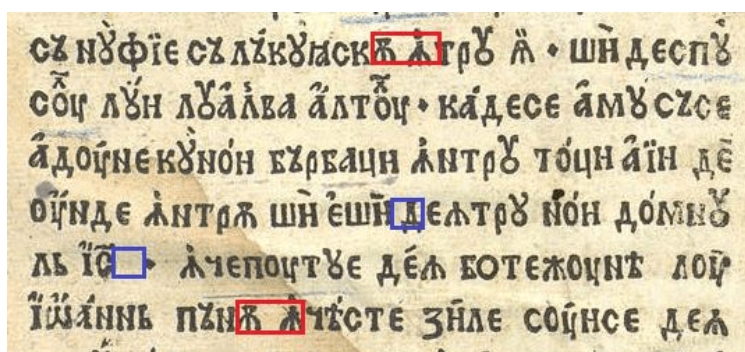


Figura 4: Spațiere normală și îngustă în *Apostolul* lui Coresi, cca. 1566

Spațierea îngustă poate fi utilizată și la transcrierea altor tipărituri coresiene. Se poate argumenta că în psaltirile de la 1570 și 1577 pot fi identificate urme ale gramaticalizării articolului hotărît românesc. I. Coteanu (*apud Nicolae, 2013*, p. 16) semnalase existența unui articol hotărît proclitic liber în formulări de felul „ispitire a ei voastre credință”. În general cliticul de dativ este deja sudat de substantiv la etapa primelor texte românești, însă anumite particularități de grafie sugerează o etapă anterioară, intermediară între cea descrisă de Coteanu și cea a normei actuale, unde demonstrativul nu se gramaticalizase complet și nici nu-și pierduse accentul. Putem exemplifica prin grafii precum *мѡніе ей* în PsVor (Fig. 1), *пеліце єи* / *ноѷстїи єи* în Ps1577 (Fig. 5 sus) sau *пеліцеєѷ* / *ноѷстїиєѷ* în Ps1570 (Fig. 5 jos). În toate cele trei exemple găsim convenții grafice care se folosesc pentru notarea vocalei inițiale: *e lat* (PsVor), *psili* (Ps1577) și *iso/psili+oxia* (Ps1570). În plus, în două dintre cazuri—PsVor și Ps1577—între lexem și articol pare să existe spațiu.

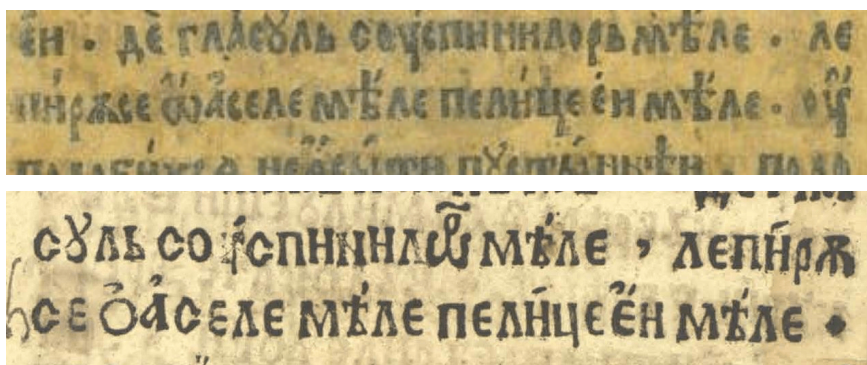


Figura 5: Ps1577 (sus) și Ps1570 (jos)

Ultimul tip de spațiu care poate fi folosit la transcriere este U+200B, Zero Width Space (break sau non-break). Îl putem utiliza pentru a separa cuvintele în textele cu *scriptio continua*. Astfel păstrăm textul inițial, dar permitem calculatorului să analizeze separat unitățile lexicale.

Asemenea tip de spațiu poate fi folosit la analize automatizate ale unor corpusuri slavone timpurii transcrise fidel.

Spațiile non-break au fost utilizate în cadrul proiectului eRomLex pentru a transcrie cazurile unde în manuscrise apar mai multe litere suprascrise pe aceeași literă de bază, iar fontul nu oferea soluții prestabilite (așa cum o face, însă, în cazul combinației slovo-tverdi suprascrise). Prima literă suprascrisă a fost notată pe litera de bază, iar celelalte au fost poziționate după litera de bază pe câte un spațiu non-break.

### 5.2. Manuscrise muzicale

Merită menționat pe scurt că site-ul proiectului Ponomar oferă câteva soluții și explicații pentru transcrierea manuscriselor muzicale. Prelucrarea lor este o sarcină dificilă, întrucât acestea conțin atât text, cât și notație muzicală, iar pentru a le transcrie împreună este nevoie de un soft dedicat sau de o soluție mai rafinată. În plus, multe dintre manuscrise nici nu sînt slavo-române, ci greco-române, motiv pentru care ies din aria noastră imediată de interes. Cei preocupați de psaltichii ar putea găsi în sursa menționată resurse care le-ar putea ușura procesul de lucru.

## 6. Concluzii

Acum 20 ani o ediție electronică slavonă perfect diplomatică era dificil de realizat. Astăzi lucrurile stau diferit. Ne bucurăm să constatăm că astăzi sîntem bine echipați în ceea ce privește setul de instrumente pentru editarea textelor slavone și slavo-române. Dispunem de *layout*-uri de tastatură pentru toate sistemele de operare, avem la dispoziție fonturi cu inventare aproape exhaustive de caractere, iar instrucțiunile de utilizare nu mai sperie utilizatorii neinițiați. Toate acestea ne permit să producem ediții de text cu un grad înalt de fidelitate față de originale. Editarea de text slavon și studiile care au la bază text slavon sînt dintre cele mai diverse și au obiective variate. Cel interesat de editarea textelor slavone, cercetător sau pasionat, poate decide care dintre instrumentele existente se potrivesc cel mai bine scopurilor propuse și mijloacelor de care dispune. O înțelegere mai bună a datelor tehnice poate scurta timpul de lucru, poate îmbunătăți transcrierea, poate micșora timpii de învățare și, per ansamblu, va ușura munca editorului. Sperăm ca articolul de față să se dovedească un ajutor cît de mic în acest sens.

## Bibliografie

### A. Referințe

- Giuglea, G. (1910). *Psaltirea Voronețeană: (mss. 693 dela Bibl. Acad. Rom.)* în „Revista pentru Istorie, Arheologie și Filologie”, 7, p. 444–467, [online].
- Macarie (1512). *Tetraevanghel slavonesc*, Târgoviște. Exemplarul cu cota CRV3 al Bibliotecii Academiei Române, disponibil în catalogul online al instituției.
- Nicolae, A. (2013). *Gramaticalizarea articolului hotărît românesc. Noi rezultate*, în „Limbă și literatură”, LIV (1–2), p. 7–19.
- Paliga, S. (2009). *Culegerea textelor vechi slave în format electronic*, în „Romanoslavica”, XLV, p. 179–188.
- Ps1570 = Coresi (1570). *Psaltire românească*, exemplarul cu cota BRV 16/1 al BCU Cluj, [online].
- Ps1577 = Coresi (1577). *Psaltire slavo-românească*, două exemplare păstrate la Biblioteca Academiei Române.
- PsVor = *Psaltirea Voronețeană*, ms. rom. 693 al Bibliotecii Academiei Române.
- Rosetti, A. (1956). *Limba română în secolele al XIII-lea – al XVI-lea*, Editura Academiei, București.
- Toma, S. (1976). *Psaltirea slavo-română (1577) în comparație cu psaltirile coresiene din 1570 și din 1589*, Editura Academiei, București.

### B. Resurse web

- [obshtezhitie.net](http://obshtezhitie.net)  
[ponomar.net](http://ponomar.net)  
[slavistik.phil-fak.uni-koeln.de](http://slavistik.phil-fak.uni-koeln.de)  
[unicode.org](http://unicode.org)