

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document**Title: Proposal to add characters used in Lithuanian dialectology to the UCS****Source: Vilnius University: Faculty of Mathematics and Informatics, in cooperation
with the Faculty of Philology of VU and the Institute of Lithuanian Language****Authors: Vladas Tumasonis; Karl Pentzlin****Status: Expert Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2010-09-24****1. Introduction**

It has always been the standard practice in the study of Lithuanian dialects to write the dialect texts in detailed phonetic transcription. The phonetic alphabet used to transcribe Lithuanian dialects, based on the so-called Copenhagen Scheme of Phonetic Transcription and Transliteration of 1925, an erstwhile rival to the IPA – see

http://books.google.com/books?ei=L0GXTLiMOIPLOMXaxlgJ&ct=result&hl=lt&id=R8lsAAAAMAAJ&dq=Phonetic+transcription+and+transliteration&q=association+internationale+#search_anchor

– was adopted for the study of Lithuanian dialects by Jurgis Gerulis (also known by the German form of his name as Georg Gerullis, 1888–1945) in 1930 (see References), and finally elaborated by the leading specialist in Lithuanian dialects, Aleksas Girdenis (b. 1937). The Lithuanian Phonetic Transcription is in many ways similar to the International Phonetic Alphabet, and shares some characters with this; but it also has significant differences. Some Lithuanian scholars have proposed switching to the pure IPA system, but it seems that it might be too late: over time, too many linguistic works have been produced using the traditional system, so that the transition would be too difficult to be practical. As it is, nowadays we refer to the Lithuanian Phonetic Transcription System (or Lithuanian Phonetic Alphabet), which is the standard means of recording Lithuanian dialects on paper for all linguistic purposes. There are still a number of characters from the Lithuanian Phonetic Transcription which have not yet been covered by the Unicode Standard. The present application is to submit these characters for the inclusion into the Unicode set.

2. Encoding Considerations

The intonation marks are given as modifier letters, in analogy to all characters in the Modifier Tone Letters block (A700...A71F). Also, they behave like letters on kerning and linebreaking, and some of them have special positions (like “middle”) which refer to letter design.

As it is to be expected for phonetics, all text specimens are found typeset in italics. However, the intonation marks (AB90...AB99) are slanted by design, retaining the symmetry of the angle of the upward and downward marks independent of the appliance of italics. On the other hand, the vertical lines (2E46, 2E47) follow the slant angle of the font, and therefore have vertical reference glyphs.

3. Proposed Characters

Block: Latin Extended-E (new block; U+AB30...U+AB8F)

The block position and size is coordinated with the "Preliminary proposal to encode Germanicist phonetic characters in the UCS" (N3907 = L2/10-346) by Michael Everson.

The code points of the characters contained here are positioned in the last column, to avoid any conflicts with that document. Especially, the code positions U+AB8C, U+AB8E, and U+AB8F denote the same characters contained there; especially the gap for the "modifier letter baseline small e" proposed there is intentional.

The placement of the proposed letters in this block follows the guideline (proposed by Michael Everson in conjunction with the aforementioned proposal N3907) that new case pairs continue to go into the Latin Extended-D block, while unpaired lowercase letters go into this block.

Additions for Lithuanian dialectology

ċ	U+AB80	LATIN SMALL LETTER C WITH PALATAL HOOK (see fig. 1930a-6, 2001a-342, 2001a-344a)
ħ	U+AB81	LATIN SMALL LETTER H WITH PALATAL HOOK → A727 latin small letter heng (see fig. 2001a-344a)
ṃ	U+AB82	LATIN SMALL LETTER M WITH INSIDE HOOK (see fig. 2001a-344a/b)
ṅ	U+AB83	LATIN SMALL LETTER N WITH INSIDE HOOK (see fig. 1930a-XIX, 1930a-26b, 2001a-344a/b)

Additions for Lithuanian and German dialectology

a	U+AB8C	MODIFIER LETTER BASELINE SMALL A (see fig. 1930a-89)
	(U+AB8D)	reserved code position for a letter contained in a separate proposal for German dialectology (N3907 = L2/10-346) (MODIFIER LETTER BASELINE SMALL E)
i	U+AB8E	MODIFIER LETTER BASELINE SMALL I (see fig. 1930a-13)
u	U+AB8F	MODIFIER LETTER BASELINE SMALL U (see fig. 1930a-XV, 1930a-13)

Block: Phonetic Extensions Supplement-B (new block; U+AB90...ABAF)

This block is intended for phonetic characters which are not Latin by nature, and therefore are not appropriate to be included in the Latin Extended-E block.

(Beside the characters proposed here, there are some characters used in Slovenian dialectology derived from Cyrillic and Greek letters, which may be proposed later. Therefore, the size of two columns seems appropriate.)

The block would have been named "Phonetic Extensions Extended-A" if the (informal) naming scheme for new blocks would be followed; but this name would sound strange.

Intonation marks for Lithuanian dialectology

↑	U+AB90	MODIFIER LETTER SLANTED UP ARROW = sharp rise in tone (see fig. 2001a-347)
↓	U+AB91	MODIFIER LETTER SLANTED DOWN ARROW = sharp fall in tone (see fig. 2001a-347)
↗	U+AB92	MODIFIER LETTER SLANTED UP ARROW ENDING IN BEND = increasing tone with falling trend at the end (see fig. 2001a-347)
↘	U+AB93	MODIFIER LETTER SLANTED DOWN ARROW STARTING WITH BEND = sharp rise and fall in tone (see fig. 2001a-347)
↗̣	U+AB94	MODIFIER LETTER SMALL MIDDLE SLANTED UP ARROW = slight rise in tone (see fig. 2001a-347)
↘̣	U+AB95	MODIFIER LETTER SMALL MIDDLE SLANTED DOWN ARROW = slight fall in tone (when at the end of a word) = overall fall in tone (when at the beginning of a phrase) (see fig. 2001a-347)
↗̣̣	U+AB96	MODIFIER LETTER SLANTED UP ARROW STARTING WITH HORIZONTAL LINE = continued rise in tone (see fig. 2001a-347)
↘̣̣	U+AB97	MODIFIER LETTER SLANTED DOWN ARROW STARTING WITH HORIZONTAL LINE = continued fall in tone (see fig. 2001a-347)
↘̣̣̣	U+AB98	MODIFIER LETTER SLANTED LINE WITH NORTH EAST ARROW FROM THE LOWER END = sharp fall in tone with rising trend at the end (see fig. 2001a-347)

↘	U+AB99	MODIFIER LETTER SHORT SLANTED LINE WITH NORTH EAST ARROW FROM THE LOWER END = slight fall in tone with rising trend at the end (see fig. 2001a-347)
ᄁ	U+ABA0	MODIFIER LETTER SMALL HARD SIGN ≈ <sup> 044A (see fig. 2001a-342, 2001a-344)
ᄂ	U+ABA1	MODIFIER LETTER SMALL SOFT SIGN ≈ <sup> 044C (see fig. 2001a-341)

Block: Supplemental punctuation

Additions for Lithuanian dialectology

⋮	U+2E46	DOTTED VERTICAL LINE → 205E vertical four dots → 250A box drawings light quadruple dash vertical → 2999 dotted fence (see fig. 2001-348 ff.)
⋈	U+2E47	WIGGLED VERTICAL LINE → 2307 wavy line → 299A vertical zigzag line (see fig. 2001-348 ff.)

Properties:

AB80;LATIN SMALL LETTER C WITH PALATAL HOOK;Lu;0;L;;;;;N;;;;;
AB81;LATIN SMALL LETTER H WITH PALATAL HOOK;Lu;0;L;;;;;N;;;;;
AB82;LATIN SMALL LETTER M WITH INSIDE HOOK;Lu;0;L;;;;;N;;;;;
AB83;LATIN SMALL LETTER N WITH INSIDE HOOK;Lu;0;L;;;;;N;;;;;
AB8C;MODIFIER LETTER BASELINE SMALL A;Lm;0;ON;;;;;N;;;;;
AB8E;MODIFIER LETTER BASELINE SMALL I;Lm;0;ON;;;;;N;;;;;
AB8F;MODIFIER LETTER BASELINE SMALL U;Lm;0;ON;;;;;N;;;;;
AB90;MODIFIER LETTER SLANTED UP ARROW;Lm;0;ON;;;;;N;;;;;
AB91;MODIFIER LETTER SLANTED DOWN ARROW;Lm;0;ON;;;;;N;;;;;
AB92;MODIFIER LETTER SLANTED UP ARROW ENDING IN BEND;Lm;0;ON;;;;;N;;;;;
AB93;MODIFIER LETTER SLANTED DOWN ARROW STARTING WITH BEND;Lm;0;ON;;;;;N;;;;;
AB94;MODIFIER LETTER SMALL MIDDLE SLANTED UP ARROW;Lm;0;ON;;;;;N;;;;;
AB95;MODIFIER LETTER SMALL MIDDLE SLANTED DOWN ARROW;Lm;0;ON;;;;;N;;;;;
AB96;MODIFIER LETTER SLANTED UP ARROW STARTING WITH HORIZONTAL LINE
;Lm;0;ON;;;;;N;;;;;
AB97;MODIFIER LETTER SLANTED DOWN ARROW STARTING WITH HORIZONTAL LINE
;Lm;0;ON;;;;;N;;;;;
AB98;MODIFIER LETTER SLANTED LINE WITH NORTH EAST ARROW FROM THE LOWER END
;Lm;0;ON;;;;;N;;;;;
AB99;MODIFIER LETTER SHORT SLANTED LINE WITH NORTH EAST ARROW FROM THE
LOWER END;Lm;0;ON;;;;;N;;;;;END
ABA0;MODIFIER LETTER SMALL HARD SIGN;Lm;0;L;<super> 044A;;;;;N;;;;;
ABA1;MODIFIER LETTER SMALL SOFT SIGN;Lm;0;L;<super> 044C;;;;;N;;;;;
2E46;DOTTED VERTICAL LINE;Po;0;ON;;;;;N;;;;;
2E47;WIGGLED VERTICAL LINE;Po;0;ON;;;;;N;;;;;

Linebreaking properties of punctuation marks:

The proposed punctuation marks are vertical lines by nature, and therefore behave in any linebreaking and related processing like U+2016 DOUBLE VERTICAL LINE.

Notes on confusables:

The letter U+AB82 LATIN SMALL LETTER M WITH INSIDE HOOK is similar to U+006D LATIN SMALL LETTER M.

Likewise, the letter U+AB83 LATIN SMALL LETTER N WITH INSIDE HOOK is similar to U+006E LATIN SMALL LETTER N.

The intonation marks, while being modifier letters, are not intended to be allowed in IDNs or identifiers.

4. Acknowledgements

Special thanks for Michael Everson for providing a font containing all the characters proposed in this document.

5. References

- [1930a] Gerullis, Georg: Litauische Dialektstudien. Leipzig 1930
- [2001a] Girdenis, Aleksas: Kalbotyros darbai (Studies in linguistics): straipsniai, studijos, esė, recenzijos. T. 3: 1988–2000. Vilnius, Mokslo ir enciklopedijų leidybos inst., 2001. ISBN 5-420-01480-7
- [2006a] Kačiušienė G. Šiaurės panevėžiškių tarmės fonologijos bruožai. Vilniaus universiteto leidykla, 2006.
- [2008a] Girdenis A. Žemaičių dzūkai: Tekstai su komentarais. Mokslo ir enciklopedijų leidybos institutas, 2008

6. Examples and Figures

The figures are numbered by the referenced work (consisting of the year of edition and the letter, as in the "references" list, followed by a hyphen the page number, and following by a second letter if more than one figure is taken from a page.
E.g.: "Fig. 2001a-344" means "See ref. [2001a], p.344").

Fig. 1930a-XV: Showing specimens for U+AB8F.

[u̯] = Petitbuchstaben neben Normaltypen weisen auf wesentlich schwächere Artikulation hin: Ž R nu̯r̯ < nóriu .

Fig. 1930a-XVIII: Showing specimens for U+AB80.

[g] = stark stimmhaftes, unaspiriertes g, gemeinlit.: schriftlit. gá·udo· = gáudo .

[ç] = palatales, stimmhaftes g, gemeinlit.: çí·βus = gývus .

Fig. 1930a-XIXa: Showing specimens for U+AB83.

[n̥] = dentales, stark stimmhaftes n, gemeinlit.: schriftlit. nu̯ .

[ŋ] = palatales n, gemeinlit.: schriftlit. sũ̯·ķei = suñkiai .

[ɲ] = velares n, gemeinlit.: schriftlit. mẽ̯·ko· = meñko .

[ɳ] = silbisches n, sehr selten: R 5 łukúos̃ < laukúosna .

Fig. 1930a-XIXb: Enlarged part of 1930a-XIXa, to show the exact form of U+AB83.

[ɳ]

Fig. 1930a-6: Showing specimens for U+AB80 (red).

<i>Zuikis ir ežys.</i>	<i>zuiķī·s ir. ežī·s</i>
<i>Dailūs, gražūs zuikēlis sāko</i>	<i>daitūs gražūs zuiķē·lis sā·ko·</i>
<i>ēžiui:</i>	<i>ē·žiui</i>
<i>— Kokiē, broliūkai, tāvo</i>	<i>kó·ķi bro·lū·kai tā·βo· rū·bai</i>
<i>rūbai dýgūs ir negrāžūs.</i>	<i>dī·gu·s ir. negrā·žu·s</i>
<i>— Tiesà, — atsākē ežys, —</i>	<i>tiesà atsā·ķē· ežī·s bēt mào·</i>
<i>bēt mào ādatos *āpgina manè</i>	<i>ā·dato·s ā·pġina ma·nē nuō šu·</i>
<i>nuo šuniēs ir vilko dantū. Aŗ</i>	<i>niēs ir. βī·ko· dantū·</i>
<i>tám tāvo gražūs kailēlis? Zui-</i>	<i>aŗ. tá·m tā·βo· gražūs kai·lē·lis</i>
<i>kēlis niēko nebeatsākē, tik suñ-</i>	<i>zuiķē·lis niēko· nebeatsā·ķē·</i>
<i>kiai atsidūso.</i>	<i>tik· suñ·ķei atsidūso·</i>

Fig. 1930a-13: Showing specimens for U+ABAE (red) and U+ABAF (green).

schwächung des Drucks tritt in Diphthongen, wie überall beim litauischen Brechton, über dem ersten Komponenten ein, nur bei *ie*, *uo* über dem zweiten.

3. Der Dehnton hat seine Druckstelle, da wir eine rein žemaitische Mundart vor uns haben, über dem ersten Komponenten des Diphthongs, nur bei *ie*, *uo* über dem zweiten.

Fig. 1930a-26a: Compare with fig. 2001a-349 which shows the same text.

tuŗ·gui meŗġā·ta atsīlī·ka nù mó·tinas
ŗi (auch ŗiñ) tā·ksta ŗēskó·dama sàβa mamī·ies
žmó·nis klá·us
kó·ķē bú·βa tàβa mā·ma
meŗġi·te veŗġdam,à sū·ka
aŗ ŗū·s ňežī·noĩ mào·na mā·ma aŗdri·sk"s tarbā·lòm apšikabī·-
nus àle ūž ňisas ġeŗā·šne

Fig. 1930a-26b: Showing a specimen for U+AB83.

kuŗ leišī·be
tupė·dams pūkštā·lis mā·di pamā·te aņt žā·mes grudā·li ir
mīšlīje sá·ų
iñ·sū ir sulėšū
atsī·tupė pūkštā·lis pŗi grú·da ir paklū·βa i. spá·stus
kuód,èl àš tūr,ù (auch tūrù) pražú·t gróū·džei sućipšė·je
pūkštā·lis vānag,ai gá·uda ġi·βus pūkštėlūs ir tai ŗiem ňiėkai
ò. àš tūrù dèl viena meñ·ka grudā·le pražú·t

Fig. 1930a-87: Showing a specimen for U+AB83 (red).

Czà roūswoš, nášwoš, pítkos čà rōū·sβā.s meš·sβā.s pī.tkā.s
ūmedes sutūpy, ú·me.đe.s sutū·pī.
Linksmūtes, gražiai áuga, niē- lipksmū.íte.s gražėi. á·uga niē-
kas jóm narūpi. kas jō·m nerū·pī

Fig. 1930a-89: Showing a specimen for U+AB8C.

būwīs (-ys) gałañ, labai nusi-
minūs (nusimine), matýdamas,
kád wisū tuřtu išznykta (wisì
tuřtai išznyko), ė (aė) tiktai

Fig. 2001a-341: Showing specimens for U+ABA1.

ḃ = labai redukuotas bespalvis priešakinis balsis; vartotinas tik tada, kai tikslesnės balsio spalvos neįmanoma nustatyti. Pvz.: š. pan. àk^(ḃ)š ~ akis, žà^(ḃ)š ~ žaliūs, -iàs;

Fig. 2001a-342: Showing specimens for U+ABA0 (green).

Ḅ = labai redukuotas bespalvis užpakalinis balsis; vartotinas tik tada, kai tikslesnės balsio spalvos neįmanoma nustatyti. Pvz.: š. pan. ràg^(Ḅ)š ~ ragūs, vzi.k^(Ḅ)š ~ vaikas, vzi.k^(Ḅ)š ~ vaikūs.

Fig. 2001a-343: Showing specimens for U+ABA0 (red) and ABA1 (green).

§ 6. Redukcija

Maža kiekybinė balsio redukcija žymima pakeltu mažesnio šrifto rašmeniu: ^a, ⁱ, ^u. Pvz.: š. žem. šâ·rk^as ~ šárkos, á·vⁱs ~ ávys, vî·r^u ~ výrų. Kai redukcija itin ryški (balsis nesudaro skiemens ir pan.) balsio rašmuo dar ir suskliaučiamas: ^(a), ⁽ⁱ⁾, ^(u). Pvz.: š. žem. an^a·pá·meř^(e) vā·k^(a) on·tá·k^(a) ~ ana („ji“) pāmetė vaiką an(t) tāko.

Redukcija, iš esmės keičianti balsio kokybę, žymima minėtais specialiais rašmenimis: α ḃ Ḅ ḅ Ḇ. Reikalui esant, jie gali būti pakeliami bei sumažinami (^(α), ^(ḃ), ^(Ḅ)) ir suskliaučiami (^(α), ^(ḃ), ^(Ḅ), ^(ḅ), ^(Ḇ)). Pvz.: š. pan. àk^(ḃ)š ~ akis, ràg^(ḃ)š ~ ragūs, vðšt^(ḅ) ~ vištà.

Fig. 2001a-344a: Showing specimens for U+AB80 (red), U+AB81 (green), AB82 (blue), AB83 (purple), ABA0 (orange).

§ 7. Priebalsių ženklai

p b β φ m w – abilūpiei. Pvz., bk βó·ras || vó·ras ~ vóras, φù·rije || furi·je ~ fùrija. Tariant β tarp lūpų būna plokščias plyšys, tariant w – apskritas.

f v m p b – lūpų dantiniai. Pvz.: bk ñimfa ~ ñimfa, apvadaĩ. ~ apvadaĩ, p̄fēnigas ~ pfēnigas. Ženklu *v* dabar įprasta žymėti atitinkamą bk ir tarmių garsą, nors dažniausiai jis yra abilūpis, – vadinasi, turėtų būti rašomas β arba w.

j j t d k ğ l r ŋ η x h – palataliniai. Pvz.: bk k̄eŋk̄é·jes ~ kenkėjas, jūŋge· ~ br̄eŋx̄itas, kretingiškių j̄a.ūt̄^u ~ jáučių¹². Pokario metais vietoj *k ğ η* ir kt. įprasta rašyti *k̄ ğ̄ ŋ̄*, beveik nebevartojami *j t d*.

ž š č ž – retrofleksiniai („šlekiuojamieji“); tariami kiek atlenktu atgal liežuvio galu. Pvz., senųjų žagariečių š̄eš̄ū· žō·žū. ~ šešių žōdžių, š̄enūks ~ senūkas (girdėti lyg tarpiniai garsai tarp *š / s, ž / z, ž̄ / ž̄*)¹³.

l m n r – skiemeniniai. Pavartojami tik fakultatyviai, – pvz., š. pan. àkl̄^s || àk̄at̄^s (kai kur àk̄at̄^s) ~ aklūs.

t d s z c ʒ n l – liežuvio priešakiniai dantiniai. Pvz.: bk žū·kas ~ dzūkas, ž̄ingulū·kai ~ dzinguliukai. Vietoj *ʒ ž̄* dažniausiai rašomi digrafai *dz d̄z*, nors jie ir nesiderina su *c č̄*.

t d š ž č ž̄ l r ŋ – liežuvio priešakiniai alveoliniai. Pvz.: tr̄ak̄t̄i ~ trakti, d̄r̄i·pso· ~ dr̄ybo, ž̄a·ŋras ~ žánras, ž̄a·r̄ž̄gina ~ džáržgina, ž̄ú·ḡauje ~ džiūgauja. Vietoj *ž̄ ž̄̄* labai dažnai rašomi digrafai *dž d̄ž̄*, nesideriną su *č č̄*.

ʒ z c r – retrofleksiniai. Pasitaiko tik kaip variantai ir todėl žymėtini tiksliai epizodiškai, – pvz.: bk sr̄uto·s ~ sr̄utos, p̄ū·zras ~ p̄ūzras.

ŋ g k γ (h) x – liežuvio užpakaliniai. Pvz., bk tángas ~ langas. Vietoj *γ* dabar įprasta rašyti *h*, nors iš tikrųjų tai ne liežuvio užpakalinio, o gérklinio priebalsio ženklas. Įsidėmėtina, kad *x* nieku būdu negalima keisti digrafu *ch*, nors kartais taip ir daroma.

Fig. 2001a-344b: Showing enlarged parts of fig. 2001a-344a, to show the exact shape of (the italic forms of) the letters U+AB82 and U+AB83.

m n

Fig. 2001a-347: Showing specimens for the intonation marks (U+AB90 ... U+AB99). In U+AB95 (red), two applications are unified into one character. U+A71D (green) is already encoded as Africanist tone mark.

§ 11. Intonacijos ir teksto sintagminės skaidos ženklai

Žymėtini tik tie svarbiausi sakinio intonacijos bruožai, kurie sutelkti centriniuose sintagmos ir frazės žodžiuose. Ženkliai (jeigu specialiai nenurodyta) kitaip, žymimi po tų žodžių.

$l \setminus \setminus$ – ryškus tono kilimas (antikadencija), kritimas (kadencija), kilimas–kritimas (akcentuota kadencija). Pvz.: bk *pàpu·îę· žvarbùš vė·jesl | iř·ę·mė· krīšti bá·tto·(s) snā·çę·s\ ||* ~ Pàpūtė žvarbùs véjas, ir ėmė krīsti báltos snaigės, *na·ir·grāžùmāz·\ çi tu·mērçinu·!* ~ Na, ir gražumas gi tų merginų!

$r \setminus$ – nežymus tono kilimas ar kritimas. Pvz.: š. žem. *tas·õ·nī^(e) | mō.n^a brūoīsl | dēdēlē· mī.līē^(e) tōn·mē·rçę·tk^(a) | uuōñē.lē\ ||* ~ Tas Añtė („Antanas“), mūno („mano“) brólis, dīdeliai („labai“) mylėjo tą mergėlką Onėlę. Dažniausiai šie ženklai rodo ir greitesnį tarimo tempą.

$J \setminus$ – tęsiamas tono kilimas ar kritimas. Pvz.: š. žem. *īęi·vā·kā· krūokl | çī·pJ | mūoī.n^(a) eš·prūot^(a) | krā·u.stūos\ ||* ~ Tie vaikai kriōka („verkia“) cypa („cypia“), mótyna iš prōto kráustos!

l° – tono kilimas su kritimo tendencija intonuojamo segmento pabaigoje. Pvz.: š. žem. *pā.s(k^(a)) | kad·ųns·mē.r^(e) | ūk,tā·příl·anūo | tarā.d^(a) ||* ~ Pāsku („paskui“), kad anas mīrė, tiktaī pri anō („pas jį“) terādo. (Baigiant sakyti *mīrė* susvyruota, ar aiškinti toliau situaciją, ar ne.)

$l \setminus \setminus$ – ryškus ir atitinkamai nežymus tono kritimas su kilimo tendencija segmento pabaigoje. Pvz.: bk *tu·vīska·l jāu·pābziçei?* ~ Tu vīską jau pābaigėi? (Ramus klausimas; emociškai klausiant ar perklausiant būtų *tu·vīska·l jāu·pābziçei?*).

$! \setminus$ – staigus bendras sintagmos tono pakilimas (žymimas frazės ar sintagmos pradžioje). Pvz.: *kōnēgā·kšīst^(a) | tas·nālē.p^(a) | ō·n·tuorūos! ||* *ōns·vāžā.ū^(a) | e·nuvāžā.ū^(a) ||* ~ Kunigáikštis – tas nelipo an(t) tvorōs! Anas važiāvo i(r) nuvažiāvo. (Pakeltu tonu pabrėžiamas temos vardininkas.)

\setminus – bendras tono kritimas (žymimas frazės ar sintagmos pradžioje). Pvz.: *sāvo·mīnī·sēl | mano·núomo·nē·s^(a) | nīēkās nēktā.ušē·\ | vīšīškzi·su ūkāū. | su·bró·lēs\ ||* ~ Sāvo mintysė (mano núomonės niēkas neklāusė) vīsiškai sutikaū su bróliais.

Fig. 2001a-348: *Showing specimens for U+2E46 (red) and U+2E47 (green).*

// – frazės pabaigos (terminalinė) pauzė (žr. ankstesnius pavyzdžius).

/ – sintagmos ribų pauzė (žr. ankstesnius pavyzdžius).

– – ilga neterminalinė pauzė. Pvz.: bk *ir_tėvas / ir_motina / ir_broliai* – *visi kaip žvėrys puola tą bernioką*.

⋮ – neryški pauzė sintagmos viduje (žr. ankstesnius pavyzdžius).

⋈ – netikėtas kalbėjimo srauto pertrūkis arba staigus intonacijos posūkis. Pvz.: š. žem. *er_õns_lõb_ s_õ.mõis / p_õ_põnkis / ⋈ k_õ.i(k)_ g_õ.us_ t_õn_d_õ.in^(a) ⋮ da.r_õb_õn_õnku* // ~ Ir anas liúoba samdýs („jis samdydavo“) po penkìs... kiek g_õus tą dieną darbinink_ (pauzė prieš darbinink_ – pridūrimo signalas).

- – atviroji sandūra. Ypač žymėtina balsių junginiuose, kurie be šio ženklo gali būti skaitomi kaip dvibalsiai. Pvz.: bk *pa-upis* ~ paupys, *su-irute* ~ suirute.

-- – netikėtas kalbėjimo srauto pertrūkis žodžio viduryje. Pvz.: *ute-niški_ spj_õ.udi.dava.(s) / žma.g_õ.lis / kai_r_õik_õ.dava. n^u. k_õbl_õ.ka. / n_õ-⋮ nuk_õb_õ.n(õ) t_õs_pu.g_õl_õ_k_õs* // ~ Spj_õudydavo žmog_õlys, kai reik_õdavo nuo k_õbl_õko n... nuk_õb_õnt' tas p_õg_õli_õkas¹⁷.

Fig. 2001a-349: Compare with fig. 1930a-26a which shows the same text.

§ 12. Keletas pavyzdžių¹⁸

Bendrinė kalba

tuř.guįę mēřęĉē.Ĺē.¹⁹ | aĉiliko nuo_mó_řino.s\ || ji_lā.ksto. | ĵieškó.da-
ma sávo. mamí.ře.s\ || žmó.ņe.s\ | ktá.ušę\ || ko.kę. būwo. tāwo. mamā? ||
mēřęĉē.Ĺē. | vęrgdamā sā.ko.\ || ařęġ jū.s | ñežino.ře.? || máno. mamā. ab-
dřřřskųšġ | řeřbē.ře. mis apřřřkabiņušġ | řet_už_vřsāž. | řeřēšņę.\ ||

Pietų žemaičiai raseiniškiai

tū.ř.guį. mēřęĉē.Ĺē. | aĉilika nu_múoři.nas\ || ji_lā.ksta. i:řkúodama | sáva
mamí.ře.s\ || žmúořis. | ktá.us\ || kuokę. | būwa. tāwa má.ma? || mēřęĉē.Ĺē.
řięgdama | sá.kal\ || ař. jū.s | ñežl.núot.? || mųna má.m. abdřřřskųšġ |
řeřbġkųms apřřřkabiņušġ | | aĹe_už_vřsāž. | řeřēšņġ.\ ||

Šiaurės žemaičiai telšiškiai

tō.ř.gō.u. mēřęĉē.Ĺē. | aĉēĹē.k^(a) nu_múořin^a.s\ || an^(a)_lā.kst^(a) i:řkúodam^a |
sá.u^(a) mamā.ř.s\ || žmúořis. | ktá.us\ || kúok. | bū.u^(a). tā.u^(a) má.mā? ||
mēřęĉē.Ĺē. řięgdama | sá.k^(a)\ || ař. jū.s | ñežēn.úot.? || mō.n^a má.m^(a).
abdrē.skųš^(a) | řeřbēĹēms apřřřkabē.nųš^(a) | řet_už_vřsāž. | řeřēšņġ.\ ||

Rytų aukštaičiai uteniškiai

tuř.guį mēřęĉē.ta. | ařlġ.ka. na_mō.čę.s\ || an_řá.ksta. | ĵieřka.damā. |
sá.řa. ma.tú.ře.s\ || žmó.ņe.s\ | ktá.ušę\ || ko.kę. bū.wa. tā.řa. mō.mā? ||
mēřęĉē.ře. | vęrgdamā. | sá.kal\ || ař. jū.s | ñežl.ná.t.? || má.na. mō.mā. |
abdřřřskųšġ | řeřbē.ta.m. apřřřkabi.nųšġ | aĹe_už_vřsāž. řeřá.sna.\ ||

Rytų aukštaičiai kupiškenai

tuř.guį mēřęĉē.ta. | ařlġ.ko. nō_mō.čę.s\ || ĵi_řá.ksta. | ĵieřko.damā. | sá.řa.
mō.tú.ta.s\ || žmó.nā.s\ | ktá.ušę\ || ko.kę. bū.wa. tā.řa. mō.mā? || márgó.itā. |
vārgdamā. | sá.kal\ || ař. jū.s | ñežl.nó.t.? || mō.nō. mō.mā. | abdřřřskųšġ |
tarbā.ta.m. apřřřkabi.nųšġ | aĹe_už_vřsāž. řeřá.sna.\ ||

¹⁸ Pasinaudota J. Gerulio (Gerullis, 1930, 11 tt.) antruoju lyginamuoju tekstu „Mama“. Širdingai dėkojame Daivai Atkočaitytei, Irenai Krupienei, Reginei Rinkauskienei, kurios padėjo mums patikslinti kai kurių tekstų transkripciją.

from p. 350:

Rytų aukštaičiai šiaurės panevėžiškiai

toř.g^(b). męrgá.ře. | aĉēĹē.ka nō_mó_řnoř.s\ || ĵōn_řá.ksta. řeškó.dam^(b) |
sá.va mořú.ře.s\ || žmó.ņe.s\ | ktá.us^(b)\ || kōk^b. bū.va. tā.va má.m^(b)? || mēřęĉē.
vęř.gdam | sá.kal\ || ař.g^(b) jū.s | ñežl.nat.? || má.na má.m^(b). | abdřřř.sk^b |
řeřbē.ře.m. apřřřkabi.n^b | řet_ōž_vřs^(b). | řeřēšņę.\ ||

Fig. 2006a-162: Showing specimens for U+AB83 and several other characters proposed here.

sałdá.inęš\ jo.š is'si.perk\ || o_viręj | vir'si_tuos | so_ęžgeržm\ ||
nò|_te_wòt paseĩ.ñ.k | ta_mó.ĩñ; dāuģá.us^(e) || paseĩ.ñ.k | mó.ĩñ; tē.v's |
so_dúon; | su_vĩ.na. | štiklųkēs || pastā.te; § e_kàs; netùr^(e) | te_vĩ.na. | àtōs\ pastā-
ta. | đvė; šteklėn^(e) § te_júos; | paseĩ.ñ.k\ || dúonō.s; | tuo(s)_su_dròsk | ušpĩ.tĩ
uñĩ_lė.kštėš; | toš_dúonō(s) | suwá.tga. | túos; | ĩžger^(e) | štiklòkš; nō_e_vl.s¹⁹ ||
dağžut|jiem | jō_nereĩ.gdawa. | jāuniesėm\ || te_vl.s | tòkš; jō | palė.ñ.gve.ñm's |
bú.daua.\ || iš (i)ú | nereikatá.uj^(e) || tek_ĩš_svõ_čos; | iš_svõ.tu. || jō; | išsipĩ-
kl.ma._ta. || pas_mòņ; | mú.su. | krāšĩ ||
dabarĩ | te_má.s; | gėžžæm\ || ka_dabar; | kad_væstũ.vėš | bāĩ.gėš | đvĩ.læktō
wā.tandō²⁰ || te_kò | te_tep_greĩt; | iššėškòrs\ || vakàr; | ka_parvažiuoĩ^(e) | bú-
dawa. | nætrō.ñ.k\ | ta_đvĩ.lækt^(a); | atėĩ.n²⁰ ||
[iš ryto] nō_žā.đñ | te_muzikòntė\ dāuģá.usę || te_tiep | má.ršš; | griež^(e) |

Fig. 2008a-271: Showing specimens for U+AB83 and several other characters proposed here.

mę(s) | žid⁽ⁱ⁾ė.ĩ\ | tokeĩ_raudúonâ.(s); žę.idėlė.s\ || pri_pàt |
žė.m^es\ || ta_tėi_dagòkâ | bō.u | dėdėl^(e); | vā.istā.\ || ta_lò.b; | tũ.n |
idies | kàt-šũ.d^(e) | idies | e_đous-ə; | gė.r^e\ || rūožė.s | ka_bũ.s_kúok^es |
kžn\ || árbà; | túokĩ.ĩ | dā.r; | tũ.n_túokũ.ñ | kiesalòku³¹²\ | bú.s-ə\ |
tuokę(s)_žuolė.l's; | šàknikėš\ dous || suriου^e || ta_túokæ.s\ lò.b |
ę_gĩ.đĩ.s^e(s)\ || vō_tė.p | nabō.u\ túokũ.m ||
xž.³¹³\ | žėnâ.v\ || toúkâ.r | nē; | aš_á.tmėn^(o) | ne_niek^(a)\ || tōu-
kâ.r; | žėnâ | kėtúokš | bōu_gi.vėnėms § dabã | gi.vėnėms ĩ.r | dėđlė.\
gėrs-ə || vō_tòukâ.r | bō.u | sũ.ñ.kūs\ || jē; | ka(s)_sė.rγə³¹⁴; sė.rgə | jē; |
nē | nē || juk_tòukâ.r | sākâu\ aš || ka_ĩrĩ.lėk^(a); | áuγúo(m)³¹⁵ |
mė.s³¹⁶ | e_šėšėj_tiktã; | tabò-om\ || aštó.un^(e) | tapaúg^uòm\ || jāut |

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: *Proposal to add characters used in Lithuanian dialectology to the UCS*

2. Requester's name: *Vladas Tumasonis; Karl Pentzlin*

3. Requester type (Member body/Liaison/Individual contribution): *Expert Contribution*

4. Submission date: *2010-09-24*

5. Requester's reference (if applicable): *Vilnius University; Faculty of Mathematics and Informatics*

6. Choose one of the following:

This is a complete proposal: *Yes*

(or) More information will be provided later:

B. Technical – General

1. Choose one of the following:

a. This proposal is for a new script (set of characters): *No*

Proposed name of script: _____

b. The proposal is for addition of character(s) to an existing block: *Partially*

Name of the existing block: *Latin Extended-E (new), Phonetic Ext. Supp. B (new), Supp. Punctuation*

2. Number of characters in proposal: *21*

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection)

C-Major extinct D-Attested extinct E-Minor extinct

F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided? *Yes*

a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*

b. Are the character shapes attached in a legible form suitable for review? *Yes*

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? _____

Michael Everson

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): _____

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*

7. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *No*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3702-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain		No
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents:	<i>One of the authors (V. T.) does itself belong to the scientific community</i> <i>See text</i>	Yes
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>See text</i>	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>See text</i>	Scientific
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>See text</i>	Yes
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	<i>To keep in line with related characters</i>	Yes Yes
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		Yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:		No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:		No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:		No
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:		No
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)		No
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:		No