

**PETRAS SKIRMANTAS**

Lietuvių kalbos institutas

## KOL NEBUVO *PALEMONO*: KOMPIUTERIAI, LIETUVIŠKI RAŠMENYS IR ŠRIFTAI

ESMINIAI ŽODŽIAI: specialieji rašmenys, kodavimas, kodlapis, ASCII, ANSI, unikodas, teksto kirčiavimas, fonetinė transkripcija.

Apie lietuviškus kompiuterinius šriftus kalba vis atsinaujina: siekiama aiškintis, kiek, kokių ir kam skirtų jų yra, kuriais atvejais jie reikalingi, kaip plačiai paplitę ir t. t. Neseniai proga apie tai pakalbėti pasitaikė dar sykį, Valstybinei lietuvių kalbos komisijai ir Lietuvių kalbos institutui surengus diskusijų seminarą *Palemonas ir kiti lietuviški šriftai: naudojimas ir perspektyvos*. Šiuo rašiniu taip pat mėginama prisiliesti prie to, kas ten kalbėta.

Pačius imtis namudinio kompiuterinių šriftų darymo ar, teisingiau pasakius, modifikavimo privertė praktiniai poreikiai, kaip sakoma – pats gyvenimas... Praeito tūkstantmečio paskutinio dešimtmečio pirmojoje pusėje, po 1990 m., tuometinių kompiuterių ėmė atsirasti ir vis labiau gausėti net ir filologų, lituanistų aplinkoje. Pirma, jie tada ėmė darytis vis labiau įperkami, antra – nemažai jų būdavo atsivežama ar atvežama kaip labdara iš užsienio, nes užsieniečiai įsigydavo naujesnių, modernesnių, o senstelėjusius atiduodavo ar palyginti nebrangiai parduodavo mums. Kompiuteriai tada buvo gerokai kitokie nei dabartiniai, daug kuklesnių galimybių. Bet kadangi ir su anų laikų kompiuteriais buvo galima, be viso kito, dar ir rašyti, kaip tada pradėjome sakyti – rinkti tekstą, tai jie labai greitai išstūmė iš apyvartos rašomąsias mašinėles. Patogumas iš to buvo didžiulis ir lemiamas: rašant ar perrašant tekstą kompiuteriu netenka prasmės tokia sąvoka kaip *juodraštis*, nes visi klaidų pataisymai ar redaguojamojo pobūdžio tvarkymai pakeičia taisomas teksto vietas, ir tiek ekrane, tiek atspausdintas visą laiką matomas tik švarraštis. Taip netrukus filologams, kaip ir nemažai daliai kitokių darbus dirbančių žmonių, teksto rašymas, teksto rinkimas tapo pagrindiniu kompiuteriu dirbamu darbu.

Kadangi kompiuteriai pas mus patekdavo įvairiais būdais ir dažnai – ne pirmos jaunystės, tai iš pat pradžių kilo problema dėl galimybės jais rašyti taisyklingai lietuviškai: norėjosi ir reikėjo, kad tiek ekrane, tiek atspausdinti būtų matyti visi mūsų dabartinėje rašyboje vartojami

rašmenys, pirmiausia – visos dabartinės lietuvių abėcėlės raidės (ir *q, ę, į, ū, č, š, ž, é, ū*), o be to – ir kitokie tekstuose pasitaikantys rašto ženklai. To reikėjo ne tik filologams ar lituanistams, bet visiems, kas panorėdavo kompiuteriu rašyti ar tvarkyti lietuviškus savo srities tekstus. O galimybių būta ne visada...

Kiek nukrypstant į šalį derėtų priminti, kad jau anuo laiku kompiuteriai ir į filologų aplinką patekdavo dažniausiai dvejopi, ir vieni nuo kitų jie gerokai skyrėsi tiek galimybėmis, tiek ir patogumu dirbti (valdymo būdais bei apskritai „bendravimo“ su žmogumi, vadinamosios sąsajos, angl. *interface*, pobūdžiu). Vieni jų buvo sukonstruoti ir gaminami firmos *Apple*, kiti – *International Business Machines (IBM)*. Dar apie devintojo dešimtmečio vidurį firmos *Apple* pradėti gaminti serijos *Macintosh* (pavadintos pagal vieną iš obuolių veislių) kompiuteriai, paskui tapę labiau žinomi kaip *Mac*, nuo pat pradžių turėjo specialų valdymo įrankį – pelę ir grafinę sąsają su žmogumi (angl. *graphic user interface*, GUI), kuri beveik viską, ko reikia dirbant kompiuteriu, sudėliodavo į apgalvotus struktūriškus sąrašus, ekrane rodomus specialiai tam piešiamuose grafiniuose languose. Užtekdavo pele parodyti į reikiamą sąrašo punktą ir kompiuteris „suprasdavo“, ką turi daryti. Languose matomas ir veiksmo rezultatas – apdorojamas tekstinis ar grafinis dokumentas, jo būseną atlikus vieną ar kitą veiksmą. Pagrindiniai grafinės sąsajos principai išplito ir tebegalioja iki dabar.

Tuo pat metu firmos *IBM* pradėti gaminti panašūs serijos PC (žodžių junginio *Personal Computer* santrumpa) kompiuteriai kurį laiką grafinės sąsajos neturėjo. Visas jų valdymo komandas reikėdavo įvesti (surinkti) klaviatūra, o sąsaja su žmogumi buvo tik tekstinė, t. y. ir įvedamos komandos, ir kompiuterio veiksmų rezultatai ekrane būdavo rodomi tekstu, sudarytu iš ženklavietėmis apribotų vienodų, standartinio dydžio ir formos, rašmenų. Ženklovietė – tai stačiakampis standartinis laukelis, sudarytas iš tam tikro taškų, suskirstytų į vienodus stulpelius ir eilutes, skaičiaus. Klaviatūra ir ženklovietėmis suskaidytas ekranas kartu vadinami tekstiniu terminalu; jis kurį laiką ir buvo pagrindinė keitimosi informacija su IBM PC tipo kompiuteriu priemonė. Todėl dirbti tokiu kompiuteriu kebloka: reikia atsiminti valdymo komandas, kurių yra nemažai, ir teisingai, be suklydimų užrašyti jas ir joms būtinus perduoti parametrus (duomenis, nuorodas). Žmogus su tokiu kompiuteriu, galima sakyti, susirašinėja: jis ekrane standartiniu tekstu užrašo komandas ir duomenis, o kompiuteris – rezultatus, gautus komandas vykdant. Tokia tekstinio terminalo principu pagrįsta sąsaja kai kuriose operacinėse sistemose ir dabar tebėra pagrindinis keitimosi informacija su kompiuteriu būdas, itin mėgstamas profesionalų. Tačiau didžiuma žmonių vis dėlto labiau mėgsta jau minėtą grafinę sąsają, laiko ją patogesne. Tad natūralu, kad palaipsniui ji sukurta ir IBM PC tipo kompiuteriams; ji gavo

taiklų vardą *Windows*, tiesiogiai nurodantį į pagrindinį sąsajos elementą – grafinius langus. Apie XX a. paskutinio dešimtmečio vidurį *Windows* išaugo iki savarankiškos operacinės sistemos ir nuo to laiko grafinė sąsaja tapo pagrindine IBM PC tipo kompiuterių sąsaja su žmogumi. Abiejų tipų – ir *Apple* gaminamų *Macintosh*, ir IBM PC – kompiuterių valdymas, apskritai darbas su abiejų tipų kompiuteriais visai supanašėjo, nors esminių „vidinių“ skirtumų tebėra.

Filologų aplinkoje, kaip ir daug kur Lietuvoje, labiau plito IBM PC tipo kompiuteriai. Jie buvo prieinamesni, pigesni, pasakytum, demokratiškesni, į juos nesuku įdiegti įvairiopus kilmės programinę įrangą, pirmiausia, reikiamas taikomąsias programas, tarp jų – net ir savadarbes, pačių susikurtas ar užsisakytas pas pažįstamus programuotojus. Greta retokų originalių, firmos *IBM* pagamintų PC tipo kompiuterių sparčiai ėmė plisti Pietryčių Azijos šalyse pagaminti ar net iš ten pagamintų blokų čia pat, Lietuvoje, sumontuoti analogai, kuriuos imta vadinti kompiuteriais, suderinamais su IBM PC... Pažymėtina, kad panašių *Macintosh* imitacijų iš esmės nebuvo.

Filologams, besinaudojantiems daugiausia IBM PC ir su jais suderinamais kompiuteriais, didesnių problemų, susijusių su taisyklinga, standartine, lietuviška rašyba, retai tekildavo. Net ir operacinės sistemos DOS (angl. *Disk Operating System*) valdomuose anuometiniuose PC tipo kompiuteriuose, su kuriais keistasi informacija per tekstinį terminalą, jau buvo taikomos gana apibrėžtos rašto ženklų, ekrane vaizduojamų ženklavietėmis, kodavimo schemas, vadinamos kodlapiais arba kodų lentelėmis (angl. *code page*, CP). Jų buvo keletas, ir jos numatė kodus ne tik anglakalbių abėcėlės raidėms, bet ir kai kurių kitų kraštų rašybos sistemose vartojamiems ženklams. Kartu su operacine sistema buvo platinami ir kompiuteriniai šriftai (angl. *fonts*) – į operacinę sistemą integruojami programinės įrangos komponentai, leidžiantys ženklus pavaizduoti regimais, žmogaus akiai matomais grafinais vaizdais. Įtaisai, galintys tai daryti, yra du pagrindiniai: ekranas ir spausdintuvas. Problemiškesnis ir išradingesnio suderinimo su kompiuteriu reikalingas įtaisas buvo tuometinis spausdintuvas: daug kas priklausė nuo jo veikimo būdo, rūšies, gamintojo ir pan. Bet svarbu pabrėžti, kad abu šie įtaisai savo darbą padaro tik kompiuterio valdomi, o reikiamą esminę informaciją kompiuteris ima iš kompiuterinių šriftų.

Savaime suprantama, taisyklinga lietuviška rašyba nebuvo svarbiausias firmos *Microsoft*, kūrusios programinę įrangą kompiuteriams, suderinamiems su IBM PC (ir tebekuriančiai ją daugeliui dabartinių kompiuterių), rūpestis. Bet, kita vertus, dabartinės mūsų rašybos pagrindas – lietuviška abėcėlė – nėra nei originali, nei labai unikali, didžiuma jos raidžių yra bendros, jas savo rašyboje vartoja ne tik lietuviai, bet ir kitos tautos. Tikrai lietuviškų, niekieno daugiau bent

jau Europoje praktiškai nevarojamų raidžių tesama turbūt tik *é, ĳ* ir *ų*. Raidės *q* ir *ę* savo rašyboje vartoja lenkai, *č, š* ir *ž* – čekai, *ū* – latviai. Vis dėlto devynios raidės – *q, ę, ĳ, ų, č, š, ž, ū, é* – dažnai išskiriamos iš kitų lietuvių abėcėlės raidžių. Ir pagrįstai – visos jos yra dvigubos, dvinarės, sudarytos iš dviejų sudedamųjų dalių: iš pamatinės raidės (*a, e, i, u, c, s, z*) ir apačioje ar viršuje pridėto pagalbinio ženkluko – nosinės, paukščiuko, brūkšnio, taško. Visos kitos 23 mūsų abėcėlės raidės – *a, b, c, d, e, f, g, h, i, y, j, k, l, m, n, o, p, r, s, t, u, v, z* – yra viengubos, jas sudaro tik vienas dėmuo (abejonių dėl tokio teiginio keltų nebent mažosios *i* ir *j*, turinčios tašką viršuje, bet prisimintina, kad didžiosios *I* ir *J* taško neturi, o tiek *é*, tiek ir *Ē* jį turi, taigi šios raidžių poros skirtingos). Pažymėtina taip pat, kad viengubos raidės yra daug „tarptautiškesnės“ už dvigubąsias, jos vartojamos daugelio Europos šalių rašto sistemose ir visos 23 įeina į angliškąją abėcėlę. Itin svarbu, kad tos dvi raidžių grupės, formuojantis rašto ženklų kodavimo kompiuteriuose schemoms, papuolė į skirtingas kodų grupes. Viengubosios mūsų abėcėlės raidės, kartu būdamos ir anglų abėcėlės raidės, kodus gavo pirmiausia, ir tokius, kurie vėliau nebesikaitaliojo, išliko pastovūs iki šiol. Jos patenka į vadinamąjį standartinį ženklų rinkinį (angl. *standard character set*), kurį reglamentuoja senas, septintajame dešimtmetyje susiformavęs standartas ASCII (angl. *American Standard Code for Information Interchange*). Pagal jį ženklų kodavimui naudojami sveikieji skaičiai nuo 0 iki 127, iš jų rašto ženklams (raidėms, skaitmenims, skirtukams) – kodai nuo 32 iki 126. Šita kodavimo schema (kartais sakoma tiesiog *koduotė*) susidarė dar prieš atsirandant „mažiesiems“ kompiuteriams, ji buvo numatyta įvairiems telekomunikacijos įrenginiams, o kompiuteriai atsiradami ją tiesiog perėmė ir pasisavino.

Tačiau, kompiuteriams gausėjant ir plintant, standartiniame (ASCII) rinkinyje esančių ženklų ėmė nebepakakti, nes visų rašmenų, kurių prireikia tekstus rašant ne tik anglų, bet ir kitomis kalbomis, jame nebuvo. Dėl augančių kompiuterinės rašybos poreikių ženklų kodavimui imta naudoti daugiau sveikųjų skaičių – nuo 0 iki 255, nes kompiuterių vidinė sąranga taip daryti leido. Pirmieji 128 (nuo 0 iki 127) kodai, reglamentuoti ASCII, nepakito, bet atsirado papildomi 128 kodai (skaičiai nuo 128 iki 255), kuriais buvo galima koduoti kitus, į standartinį rinkinį nepatenkančius ženklus. Tokia padidinta kodų aibė (kodų erdvė) anglakalbių neretai yra apibūdinama kaip *Extended ASCII*, tačiau tai nėra ASCII (standarto ar juo reglamentuotos kodavimo schemas) papildymas ar išplėtimas – tai kodų aibės išplėtimas taip, kad ji aprėptų nepakeistą ASCII kodavimo schemą ir kartu pateiktų 128 papildomus kodus kitiems reikalingiems ženklams, kurių kodavimo ASCII nereglamentuoja. Taip atsirado 256 kodų paskirstymo modelis, visą kodų aibę suskirstantis į du lygaus dydžio poaibius – pastovųjį, reglamentuotą ASCII ir apimantį kodus nuo 0 iki 127, ir laisvąjį, apimantį kodus nuo 128 iki

255. Jį, tą modelį, kaip tik ir žymėtų angliškas apibūdinimas *Extended ASCII*. O jo įtvirtintas laisvasis kodų poaibis sudarė prielaidas modelio pagrindu atsirasti daugybei skirtingų kodavimo schemų, vadinamų, kaip jau minėta, kodlapiais arba kodų lentelėmis ir žymimų pačiais įvairiausiais žymenimis. Visų jų pastovusis poaibis, pastovioji dalis buvo vienoda ir sutapo su ASCII, bet laisvosios dalies kodų priskyrimas ženklams labai įvairavo, nelygu kokiam ženklų rinkiniui kodlapis būdavo sudaromas.

Vis gausėjančioje kodlapių įvairovėje ėmė rasti ir tokių kodlapių, kurių laisvojoje dalyje būdavo užkoduotos ir lietuvių abėcėlės sudėtinės (dvigubos) raidės, dažniau – tik tos iš jų, kurių reikia ne tik mūsų rašybai. Atsirado ir visoms lietuvių abėcėlės raidėms kodus numatančių kodlapių; prie to yra prisidėję taip pat ir mūsų kompiuterininkai. Kartu radosi ir tuos kodlapius atitinkančių kompiuterinių šriftų, leidžiančių numatytais kodais užkoduotus rašto ženklus iš tikrųjų naudoti kompiuteryje. Būta ir painingos: pirmojoje paskutinio dešimtmečio pusėje jau buvo net keli DOS operacinei sistemai skirti kodlapiai su visomis lietuvių abėcėlės raidėmis – jų numeriai buvo nuo 770 iki 774, paskui prisidėjo 775, 776, 777 ir 778. Lietuviškų raidžių kodai kai kuriuose iš tų kodlapių sutapo, kituose – skyrėsi. Gausėjant tuose kodlapiuose užkoduotus ženklus vaizduojančių kompiuterinių šriftų, reikėdavo susigaudymo ir apdairios. Dar reikėjo tuos įvairuojančius devynių dvigubų mūsų abėcėlės raidžių, neretai pavadinamų savitosiomis ar specifinėmis lietuviškomis raidėmis, kodus susieti su klaviatūra, atitinkamais jos klavišais (viengubų, standartinio ASCII ženklų rinkinio apimamų raidžių kodų sąsaja su klaviatūra dėl suprantamų priežasčių jokių problemų niekada nekėlė). Bet, kita vertus, visa tai sudarė galimybes taisyklingai lietuviškai rašyti net ir tais kompiuteriais, kuriuos valdė operacinė sistema DOS.

IBM PC tipo kompiuteriams skirta grafinė sąsaja *Windows*, kuri iš DOS priedo netruko pavirsti savarankiška operacine sistema (ja pirmoji tapo *Windows 95*), iš pagrindų pakeitė ir rašmenų vaizduoseną ekrane, ir kompiuterinių šriftų, ir apskritai rašymo kompiuteriais padėtį. Neliko ženklaviečių, įsitvirtino kitokie, naujo tipo, šriftai, vadinami vektoriniais, kurie regimąjį rašmenų vaizdą kuria vektorinės kompiuterinės grafikos priemonėmis – Bezjė kreivėmis nužymėdami rašmenų kontūrus. Gausėjo rašalinių ir lazerinių spausdintuvų, kuriais galima popieriuje tuos kontūrus atkurti labai tiksliai, su smulkiausiomis grafinėmis detalėmis. Dėl to tekstai ir dokumentai, parengti kompiuteriu ir atspausdinti tokiu spausdintuvu, tapo panašūs į spaustuvių spausdinamus dokumentus. Nemažai naujo tipo šriftų plito kartu su *Windows* grafine terpe, jie buvo integrali jos dalis, bet kartu būta galimybės prireikus „sisteminį“ šriftų rinkinį papildyti kitais, atskirai įsigytais šriftais.

Ženklų kodavimas taip pat ėmė nusistovėti. Nepaisant lokalių kodlapių įvairovės, vis svarbesni darėsi universalūs tarptautiniai kodavimo standartai. Labiausiai minėtini du: iš keliolikos palaipsniui kurtų dalių sudarytas ISO/IEC 8859 ir ISO/IEC 10646. Pirmuoju nustatyta, kaip didžiumai europinių rašto sistemų reikalingi ženklai būtų koduojami kodlapiais, turinčiais po 256 kodus, o antruoju daugybę kartų padidinta sveikųjų skaičių, galinčių būti ženklų kodais, aibė – ji tapo didesnė nei du milijardai! Didžioji jos dalis tebėra nepanaudota, iš viso tos aibės kodais užkoduotų ženklų kol kas tėra tik netoli 140 tūkstančių... Lygia greta radosi ir kitas ženklų kodavimo standartas (ne tarptautinis) – *Unicode*. Jis taip pat buvo skirtas padidinti ženklų kodų aibei. Jo apmatai parengti 1988 m., 1991 m. pradžioje įsteigtas *Unicode Consortium*, o rudenį išleistas unikodo standarto pirmasis tomas, pagal kurį ženklų kodais gali būti 65536 sveikieji skaičiai – nuo 0 iki 65535. Netrukus standartų ISO/IEC 10646 ir *Unicode* rengėjai sutarė dirbti koordinuotai, taigi abu standartai buvo (ir tebėra) derinami, sinchronizuojami, ir ženklų kodavimas jais nuo to laiko pasidarė visiškai vienodas, nors ir vieno, ir kito standarto paskirtis ir funkcijos liko savitos. Dėl įvairių priežasčių ir tam tikrų reikalavimų bendras iš tikrųjų naudotinių kodų skaičius, paprastai vadinamas kodų erdve, labai sumažėjo ir šiuo metu yra kiek didesnis nei 1,1 milijono; iš jų ženklams jau yra priskirti, kaip sakyta, apie 140 tūkstančių, taigi apie 12,5 procento, kodų. Abiejų standartų sąlaja paprastai vadinama unikodu. Pavadinimas motyvuotas, nurodantis į pastangas unifikuoti bet kokių kompiuteriuose naudojamų ženklų kodavimą, padaryti jį vienodą visame pasaulyje.

Minėtieji ženklų kodavimo standartai yra universalūs, orientuoti į informacijos technologijos poreikius – jie nebuvo kuriami tik IBM PC tipo kompiuteriams ar jų grafinei sąsajai. Tačiau, rengiant įvairių tipų bei paskirties kompiuteriams ir įvairioms operacinėms sistemoms skirtus kompiuterinius šriftus, stengiamasi kaip įmanoma labiau laikytis šių standartų nustatomo ženklų kodavimo, nes tai naudinga: unifikuotas ženklų kodavimas užtikrina, kad rašytinė, rašto ženklais perteikiama, informacija bus siunčiama, priimama ir atkurama tiksliai, nepriklausomai nuo to, kokia naudojama techninė ar programinė įranga.

Programinės įrangos gamintoja *Microsoft*, sukūrusi, tobulinusi ir platinusi grafinę sąsają *Windows*, kuri evoliucionavo taip sparčiai, kad po kurio laiko, vaizdžiai sakant, prarijo operacinę sistemą DOS ir pati virto operacine sistema, kartu platinamiems vektoriniams šriftams naudojo savus kodlapius, aprėpiančius po 256 ženklų kodus. Jie gavo savitus pavadinimus, bet vienas nuo kito ir nuo DOS kodlapių skyrėsi tik laisvuju kodų poaibiu, t. y. kodų nuo 128 iki 255 priskyrimu ženklams. Pastovusis, ASCII reglamentuojamas poaibis irgi nepakeistas (beje, nepakeistas jis ir visose ISO/IEC 8859 dalyse bei unikode). Su pirmosiomis *Windows* versijomis

buvo pradėtas naudoti kodlapis, pažymėtas 1252 numeriu (*code page 1252*). Vėliau jis vadintas ir sutrumpintai *CP-1252*, ir *Windows-1252*, ir tiesiog *Latin 1*. Paskui kodlapių daugėjo, kartu daugėjo ir jais koduojamų rašmenų, grindžiamų tiek lotyniškuoju, tiek kitais raidynais. Siekiant pabrėžti, kad tai ne DOS, o *Windows* terpei skirti kodlapiai, kartais jie vadinami ANSI kodlapiais, nes pirmojo iš jų, CP-1252, pirminė versija buvo grindžiama Amerikos nacionalinio standartų instituto (*American National Standards Institute*, ANSI) rengto kodlapio projekto pirminiu variantu. Tas projektas vėliau išaugo į kodlapį ISO/IEC 8859-1, bet ANSI vardas vis tiek prigijo darbinėje *Microsoft* terminijoje ir tapo grupiniu *Windows* kodlapių pavadinimu. Pridurtina, kad operacinei sistemai DOS skirti kodlapiai taip pat turėjo bendrą, grupinį pavadinimą – OEM.

Kodlapiai, kaip jau ne kartą minėta, yra ženklų kodavimo schemas, numatančios, kokiam ženklui kokį kodą – sveikąjį skaičių, ne didesnę kaip 255 – priskirti. Kodas – ženklo skaitmeninis identifikatorius kompiuterinėje sistemoje, jo atpažinimo sutartinis žymuo – yra beveik tas pat, kas šaukinys radijo stočiai ar asmens kodas Lietuvos piliečiui. Anglakalbėje kompiuterijos literatūroje kodą dabar įprasta vadinti taikliu, vykusiai padarytu žodžiu *codepoint* (kartais parašoma ir *code point*), taip nurodant, kad tai kodų aibės punktas, vienas iš tos aibės skaičių. Vis dėlto pabrėžtina, kad kodas nėra koks imanentiškas ženklo atributas, slypintis jo „viduje“ – ženklui jis parenkamas ir priskiriamas laisvai, remiantis tam tikru susitarimu. Todėl tas pats ženklas, jeigu jis patenka į skirtingus kodlapius ir (ar) skirtingus kodavimo standartus, iš principo gali gauti skirtingus kodus, nors tai nepageidautina ir vengtina. Kodai reikalingi ne tik ženklams koduoti kompiuteriniuose šriftuose, jų pirmiausia reikia tam, kad rašmenimis išreikšta informacija galėtų koordinuotai cirkuliuoti kompiuterinės sistemos viduje, kad įvairūs informaciją siunčiantys ir priimantys programiniai bei aparatiniai kompiuterio dėmenys ją išsaugotų nepakeistą, jos neiškraipytų, nepaverstų atsitiktinių signalų srautu, t. y. informaciniu triukšmu, entropija.

Kai kompiuteriai patobulėjo ir atsirado galimybė kodams naudoti daug kartų už 255 didesnius sveikuosius skaičius, kodlapių svarba kiek pablėso. Unikode, kodų erdvę labai išplėtusiame standarte, kodlapio sąvoka nebeartojama, skiriamos kitokios rašmenų grupės, susijusios su rašmenų paskirtimi bei vartoseną, ir jos vadinamos skriptais, o kodų erdvės padalos dažniau vadinamos blokais. Tačiau *Microsoft* nuosekliai laikosi vadinamojo suderinamumo iš apačios į viršų (kartais sakoma – istorinio suderinamumo) principo: siekiama, kad ir naujuose, modernių *Windows* versijų valdomuose kompiuteriuose galėtų veikti taip pat ir senosios, praeityje naudotos programos. Dėl to dabartinės *Windows* versijos palaiko dvejopą ženklų



kodavimą: ir naujovišką, suderintą su unikodu, ir senovišką, atitinkantį anksčiau sudarytus ANSI kodlapius. Minėtą informacijos cirkuliaciją kompiuterio viduje valdo sisteminėse bibliotekose esantis programinių funkcijų rinkinys *Windows API* (API – iš angl. *Application Programming Interface*), ir daugelio jo funkcijų, skirtų darbui su simboliais bei jų sekomis, esama po dvi versijas: A, skirta ANSI kodais koduotiems simboliams, ir W (angl. *wide*), skirta unikodo kodus turintiems simboliams ir jų sekoms. Kadangi atitiktis tarp įvairių kodlapių ženklui priskirtų ANSI kodų ir unikodo priskirto kodo yra vienareikšmiška, nesunku vienokius kodus „abiem kryptimis“ keisti į kitokius – tereikia žinoti, su kuriuo ANSI kodlapiu jais keistis. Ženkliai į ANSI kodus prireikus perkoduojami dėl kai kurių vidinių sistemos ypatumų ir dėl vis dar pasitaikančių senų taikomųjų programų, kurios nepripažįsta kodų, didesnių už 255. Šiuolaikinės tekstams skirtos taikomosios programos paprastai manipuliuoja ilgaisiais (angl. *wide*) kodais užkoduotais rašmenimis. Unikodinis rašmenų kodavimas dabar yra norma.

Taisyklinga lietuviška rašyba, įsitvirtinant vektoriniams šriftams, buvo pasirūpinta. Tai irgi mūsų kompiuterininkų nuopelnas. Laikas, kai kompiuterių, išstumiančių rašomąsias mašinėles, ėmė vis gausėti ant mūsų filologų rašomųjų stalų, maždaug sutapo su grafinės sąsajos *Windows* pirmųjų versijų įsitvirtinimo laiku. Tiesa, normaliai ir naudingai ta sąsaja galėjo veikti tik naujesniuose kompiuteriuose, turinčiuose ne menkesnę kaip *Intel 80386* procesorių ir VGA standarto vaizdo adapterį (vaizdaplakštę) bei ekraną, bet tokie būdavo toli gražu ne visi filologų tuomet įsigjami kompiuteriai. Besikeičiant *Windows* versijoms reikalavimai kompiuterio aparatūrai (vadinamajai techninei įrangai) vis augo ar, tiksliau pasakius, atvirkščiai: didėjant techninės įrangos galimybėms, susidarydavo prielaidos naujoms, patobulintoms, grafinės sąsajos *Windows* versijoms atsirasti. Kalbamuojau laiku jau ėmė plisti versija *Windows 3.1*, kiek vėliau – *Windows 3.11*, ir jau būta vektorinių šriftų su visomis lietuviškomis raidėmis. Pirmiausia čia minėtinas *Fotonijos* parengtas šriftas *TimesLT*, labai paplitęs ir lietuvių ilgai naudotas. Jo raidžių piešinys imitavo itin populiarią garnitūrą *Times*. Netrukus atsirado ir *TimesLT* atitikmenų su kitokias garnitūras imituojančiais raidžių piešiniais, lietuviškas raides koduojančių taip pat, kaip ir *TimesLT*. Ženkliai šiuose šriftuose jau buvo koduojami pagal Lietuvos ir Latvijos kompiuterininkų tuomet kurtą bendrą kodlapį užmojo kupinu pavadinimu *BalticRIM*, atseit Baltijos apvadas ar Baltijos žiedas... Gražu, žinoma, buvo ir tai, kad jame netikėtai atsispindėjo tuomet veikusios ir informacijos technologijos reikalus kuravusios Ryšių ir informatikos ministerijos abreviatūra. To kodlapio ANSI dalyje buvo sutelkti rašmenys, reikalingi didžiumos Baltiją supančių šalių rašybai, bet nepatenkantys į jo ASCII dalį. Be abejo, tarp jų buvo ir devynios savitosios, dvigubosios mūsų raidės. O pradėdant *Windows-95 PE* versija (PE – iš



angl. *Pan European*) ši kodlapį *Microsoft* faktiškai įteisino kaip Baltijos kraštams skirtą kodlapį *Windows-1257* (pavadinama *Windows Baltic*, *CP1257*, *CP1257 BalticRIM* ir pan.). Juo lietuviškos raidės ir visi Baltijos kraštų rašmenys koduojami tokiais pat kodais, kokie jau buvo įsitvirtinę Lietuvoje su *TimesLT* bei kitais tą kodavimą atitinkančiais šriftais. Taip susiklostė ir įsitvirtino su pavadinimu *Baltijos žiedas* asociacijų keliantis ženklų rinkinys, nugulęs į dar du kodlapius – tuomet tebe naudotai operacinei sistemai DOS skirtą 775 ir vėliau, 1998-aisiais, paskelbtą tarptautinio standarto ISO/IEC 8859 tryliktąją dalį (*ISO/IEC 8859-13:1998 Information technology – 8-bit single-byte coded graphic character sets – Part 13: Latin alphabet No. 7*), kuri dažnokai neoficialiai taip pat vadinama *BalticRIM*, *Baltic RIM* ar *Baltic Rim*. Ženklų rinkinys yra toks pat, o kodavimas, išskyrus kabučių bei apostrofo kodus, iš esmės sutampa tik *Windows-1257* ir ISO/IEC 8859-13; ženklų kodavimas OEM kodlapiu 775 gerokai skiriasi.

Vektoriniai kompiuteriniai šriftai iš esmės pakeitė kompiuteriu rengiamų tekstinių dokumentų išvaizdą. Jie ėmė labai skirtis nuo mašinėle perrašytų dokumentų, arba mašinraščių, ir vis labiau panašėti į spaustuvėse poligrafijos priemonėmis spausdinamus spaudinius. Rašmenų pavidalui kurti pasitelkiama vektorinė kompiuterinė grafika sudarė galimybę pagal poreikį, sumanymą ar norą plačiame diapazone keisti raidžių dydį (kegelį) nė kiek nedeformuojant jų kontūrų, taip pat tapo įmanoma, turint tinkamą vektorinių šriftų komplektą, keisti bet kokio dydžio raidžių ryškumą (jas suformuojančių brėžių plotį), norimas raides pateikti kursyvu, manipuliuoti intervalais tarp raidžių ar jų eilučių. Išraiškinga, poligrafiška kompiuteriu parengtų dokumentų išvaizda žmonėms patiko ir įtiko, filologams – taip pat, tad jų pagrindine rašymo kompiuteriais priemone kone iškart tapo ir iki šių dienų tebėra viena iš *Microsoft* platinamo taikomųjų programų paketo *Microsoft Office* sudedamųjų dalių – tekstams skirta programa *Microsoft Word*, nes ji leidžia minėtas vektorinių šriftų galimybes palyginti lengvai panaudoti praktiškai. Dar ir šiandien pasitaiko filologų, kurių nepavyksta įtikinti, kad testus galima rašyti ne tik programa *Word*, kad kai kurias atvejais racionaliau būtų juos rengti kitomis programinėmis priemonėmis. Geležinis argumentas – su *Word* išeina gražiausiai...

Kompiuterinių šriftų lietuviškos raidės ir kartu galimybė lietuvišką tekstą rašyti taisyklingai, nepažeidžiant dabartinės rašybos normų, tenkino ir tebetenkina daugelio raštus rengiančių įstaigų ir daugelio žmonių, kuriems rašymas yra sudedamoji, o neretai ir pagrindinė jų profesinės veiklos dalis, praktinius poreikius. Rašytojui, žurnalistui, literatūros kritikui, kino scenaristui ar panašios profesijos žmogui tos galimybės dažnai ir pakanka, ko nors daugiau kaip ir nereikia. Tačiau filologai, ypač kalbininkai savo rašytiniuose darbuose naudoja ne tik įprastas lietuvių abėcėlės raides, bet ir daug visokių kitokių ženklų, įvairių įvairiausių rašmenų, kartais

net ir labai specifiškų bei neįprastų. Suprantama, jog anksčiau naudotos rašomosios mašinelės negalėjo tų specialiųjų ženklų atspausdinti, todėl buvo įprasta į filologinio teksto mašinraštį juos įrašyti ranka. Mašinraštis dažniausiai būdavo (išskyrus galbūt disertacijas ir vieną kitą mašinėle perrašytą kitokio pobūdžio veikalą) tik tarpinis teksto variantas, lyg ir pagalbini priemonė, o galutinį grafinį pavidalą tekstas įgaudavo spaustuvėje, kur jis būdavo pagal mašinraštį ištisai perrenkamas metalinėmis literomis. Jų įvairovė buvo daug didesnė už mašinelės apimamą ženklų rinkinį. Dažniausiai spaustuvė rasdavo metalinių literų, atitinkančių ir ranka įrašytus specialiuosius ženklus, o jeigu nerasdavo – arba pagamindavo jas atskirai, pagal užsakymą, arba prašydavo autoriaus tokį ženklą pakeisti kitu, turimu. Pereinant nuo mašinelės prie kompiuterių pamatę, kad kompiuteriu į tekstą galima įkelti gerokai daugiau ir daug įvairesnių ženklų, negu jų būdavo rašomojoje mašinėle, filologai pradėjo norėti, kad jų tekstuose būtų galima naudoti visus jiems reikalingus ženklus – visokius visokiausius. Įgyvendinti tokio pobūdžio norus nebuvo nei neįmanoma, nei labai sudėtinga: užteko į turimą reikiamos garnitūros (dažniausiai *Times* šeimos) šriftą koku nors būdu, kad ir savamokliškai, įterpti papildomus, jame nesančius ženklus. Rasta tai padaryti leidžiančių būdų bei priemonių. O pademonstravus praktiškai, kad tokios galimybės esama, kad ženklų rinkinį kompiuteryje iš principo galima praplėsti ir papildyti savo nuožiūra, norai ir pageidaujamų specialiųjų ženklų skaičius ėmė didėti kone eksponentiškai... Taip prasidėjo keistoka kompiuterinių šriftų modifikavimo praktika ir ji truko ne vienus metus.

Filologų darbams, labiausiai – įvairių lingvistikos sričių, reikalingus specialiuosius ženklus galima suskirstyti į keletą grupių. Paminėtinos šios:

1) *lietuviškos kirčiuotos raidės*. Jos reikalingos, beje, ne vien tik filologų moksliniams veikalams, bet ir kitų sričių raštijai – vadovėliams ir kitai mokomajai literatūrai, įvairiems žodynams, žinyams ir t. t. Skiriant didžiąsias ir mažąsias, iš viso susidaro 68 kirčiuotos lietuviškos raidės. Įprastinėje lietuvių abėcėlėje jų nėra nė vienos, bet kai kurios iš jų randamos kitų kraštų abėcėlėse.

2) *lietuvių kalbininkų vartoti ir vartojami fonetinės transkripcijos ženklai*. Lietuvių kalbininkai tarminei kalbėsenai užrašyti dažniausiai naudoja ne tarptautinės fonetinės abėcėlės (TFA, angl. *International Phonetic Alphabet*, IPA) ženklus, bet savitą fonetinę rašybą, kurios ištakos – vadinamoji kopenhaginė transkripcija ir kurią lietuvių tarmėms pritaikė Jurgis Gerulis (vok. Georg Gerullis), o vėliau detalizavo ir papildė Aleksas Girdenis.

3) *raštijos paveldui publikuoti reikalingi specialieji ženklai*. Senojoje lietuvių raštijoje randama daug tokių rašto ženklų, kurie į dabartinę lietuvių abėcėlę nepateko. Bet cituojant

rašytinį palikimą įprasta tuos ženklus, jeigu tik techniškai įmanoma, atkurti ir pavaizduoti tokius, kokie jie yra originale. Kai kurie kalbininkai, pavyzdžiui, profesorius Vincas Urbutis, į tai žiūrėjo gana skrupulingai.

4) *specialieji ženklai, reikalingi istorinės gramatikos, baltistikos bei indoeuropeistikos darbams*. Dažniausiai tai būna įvairių abėcėlių (lotyniškosios, graikų, kirilicos) rašmenys, papildyti įvairiais specifiniais pagalbiniais ženklais.

Filologų darbų galutinis tikslas buvo ir daugiausia tebėra spausdintas pavidalas: visiškai pabaigtu darbas laikomas tada, kai jis tampa knyga ar straipsniu kokiame nors leidinyje. Kol spaustuvės spaudinius gamindavo senoviškai – perrinkdavo tekstą metalinėmis literomis, paskui jį sulaužydavo į puslapius (sumaketuodavo) ir iš jų, sudėliotų reikiama tvarka, pagal lankus, liedavo spaudos formas, tol spaustuvei skirtas tekstas, net ir kompiuteriu parengtas bei atspausdintas, turėjo iš esmės tokį pat statusą, kaip anksčiau mašinraštis. Todėl ir jame įspausdinti, o ne ranka įrašyti specialieji ženklai tebuvo daugiau grožio ar kiek savitai suprantamo prestižo dalykas – spaustuvė juos vis vien perrinkdavo savomis literomis. Kitaip sakant, specialius ženklus dar vis buvo galima įrašyti ir rankomis – spaudinyje to nebuvo matyti. Tačiau apie paskutinio dešimtmečio vidurį padėtis ėmė iš esmės keistis: kadangi tobulesni šriftai ir lazeriniai spausdintuvai jau leido spausdinti taip, kad atspaudų išvaizda ėmė prilygti tradicinei poligrafinei, spaustuvės pradėjo atsisakyti „metalinio“ spaudinių perrinkimo ir spaudos formas ėmė gaminti tiesiai iš kompiuteriu spausdintų puslapių. Puslapių maketavimas to lūžio metu pasidarė nebe spaustuvių, o leidyklų ir leidėjų problema, spaustuvės pageidavo gauti atspausdintus gatavus puslapius. Pradžioje buvo reikalaujama juos spausdinti ant specialių plėvelių (skaidrių), vėliau pakakdavo virtualiu spausdintuvu (tokia specializuota programa) spausdinti į PDF formato failą. Trūkstantį specialųjį ženklą ar jo komponentą, pavyzdžiui, priegaidę žymintį ženkliuką, iš principo dar būdavo galima ranka įrašyti ir puslapio skaidrėje, bet PDF faile to padaryti nebeįmanoma. Ranka įrašyti ar papildyti rašmenys tokie patekdavo ir į tiražą, būdavo matomi kiekviename spaudinio egzemplioriuje ir, išsiskirdami iš kitų ženklų, gadindavo bendrą spaudinio vaizdą. Taip visų ženklų, įskaitant ir specialiuosius, turėjimas kompiuteryje, anksčiau buvęs labiau prabanga, tapo būtinybe. Be jų parengti kokybiškų spaudinių nebebuvo galima.

Spaudinių, kuriems specialiųjų ženklų nereikia, rengyba tokių spaudos technologijos pokyčių labiau nepajuto. Tiesiog leidyklose atsirado papildomo darbo, maketavimo, bei jį dirbančių žmonių, atsirado techninė įranga, leidžianti spausdinti kokybiškas parengtų maketų skaidres. Bet daug ir įvairių specialiųjų rašmenų reikalingų lingvistinių leidinių rengėjus ir

leidėjus tas lūžis palietė gana skaudžiai. Leidyklos savo kompiuteriuose daugelio tų rašmenų dažniausiai neturėjo, ir nebuvo aišku, kiek ir kokių jų gali prireikti. Todėl kai kurios iš jų ėmė linkti prie minties, kad joms, kaip ir spaustuvėms, geriau ir patogiau būtų gauti jau gatavus, sumaketuotus kalbinius leidinius, faktiškai – kompiuterinius jų maketus. O tos, kurios pačios ryždavosi leidinius maketuoti, vertėsi kaip išmanydamos. Dėl tų priežasčių Vilniaus universiteto Filologijos fakultete imta rūpintis nebe tik dalykine, bet ir technine kalbotyros mokslo darbų žurnalų *Kalbotyra* ir *Baltistika* rengyba, jų maketavimu. Kartu imta galvoti ir apie kitų leidinių, kai kurių lingvistinių knygų, techninį parengimą spaudai. Taigi be specialiųjų lingvistinių ženklų – nė žingsnio...

Pradėta nuo lietuviškų kirčiuotų raidžių. Tiesa, jų stygių ankstėliau jau buvo pajutusi UAB *Fotonija*, rengusi įvairias priemones lietuviškai rašybai kompiuteriu palaikyti ir pradėjusi rengti pirmuosius kompiuterinius žodynus. Ji modifikavo kai kuriuos DOS kodlapius su lietuviškomis raidėmis, įterpdama į juos kirčiuotas lietuviškas raides ir žodynams reikalingus fonetinės transkripcijos ženklus (taip atsirado kodlapiai 776, 777 ir 778). Labai populiari tapo ir daug kieno ilgai buvo naudojama *Fotonijos* šrifto *TimesLT* atmaina su kirčiuotomis lietuviškomis raidėmis *TimesLTAccented*, irgi, suprantama, turinti žodynams reikalingus transkripcijos ženklus. Nekirčiuotos lietuvių abėcėlės raidės *TimesLTAccented* šrifte išlaikė tokius pačius kodus kaip ir *TimesLT*, t. y. kodlapio *Windows-1257* kodus, o kirčiuotos į ANSI dalį buvo įtrauktos vietoj kitų kraštų rašmenų ir taip „pasisavino“ pastarųjų kodus. Todėl *TimesLTAccented* šriftu parengtas lietuviškas kirčiuotas tekstas gerokai skirdavosi nuo nekirčiuoto: visos kirčiuotos raidės gaudavo kitokius kodus negu nekirčiuoti jų variantai, ir tokio teksto nebeįveikdavo nei rašybos tikrinimo, nei skiemonavimo-kėlimo programinės priemonės. Tai buvo akivaizdus kliuvinys kompiuterinei lingvistinių tekstų leidybai. Tad imta galvoti apie kitokį kirčiuotų tekstų rengimo būdą.

Mintį, tiesą sakant, pakišo pats *MS Word*, kuriame paryškintieji (angl. *bold*) ir kursyviniai (angl. *italic*) fragmentai gaunami keičiant šriftus. Rašmenų kodai tuose šriftuose nesikeičia, bet keičiasi rašmenų piešinys, grafinis jų vaizdas. Nei rašybos tikrinimui, nei automatiniam kėlimui tai nė kiek netrukdo. Taip kilo mintis, kad keičiant šriftą kirčiuota raidė galėtų pakeisti nekirčiuotą ir atvirkščiai. Sumanyta padaryta: sukurta *Times* garnitūros šrifto modifikacija *LT-Times*, apimanti visas keturias stilistines atmainas – normalią statmenąją, normalią kursyvinę, paryškintą statmenąją ir paryškintą kursyvinę. Raidės koduotos pagal *Windows-1257* kodlapį, todėl buvo gerai rodomi ir tekstai, surinkti *TimesLT* šriftu. Porai papildomų ženklų priskirti *Windows-1257* neapibrėžti ANSI kodai. Tačiau svarbiausia, kad prie tos ketunarės *LT-Times*

šeimos buvo pridėtos dar trys lygiai tokios pat šeimos išplėstais pavadinimais *LT-Times-GR*, *LT-Times-AK* ir *LT-Times-CF*. Papildymas GR čia nurodo į kirčio ženklą gravį, AK – į akūtą, CF – į cirkumfleksą. O atitinkamai pavadintuose šriftuose visoms lietuviškoms raidėms, galinčioms turėti tą kirčio ženklą, jis buvo pridėtas, nupieštas ant viršaus. Taip susidarė 16 šriftų šeima, leidžianti raidžių kirčiuotumą keisti lygiai taip pat, kaip keičiamas jų ryškumas ar kursyviškumas. Tai leido lengvai kirčiuoti jau surinktą tekstą, o sukirčiuotą vienu ypu paversti nekirčiuotu, labai paprastai vienokį raidės kirčio ženklą keisti kitokiu ir t. t. Rašybos tikrinimui ar automatiniam žodžių kėlimui toks kirčiavimas nė kiek netrukdo, tam skirtos pagalbinės programos jo nė nemato. Ir kaip matys – juk kirčiuotos ir nekirčiuotos raidės iš tikrųjų tesiskiria tik piešiniais, regimaisiais atvaizdais, kurie matomi ekrane arba spausdinami, o kompiuterio viduje jos yra tokios pačios, jas išreiškia tie patys kodai. Kirčiavimas, aiškiau sakant, čia yra ne paties teksto, bet jo vaizdavimo kai kuriose programose, tokiose kaip *MS Word* ar maketavimo programos, dalykas, panašus į rašmenų formatavimą. Lingvistinių veikalų rengybai ir leidybai visa tai tiko ir buvo gana paranku. Pavyzdžiui, dėl kirčiavimo nieko nereikėjo keisti kompiuterio klaviatūroje, nes tekstas būdavo kirčiuojamas naudojant *MS Word* turimas rašmenų formatavimo priemones, papildytas savadarbiais, VBA programėlių (skriptų) pagrindu veikiančiais įrankiais. Be viso kito, toks kirčiavimo būdas kiek priminė kalbos moksluose taikomą požiūrį į kirtį bei priegaidę, kad jie esą ne garsų, o skiemenų požymiai ar ypatybės, su garsais nepainiotini ir priskirtini atskiram, supersegmentiniam, lygmeniui. Taip kirčiuojant, kirčius ir priegaides žymintys ženklai ne priskiriami tekstui, o iškeliami į kitą, faktiškai – formatavimo, lygmenį, kuriame jie tik padedami ties kirčiuojamo skiemens centrą žyminčia raide. Tiek kirčiuoto, tiek ir nekirčiuoto teksto raidės kompiuterio „viduje“ yra tos pačios, koduojamos vienodai. Beveik visuotinai išplitęs ir dažnai kartojamas žodžių junginys *kirčiuotos raidės* nėra labai tikslus, nes raidės – tai grafiniai garsų atitikmenys, taigi, nuosekliai laikantis minėtojo požiūrio, daug kur labiau tiktų sakyti *kirčiuoti skiemenys*. Deja, kompiuterinėje rašyboje ir spaudoje priemonių, kurios leistų kirčiuotumą kaip nors susieti būtent su skiemenimis, o ne su atskiromis raidėmis, nėra.

Aprašytasis kirčio ženklų perteikimo būdas, tikęs anuometinei „popierinei“ leidybai, turėjo ir didžiulį trūkumą: kirčiavimą iškeliant į formatavimo lygmenį, iš principo nebuvo galima kirčiavimo perteikti ten, kur formatavimo nebuvo! Kitaip sakant, išvis nebuvo galimybės kirčiavimą perteikti vadinamuoju grynuoju (angl. *plane*) tekstu. Todėl niekam, išskyrus spaudinių rengimą, toks kirčiavimo būdas netiko. Jo praktinio pritaikymo sritis buvo iš tikrųjų siaura ir ribota. O iš sudarytojo 16 šriftų *LT-Times* rinkinio visaverčiai tebuvo tik keturi vienos

šeimos šriftai – tie, kurių visos lietuviškos raidės buvo be kirčių. Kiti 12 šriftų buvo šiaip jau pagalbines tų keturių šriftų modifikacijos, skirtos aprašytam kirčių raiškos būdai perteikti. Beje, pastarosiomis nepaisyta vieno pačių didžiausių šios srities tabu – ASCII dalies neliečiamumo: kirčio ženklai įterpti ir virš viengubųjų *a, e, i, o, u, l, m, n, r*, patenkančių į standartinį ženklų rinkinį ir koduojamų ASCII kodais...

Analogiškai buvo sudaromi ir kiti leidybai reikalingų specialiųjų lingvistinių ženklų rinkiniai, tiksliau – tiems rinkiniams skirti šriftai: į vieną, pradinį šriftą buvo sukeliama pamatiniai rinkinio rašmenys, o jų variantai, papildyti pagalbiniais ženkliais, diakritikais ir t. t., buvo dedami į kitus to rinkinio šriftus. Kiekvienam tokiam specialiajam ženklui paprastai skiriama po atskirą šriftą. O nesistemiški, atsitiktiniai, šiaip labai reti ar tiesiog keisti specialieji ženklai, nepritampantys nė prie vieno rinkinio, būdavo dedami į atskirą pagalbinių šriftą, dažnai vadinamą *Varia*. Einant laikui tų *Varia* ėmė rasti ne vienas, pavadinimus teko pildyti ir kaitaloti, ir netruko atsirasti įprotis po panašų pagalbinių šriftą daryti kone kiekvienam didesniai kalbiniam veikalui ar leidiniui.

Kone kartu su kirčiuotomis raidėmis kompiuteryje pasigesta ir lietuvių kalbininkų vartojamų fonetinės transkripcijos ženklų. Jų, be abejo, reikėjo tiek *Kalbotyrai*, tiek ir *Baltistikai*, bet poreikis vis dėlto buvo nesistemiškas, jų teprareikdavo kartkartėmis, ir buvo galima verstis pagalbiniais šriftais, kiekvienam tomui sudarant po savą *Varia*. Tų žurnalų leidyba, po 1990 m. trumpam stabtelėjusi, 1993–1994 m. buvo atnaujinta ir ėmė vėl įsibėgėti. Į 1994 m. pabaigą ir 1995-aisiais juose jau nebebuvo nei ranka piešiamų kirčio ženklų, nei kitokių ranka įrašomų rašmenų. Maždaug tuo pat metu amžinatilsį profesorius Aleksas Girdenis, daugelio čia aprašomų dalykų įkvėpėjas, palaikytojas ir tikrasis jų *spiritus movens*, rengė spaudai savo gimtosios tarmės transkribuotų tekstų knygą, ir kai darbas ėmė artėti prie pabaigos, kilo klausimas: o kaip ją išleisti, kaip sumaketuoti ir išpausdinti? Tarminės transkripcijos ženklų, tokių, kokių nėra nė viename kompiuteryje, tame darbe vartota daugybė, didesnioji knygos teksto dalis parašyta vien tik tai jais. Leidykloms tokio teksto nė nesiūlyk! Teko imtis specialaus šriftų komplekto, kuris leistų pavaizduoti reikiamus tarminės transkripcijos ženklus, sudarymo. Pagal knygos autoriaus eskizus ir žodinius apibūdinimus buvo parengti tų ženklų vektoriniai piešiniai, kodai ženklams parinkti savo nuožiūra ir „be tabu“ – tiek iš ANSI, tiek ir iš ASCII zonos. Taip atsirado šriftas su pamatiniais fonetinės transkripcijos ženklais. Modifikuojant jo kopijas, prie pamatinių ženklų prilipdyti papildomi ženkliai, diakritikai, kiekvienas perteikiantis vis kitokį ženklų žymimo garso tarimo atspalvį. Susidarė aštuonių kompiuterinių šriftų rinkinys su fonetinės transkripcijos ženklais, kiekvienas šriftas perteikė vis kitus atspalvius. Kadangi gimtoji profesoriaus tarmė yra

žemaičių ir į knygą, suprantama, buvo sudėti transkripcija jo užrašyti šnekos pavyzdžiai iš gimtųjų apylinkių, tai ženklų rinkinys pavadintas *Zemtra* – imant pavadinimo *Žemaitiška transkripcija* žodžių pradžią. Naudojantis juo ir anksčiau aptartu šriftų komplektu *LT-Times*, žemaitiškų transkribuotų tekstų knyga surinkta, sumaketuota ir atiduota leidyklai, o ji 1996 m. knygą pavadinimu *Taip šneka tirkšliškiai* išleido. Visa tai parodė, kad knygų rengimo spaudai naudojant kompiuterį galimybės yra išties didelės ir lanksčios. Tokios knygos išleidimas taikant „metalinės“ spaudos technologiją būtų kėlęs itin didelių problemų ir kažin ar iš viso būtų buvęs galimas: greičiausiai būtų tekę transkripcijos ženklų rinkinį siaurinti, šnekos transkribavimą derinant prie spaustuvės galimybių, o dabar buvo galima knygos ženklus ir jų įvairovę derinti prie autoriaus sumanymo ir pasirinkimo. Kitaip sakant, ženklus jau buvo galima derinti prie poreikių, ženklų parūpinant tokių, kokių tik prireikė. Anksčiau tekdavo poreikius derinti prie turimų ženklų.

Galima spėti, jog ši A. Girdenio kruopščiai parengta transkribuotų tekstų knyga buvo postūmis ir pavyzdys, skatinantis rengti analogiškas kitų tarmių, patarmių ir šnektų tekstų knygas. Daugėjo tokios veiklos besiuimančių kalbininkų, aktyviai įsitraukė Lietuvių kalbos institutas. Fonetinės transkripcijos raštu užrašomų gyvosios žmonių kalbos pavyzdžių geografija sparčiai plėtėsi, apimdama įvairius Lietuvos regionus, ir šiaurinių žemaičių šnekai transkribuoti pritaikyto *Zemtra* komplekto ženklų ėmė nebepakakti, bent jau ne visi jų ir ne visais atvejais tiko. Derinant *Zemtros* rašmenis su kitų savadarbių šriftų rašmenimis, šiaip taip verstis buvo įmanoma, bet neparanku. Lietuvių kalbos institutas susikūrė savų šriftų su tarmių transkribavimui ir kitoms reikmėms skirtais ženklais (*TimesDialect0 – TimesDialect2, Alt, Altt* ir kt.), universitete buvo kuriami vis nauji, papildomi *Varia... Žodžiu*, aiškėjo, kad reikia kito fonetinei transkripcijai skirto šriftų komplekto – tokio, kurio ženklais būtų galima išreikšti bet kurios lietuvių kalbos tarmės ar šnektos fonetikos subtilybes. Profesorius A. Girdenis kaip naujo tarminės rašybos ženklų rinkinio pagrindą parinko J. Gerulio transkripciją, ją papildė ir aiškiai nurodė, kaip kuris fonetinis ženklas turėtų atrodyti. Tų ženklų vektoriniai piešiniai, susieti su vėlgi savo nuožiūra ir „be tabu“ parinktais kodais, buvo sudėti į dešimt šriftų, jie pažymėti skiriamaisiais numeriais nuo 0 iki 9, o visas kompletas pavadintas *Fontra* – sujungus pavadinimo *Fonetinė transkripcija* pradžias. *Fontra* apėmė iš esmės visus *Zemtros* ženklus ir turėjo kitokių, reikalingų kitų regionų šnekai perteikti. Šis kompletas labai palengvino beįsibėgėjančią tarminių tekstų knygų ir kitų kalbinių leidinių rengybą ir leidybą. 1998 m. *Kalbotyroje* (t. 47, sąs. 1, p. 107–121) buvo išspausdintas straipsnis, kuriame aiškinta, kaip *Fontra* sudaryta, kam ji skirta, kokia jos ženklų paskirtis, vartosenos ypatumai. Išstūmusi *Zemtrą*,



*Fontra* įsitvirtino ilgokam laikui, kone dešimtmečiui – iki 2004–2005 metų, o tada ją pradėjo išstumti modernus, iš pagrindų kitoks *Palemonas*. Į ano tūkstantmečio pabaigą ir naujojo pradžia fonetiniu raštu, transkripcijos ženklais, spausdinamų knygų gerokai pagausėjo, jas imta leisti vieną po kitos, ir *Fontra*, ko gero, tam bus turėjusi įtakos.

Beje, profesorius A. Girdenis *Fontra* labai mėgo, laikė ją vykusiu fonetinės rašybos įrankiu, labiausiai tinkančiu transkribuotiems tarmių tekstams spausdinti, norėjo, kad jo parengti tekstai būtų spausdinami būtent ja. *Palemono* A. Girdenis nepamėgo... Todėl 2008 m. išleista antroji profesoriaus rengta tarminių tekstų knyga *Žemaičių dzūkai* išspausdinta *Fontra*, nors tada jau daug kur ją keitė *Palemonas*. Tik tai trečiasis žemaitiškų tekstų rinkinys *Šiaurės žemaičių Skuodo zona*, išleistas jau profesoriui mirus (†2011), jo mokinių buvo konvertuotas į *Palemoną*... Jis išspausdintas 2013 m.

Trečiasis lingvistikos darbams reikalingus specialiuosius ženklus perteikiančių šriftų komplektas, sudarytas lygiai taip pat, kaip du jau aprašyti ir pavadintas *IndoBalt*, yra skirtas indoeuropeistikos, baltistikos bei istorinės gramatikos veikaluose pasitaikantiems savitiesiems rašmenims ir simboliams. Jis gerokai palengvino *Kalbotyros* ir *Baltistikos* tekstų rengimą ir buvo naudojamas leidžiant kitus kalbotyros darbus. Turint tris minėtus šriftų komplektus – *LT-Times*, *Fontra* ir *IndoBalt* – kalbotyros darbams reikalingų specialiųjų ženklų trūkumo Vilniaus universiteto Filologijos fakultete nebejausta, juolab kad prireikus galima nesunkiai pasidaryti pagalbinį šriftą, dar vieną *Varia*... Raštijos paveldo Filologijos fakultetas tada beveik nepublikavo, tuo iš esmės užsiėmė Lietuvių kalbos institutas, taigi ir su specifiniais paveldo rašmenimis susijusios problemos gulė ant jo pečių. Spręstos jos panašiai – pasidarant šriftų su reikiama ženkla.

Tokia būtinumo skatinama „namudinės“ šriftadarystės praktika, primenanti itin populiarią dar nesenų tarybinių laikų nuostatą „pasidaryk pats“, leido palyginti sėkmingai dorotis su lingvistinės literatūros rengybos ir leidybos iššūkiais, kilusiais iš kompiuterių plėtros nulemtų spaudos technologijos pokyčių. Tapo įmanoma parengti bei išleisti beveik bet kokio poligrafinio sudėtingumo lingvistinius veikalus. Tačiau pradėjo ryškėti ir neparankumai: savadarbių šriftų vis daugėjo, didėjo bendras specialiųjų ženklų skaičius, ženklai ėmė dubliuotis, kartotis... Sužiūrėti ir suvaldyti vis besiplečiantį šriftų ir ženklų ūkį darėsi kebloka. Sulėtėjo tekstų rinkimas: klaviatūros klavišų visiems ženklams niekaip nebeužteko, dažnai tekdavo raustis po vis storėjantį ženklų lentelių aplanką ieškant kokio keistesnio rašmens... Itin problemiška buvo tekstus, kuriuose naudoti savadarbiai „specialios paskirties“ šriftai, perkelti iš kompiuterio į kompiuterį, persiųsti. Juk reikėjo, kad kompiuteryje būtų įdiegti visi perimamame

tekste panaudoti šriftai. Iš tikrųjų tai reiškė, kad rengiamas ar parengtas tekstas, kompiuterinio dokumento failas, turi būti siunčiamas kartu su jame panaudotų šriftų komplektu, ir tai buvo negerai: kas gi norės į savo kompiuterį įdiegti keletą ar kartais net keliolika neaiškios kilmės ir kokybės, nesertifikuotų šriftų vien tam, kad galėtų teisingai perskaityti vieną kitą dokumentą? O kompiuterių ryšiai plėtėsi, vis funkcionalesnis ir landesnis darėsi internetas, ėmė rasti tiek laikmenomis platinamų, tiek ir internetui skirtų elektroninių žodynų, kiek vėliau – elektroninių knygų. Greta tradicinių publikacijų, leidžiamų kaip spaudiniai, ėmė rasti elektroniniai variantai. Ta linkme pradėjo sukti ir kalbiniai leidiniai. Ir savadarbiai specializuoti laisvo „be tabu“ kodavimo šriftai, visiškai tikę tradicinei leidybai, ėmė darytis didelė ir itin sunkiai įveikiama kliūtimi.

Darbai su tekstais skirta programinė įranga taip pat patyrė permainų. Populiariausia ir filologams labiausiai įprasta rašymo priemonė *Microsoft Word* nuo aštuntosios versijos, labiau žinomos kaip *Word 97*, leido naudoti kompiuterinius šriftus su daug daugiau nei 255 ženklų: juose, kaip numatė tada jau sutvirtėjęs, autoriteto įgavęs unikodas, jų galėjo būti šimtai ir tūkstančiai. Savadarbiai aprašytų komplektų šriftai tokie nebuvo, bet besiverianti didžiųjų šriftų perspektyva atrodė patraukli. Juk viename tokiaame šrifte galėtų tilpti visi filologų darbams reikalingi specialieji ženklai! Dar liktų vietos ir paprastoms, tradicinėms įvairių alfabetų raidėms bei jų variantams, nes į vieną tokio formato šriftą, nors jis irgi nėra beribis, galima sudėti daugiau kaip 65,5 tūkstančio skirtingų ženklų. Tą perspektyvą dar realesnę padarė ties tūkstantmečių riba firmos *Adobe* išleista nauja spaudinių maketavimo programa *InDesign*, kuri taip pat palaikė unikodinį rašmenų kodavimą ir juo grindžiamus šriftus (prieš tai naudotos maketavimo programos *Adobe PageMaker* visos versijos, išskyrus paskutinę, septintąją, išleistą jau po *InDesign* pasirodymo ir greit nunykusią, tepripažino tik ANSI tipo kodus). Vieno visus lingvistikai reikalingus ženklus turinčio šrifto idėją į priekį stūmė dar ir tai, kad kompiuteriu transkribuoti tarminiai tekstai pradėti ne tik leisti knygomis, bet ir kaupti kompiuterinėse saugyklose, duomenų bazėse: šriftų su fonetinės transkripcijos ženklais gausa ir įvairovė pradėjo virsti į maišalynę, o ženklų kodavimo laisvė vedė, tiesiai sakant, į chaosą. Tai, suprantama, varžė ženklų rinkinių plėtojimo galimybes, kėlė abejonių dėl jų funkcionalumo bei naudos. Žodžiu, naujo, daugiafunkcio, t. y. tinkančio ir spaudai, ir internetui, ir duomenų bazėms, šrifto poreikis buvo jaučiamas visur.

Valstybinė lietuvių kalbos komisija ėmėsi iniciatyvos, kad toks šriftas atsirastų. 2002 m. rudenį į pirmąjį darbinį posėdį susirinko jos suburta darbo grupė naujam šriftui sukurti. O 2006 m. rudenį, per parodą *Infobalt 2006*, užbaigtas, patikrintas, iš esmės jau nuo 2004 m. knygų

leidybai naudojamas naujasis šriftas *Palemonas* buvo pristatytas visuomenei. *Palemoną* kūrusi grupė, be to, parengė ir 2005 m. išleido knygą *Lituanistinis šriftas Palemonas*, kurios didžiąją dalį sudaro konkreti techninė informacija – pagal pamatines raides sugrupuoti ženklų sąrašai su nurodomu kiekvieno ženklo unikodiniu pavadinimu ir jo kodavimu, taip pat schemas, padedančios reikiamą ženklą greitai surasti. *Palemonas* yra bene vienintelis taip išsamiai dokumentuotas lietuviškas kompiuterinis šriftas.

Bedirbant prie *Palemono*, apie 2003 metus pasirodė pirmeivis – *Fotonijos* sukurtas tokio paties formato komercinis šriftas *Aistika*, jis buvo viešai pristatomas kaip pirmasis lietuviškas kompiuterinis šriftas. Nors dėl to galima ir paabejoti (argi tos pačios *Fotonijos* anksčiau platinti *TimesLT* ir *TimesLTAccented* buvo nelietuviški?), *Aistika* vis tiek buvo pirmasis tikrai vykęs šios srities darbas, lietuviškus rašmenis, ne tik įprastinius, bet ir nemenką dalį specialiųjų, įkūnijantis moderniu unikodinio kodavimo šriftu. *Aistika* buvo skirta kitiems *Fotonijos* produktams, jos platinamiems kompiuteriniams žodynams, bet kartu leido kokybiškai rašyti tiek standartinėmis, tiek ir kirčiuotomis lietuviškomis ar rusiškomis raidėmis, apėmė graikiškus rašmenis, nemažai transkripcijos ženklų ir kitokių simbolių. Tačiau visų filologiniams leidiniams reikalingų ženklų nenumatė ir *Aistika*. Todėl *Palemonas* tebėra reikalingas ir pasirodžius *Aistikai*, juolab kad jis buvo sumanytas kaip nemokamas, laisvai platinamas šriftas ir toks tebėra iki dabar. Abu šie unikodinio kodavimo šriftai tarpusavyje dera, suderintas ir visų kirčiuotų raidžių (net ir tų, kurios neturi unikodinių kodų) kodavimas. Todėl ir paprastas, ir kirčiuotas lietuviškas tekstas, surinktas vienu kuriuo iš jų, gali būti atkuriamas kitu jo neiškraipant.

Sukūrus *Palemoną*, lingvistiniams ir lituanistiniams darbams reikalingų specialiųjų rašmenų problema išspręsta iš esmės.

Gauta 2019 12 15

BEFORE THERE WAS *PALEMONAS*: COMPUTERS,  
LITHUANIAN GRAPHEMES, AND FONTS*Summary*

The article is an account of the difficulties that had to be overcome in preparing philological publications on Lithuanian linguistics, containing special graphemes (beyond the standard Lithuanian alphabet) in the last decade of the 20th century, when the typesetting technologies underwent radical changes with the advent of computer typesetting. An emphasis is placed on the importance and possibilities opened by the graphical interface in the IBM PC-type computers and the vector fonts that came with it. The topics discussed include character encoding schemes current at the time and the situation of the Lithuanian alphabet therein. The main focus of the article however is on the special linguistic characters used in Lithuanian philology (beyond the standard Lithuanian alphabet letters) and the technical challenges associated with them. The topics discussed include in-house-made special computer fonts and collections of fonts, used at the University of Vilnius Faculty of Philology at the time, for the publication of the periodicals *Baltistica* and *Kalbotyra* as well as linguistic books containing phonetically transcribed texts in Lithuanian dialects, with Lithuanian accentuation (syllable tone) marks, the so-called 'Gerulian' phonetic alphabet (i. e. phonetic transcription system based on the so-called Copenhagen Scheme of Transliteration and Phonetic Transcription and introduced into the study of Lithuanian dialects by a prominent scholar Jurgis Gerulis) as well as other characters necessary for this type of publications that were not available in the standard computer fonts of the time. It must be admitted that the sphere of usage of these in-house-made fonts was rather narrow and confined to publishing of linguistic texts on paper; thus, these sets of fonts were not a universal instrument of information exchange and this circumstance rendered them inferior to other, standard fonts. Difficulties arising from this non-universality and the proliferation of varieties of these privately made fonts are also discussed; the issue especially came into prominence when the time had come to publish linguistic texts on the internet and to assemble them into sortable and searchable databases. These difficulties and the new demands arising thereof are the main reason for creating a modern, universal computer font that should encompass all graphemes necessary for linguistics texts, especially those of Lithuanian linguistics.

KEYWORDS: font, special characters for linguistics, encoding, character set, code page, encoding standards, ASCII, ANSI, Unicode, accent notation in writing, phonetic transcription

PETRAS SKIRMANTAS  
Lietuvių kalbos institutas  
Petro Vileišio g. 5, LT-10308 Vilnius  
[petras.skirmantas@gmail.com](mailto:petras.skirmantas@gmail.com)