

Additional proposal to encode Latin characters for theta and delta

Denis Moyogo Jacquerye <moyogo@gmail.com>

2014-08-07

1. Introduction

Some orthographies (Cypriot Arabic, Rromani, Tuscarora, Wakhi) using the Latin script have the delta or theta letters. This document supports the proposal to disunify the Greek theta characters made in *Proposal to encode “Unifon” and other characters in the UCS* by Michael Everson (L2/12-138) which showed theta is used in the Tuscarora language orthography and some Latin based phonetic transcription systems. The addition of a capital delta as an uppercase of the existing U+1E9F LATIN SMALL LETTER DELTA is also being proposed.

2. Background

The Rromani common alphabet was developed to be a cross-dialect standard orthography by Marcel Courthiade in 1989. [1] The alphabet was accepted by the International Romani Union at the 1990 World Romani Congress at Warsaw. The original document is replicated and translated in Hungarian and English in M. Courthiade and M. [2] The alphabet is in used in an education programme supported by the Ministry of Education and Unicef in Romania. [3] It can be seen in school primers. [4] [5]

In the Rromani alphabet, the Latin theta is used to represent a positional letter which is pronounced [d] after n and [t] in other positions. [2]

The Cypriot Maronite Arabic alphabet also uses a Latin theta along with a Latin delta. Both are in use in some newspapers. [8] [9]

The alphabet was presented by Alexander Borg in 2007 and a Committee of expert for the codification of CMA adopted the alphabet in 2009. The alphabet doesn't have an official status yet. [7]

Wakhi language orthographies using the Latin alphabet also uses the Latin theta and the Latin delta. [10] [11] [12]

3. Rationale

The disunification of the Latin letter theta is proposed in the Everson 2012 (L2/12-138). [13] There is evidence of a wider use of the Latin letter theta (Cypriot Arabic, Rromani —used by more than 1 million speakers in Romania where the Rromani common alphabet is being taught—, Tuscarora, Wakhi).

There is evidence of use of the Latin delta in some orthographies. The lowercase form has already been encoded as a phonetic symbol for medieval Welsh transcription, proposed in Everson *et al.* (L2/06-027) [14], and could be unified with the orthographic Latin delta. The uppercase used in orthographies would then need to be encoded.

4. Proposed characters

Characters proposed

| Glyph | Code | Character name |
|--------------|-------------|----------------------------|
| Δ | X+A7B8 | LATIN CAPITAL LETTER DELTA |
| Θ | X+A7B8 | LATIN CAPITAL LETTER THETA |
| θ | X+A7B9 | LATIN SMALL LETTER THETA |

5. Examples

5.1. Rromani examples

THE 3 POSTPOSITIONNAL LETTERS

| | Position in word | Karp. Fen. Gurv. Cerh. | Colari | Lovari | outside Hungary |
|---|--------------------|------------------------|------------------|--------|-----------------|
| q | after n | g | g | g | g g' d' dʒ d |
| | any other position | k | k (-sq- => -kk-) | k | k kk k' t' tʃ Ø |
| θ | after n | d | d | d | d Ø etc... |
| | any other position | t | t (-sθ- => -tt-) | t | t tt t' |
| ϕ | after n | ts | ts | ts | c dʒ |
| | any other position | h | h | s | s h j v Ø |

Figure 1: Example from Courthiade and Rezműes 2009, inside cover, showing
LATIN SMALL LETTER THETA.


| | | |
|--|---|------------------------------------|
| <p>i korri jakh MA halánték; EN temple; FR tempe; ES sien; DE Schläfe; UKR скроня; RO tâmplă; HR sljepoočnica; SLK spánky; GK κρόταφος.</p> <p>lel/marel jakhaθe/jakhalo MA szemmel ver; EN to cast the evil eye, to put a jinx; FR jeter le mauvais œil; ES aojar, echar mal de ocho; DE mit dem bösen Blick treffen; UKR дурно диавитися; RO a lua în ochi de rău, a deochea; HR štetiti urokljivim okom; SLK uhranút'; GK ματιάζει.</p> <p>bi jakhenqo MA szemtelen; EN impertinent, arrogant;</p> | <p>κλιτ', začarovat'; GK γοητεύει, μαγεύει, κάνει μάγια.</p> <p>i/e jakhvalin, -ă- ʒ, -ă, -ěn- MA szemüveg; EN glasses; FR lunettes; ES gafas; DE Brille; UKR окуляри; RO ochelari; HR naočale; SLK okuliare; GK γυαλιά.</p> | |
| |  | <p>o/e jakhvalinã khamesqe</p> |

Figure 2: Example from Courthiade and Rezműes 2009, page 180, showing showing
LATIN SMALL LETTER THETA.

O ROMANO HÏMNO

GELEM, GELEM

- Gelem, gelem lungone dromença,
Maladilem baxtale rromença,
A, rromalen, kaθar tume aven,
E caxrença, bokhale çhavença?
R.: *Haj, haj, rroma!len*
Haj, haj, çhava!len... (duj var)

Figure 3: Example from Borcoi et al. 2009, page 4 showing LATIN SMALL LETTER THETA.

NEVE SIKLÖVIMATA

1. I ANALIZA JEKHE TEKSTOSQI ANΘAR I RIG E LINGVISTIKAQI

(adaptàcia palal o Marcel Courthiade)

O Ali Krasnić lekhavel jekh but gugli rromani čhib, anθar e lèkta e Kosovaqe Gurbetenqi, savi perel an-e “stràta 2” e rromane dialektenqi (e “stràta 2” si kaja kaj phenel “sem”, “paj”, “lùmia” thaj kola, zikaj e “stràta 1” phenel “sinom, hium, som”, “pani”, “sundal, svèto” thaj kola. an-e “stràta 1” peren e lèkte e Arlienqe ta an-e “stràta 3” e lèkte e Kelderarenqe ta e Lovarenqe; e “stràta 3” xasarda e duj fonème 3 thaj čh).

O Komisariàto vaš-e Čhib ta e Čhibutne Xakaja (Drezipen e Maškarthemutne Rromane Uniaqo – Praha) dias vast te lačhardöl e palutni vèrsia e tekstosqi. E ediciaqi grùpa nakhada e paramiča p-e sa-themenqi alfabèta e rromane čhibaqi, so sas oficialosardi an-o b. 1990-to an-e Varšavaqo Kongrèso thaj fonkcionisarel pal-o princìpo so savorre lekhavas sa e rromane lèkte palal jekh grafiko sistèma (i alfabèta khethani sa e Rromenqe), aj das patìv sa e dialektalone variantenqe, godlesqe so si vi von jekh barvalipen anθ-e čhibaqo mangin.

Figure 4: Example from Sarău et al. 2009, page 12 showing LATIN CAPITAL LETTER THETA and LATIN SMALL LETTER THETA.

5.2. Cypriot Arabic examples

In view of its relatively small speaker community it has become clear that the survival of CMA cannot be guaranteed via generational transmission as in the past, particularly, since all CMA speakers are natively fluent in Greek and are liable to shift to this majority language within the next few decades. The CMA community has therefore decided to cultivate its language as a written medium and the following alphabet has been proposed for this purpose:

A, a – B, b – C, c – Ć, ć – D, d – Δ, δ – E, e – F, f – G, g – Ğ [or G, g] – X, x – I, i – J, j – K, k – L, l – M, m – N, n – O, o – P, p – R, r – S, s – Š, š – T, t – Θ, θ – U, u – V, v – W, w – Y, y – Z, z.

Figure 5: Example from Borg 2008, page 1, showing LATIN CAPITAL LETTER DELTA, LATIN CAPITAL LETTER THETA and LATIN SMALL LETTER THETA.

Θenixxar pale mannixunu os te
šil afro. L- afro maxxavvisu ma
exte mišle

Maxxavvisu os tentamma. Ta ma
tte kullu afro, l- casel o mintamm.

Mša ttanacaref kif l' casel
o mintamm, maxxott exte malaka
xost u mantira.
Mmalaka e sauta?
L- casel o mintamm.

Amma ma pikkistvi kaes l- casel, pisur
xameδ

Την άλλη μέρα πάλι το βράζουμε μέχρι να
βγάλει αφρό. Το αφρό τον μαζεύουμε με
μια τρυπητή

Το μαζεύουμε μέχρι να τελειώσει. Όταν δε
θα έχει πια αφρό το μέλι είναι τελειωμένο

Για να ξέρούμε ότι το μέλι
είναι τελειωμένο, βάζουμε ένα κουτάλι
μέσα και το γυρίζουμε
Το κουτάλι είναι μαύρο?
Το μέλι είναι τελειωμένο

Όταν δε ψηθεί καλά, το μέλι γίνεται
ξινό(ξινίζει)

Figure 6: Example from Κοινοτικό Βήμα, December 2012, page 8, showing LATIN CAPITAL LETTER THETA.

Δδεια farxet tel Antoni P Skullo

Mi ta mannaklep Mmurto taxt, u manacanti xost l-trep te l-δeca
kiptel l-zman, kiptel l-xar, l-xava, l-sama, u mantes pittiptel u ttinye kulla.
Mannilteka kintam fi exte ikona ta ftaillak kan ispaca Rrabbi ma kullon
l-şkal xromata. Xoδer, asfar, apiaδ, axmar.

Farxu δδentra kullon, l-fkiora, l-şauk, l-lingres, tterraş, l-kantul, l-zacataru
u kan xazir te farxu u l-xjar. Pitnazel l-ćami tel aftokinito pşian te cantu
l-rixat java u pros stiymis pitir δexnak u pkyntece nofesak. Akke ta
oδiyas, kintarru kintamak allik l-şiliaes lingres, akke efkin ta innen
mnauęe u mnauna mi ttarp, u pipanullak kif kalosorizuse cintak. Pi sauk
atxis akke mix ta kan kistantru nnes te δδeca şi ypir intsan, yia rais yia
iskof, u kan itecelu kintamu mnauna xok Kkamara u kan ifarizu rixan xost
l-trup. L-xkali, mantes u mattullon peftćia. L-xarupat txasnu, l-zaitunat
tlapsu carusat u mmonopakia ntalu lazarus, paparunes u fkiora tel iMsix.
Piθkaem moxak sop l-sama, pitxott slipak u piθkul:
- Δoksazose ya Rrabbi.

Figure 7: Example from *O Týpos των Μαρωνιτών*, December 2012, page 8, showing
LATIN CAPITAL LETTER DELTA and LATIN SMALL LETTER THETA.

5.3. Wakhi examples

| Special Symbols/Alphabet Keys | | | | | | | |
|-------------------------------|-----------|-------|------|---------|-----------|-------|-----------|
| Capital | Keys | Small | Keys | Capital | Keys | Small | Keys |
| Ĉ | Shift + 1 | ĉ | 1 | Š | 0 | š | 9 |
| Ĉ | Shift + 2 | ċ | 2 | Ŧ | Shift + [| ŧ | [|
| Đ | Shift + 3 | đ | 3 | θ | ` | θ | ~ |
| Δ | Shift + 4 | δ | 4 | Ĭ | Shift +] | ĭ |] |
| Γ | Shift + 5 | γ | 5 | Ž | Shift + \ | ž | \ |
| Ĵ | Shift + 6 | ĵ | 6 | Ž | Shift + > | ž | Shift + < |
| Ĵ | Shift + 7 | ĵ | 7 | 3 | Shift + - | 3 | - |
| Š | Shift + 8 | š | 8 | ł | Shift + = | ł | = |

Figure 8: Example from *Khan Xikwor Rasmulxat* 2006 LATIN CAPITAL LETTER DELTA and
LATIN SMALL LETTER THETA and LATIN SMALL LETTER THETA.

6. References

6.1. Rromani

- [1] Courthiade, Marcel. 1989. "La langue Rromani (Tsigane): Évolution, standardisation, unification, réforme". In: *Language Reform. History and Future*, Vol IV, edited by Fodor, I. & Hagège, C. Hamburg: Helmut Buske Verlag. 87-110
- [2] Courthiade, Marcel and Melinda Rezműes. 2009. *Morri angluni rromane čhibăqi evroputni lavustik*. Budapest: Cigány Ház.
- [3] Unicef Romania, *Limba și literatura rromani*, <http://www.unicef.ro/publicatii/limba-si-literatura-rromani>
- [3a] Sarău, Gheorghe. 2007. *I rromani čhib thaj i literatūra vaš o paŋžto siklövipnasqo berš (Limba și literatura rromani, pentru anul V de studiu)*. Buzău: Alpha MDN. ISBN 978-973-139-013-0 <http://www.scribd.com/doc/212312474/Manual-de-Limba-Rromani>
- [3b] Cordovan, Ionel and Noemi Palfi-Tamaș. 2008. *I rromani čhib thaj i literatūra vaš o eftato siklövipnasqo berš (Limba și literatura rromani, pentru anul VI de studiu)*. Bucharest: Vanemonde. ISBN 978-973-1733-06-7. http://www.unicef.ro/wp-content/uploads/manual_vi-a_1.pdf
- [3c] Sarău, Gheorghe, Gabriela Moise and Marius Căldăraru. 2009. *I rromani čhib thaj i literatūra vaš o eftato siklövipnasqo berš (Limba și literatura rromani, pentru anul VII de studiu)*. Bucharest: Vanemonde. ISBN 978-973-1733-07-4. http://www.unicef.ro/wp-content/uploads/manual_vii-a_2.pdf
- [3d] Borcoi, Jupiter, Nicolae Pandelică and Gheorghe Sarău. 2009. *I rromani čhib thaj i literatūra vaš o oxtoto siklövipnasqo berš (Limba și literatura rromani, pentru anul VIII de studiu)*. Bucharest: Vanemonde. ISBN 978-973-1733-09-8. http://www.unicef.ro/wp-content/uploads/manual_viii_a_1.pdf
- [3e] Sarău, Gheorghe and Ionel Cordovan. 2009. *I rromani čhib thaj i literatūra vaš o dešto siklövipnasqo berš (Limba și literatura rromani, pentru anul X de studiu)*. Bucharest: Vanemonde. ISBN 978-973-1733-23-4 http://www.unicef.ro/wp-content/uploads/manual_rromani_clsx.pdf
- [3f] Stănescu, Camelia. 2002. *Alfabetizaciaqo lil anθ-i rromani čhib vaš e čhave, terne thaj baremanuša*, Bucharest, Vanemonde.
- [4] Cordovan, Ionel and Noemi Palfi-Tamaș, *I alfabeta e rromane čhibăqi*, <http://rromani.org/>
- [5] Sarău, Gheorghe and Camelia Stănescu. 2002. *Alfabetizaciaqo lil anθ-i rromani čhib vaš e čhave, terne thaj baremanuša*. Bucharest: Vanemonde. ISBN 973-85501-5-7 http://www.isjbotosani.ro/img/uploads/petronia.scripcariu/file/Manual_de_alfabetizare_in_limba_rromani.pdf

6.2. Cypriot Arabic

- [6] Borg, Alexander. 2008. *An Alphabet for Cypriot Maronite Arabic*. http://sana.squarespace.com/storage/The_CMA_Alphabet.pdf
- [7] Council of Europe, *European Charter for Regional or Minority Languages : Third periodical presented to the Secretary General of the Council of Europe in accordance with Article 15 of the Charter*. 18 January 2011. http://www.coe.int/t/dg4/education/minlang/report/PeriodicalReports/CyprusPR3_en.pdf
- [8] Κοινοτικό Βήμα. December 2012. see page 8. <http://www.maronitesofcyprus.com/upload/20130104/1357296660-08226.pdf>
- [9] Ο Τύπος των Μαρωιτών. April 2014. see page 21 in Ash Xabar section. <http://www.typosmaroniton.com/upload/20140417/1397740841-32092.pdf>

6.3. Wakhi

- [10] *Wakhi alphabet* <https://wakhi.wordpress.com/about/>
- [11] Ager, Simon. 1998-2014. "Wakhi (xīk zik)", *Omniglot* <http://www.omniglot.com/writing/wakhi.php>
- [12] Khan, Zulfiqar Ali. 2006. *Xīkwor Rasmulxat*, http://www.gojal.net/languages_wakhi_%20alphabets.htm

6.4. Other

[13] Everson, Michael. 2012-04-29. *Proposal to encode “Unifon” and other characters in the UCS*. L2/12-138. <http://www.unicode.org/L2/L2012/12138-n4262-unifon.pdf>

[14] Everson, Michael (editor), Peter Baker, António Emiliano, Florian Grammel, Odd Einar Haugen, Diana Luft, Susana Pedro, Gerd Schumacher, Andreas Stötzner. 2006-01-30. *Proposal to add medievalist characters to the UCS*. L2/06027. <http://www.unicode.org/L2/L2006/06027-n3027-medieval.pdf>