

Greek and hyperref

Günter Milde

2020/10/30

Contents

1 Transcription: λογος, λογος	4
2 Macros: λογος, λογος, λογος, λογος,	4
3 Kerning: ΑΥΛ ΑΥΛ ΑΥΛ	4
4 Literal Unicode input	4
4.1 ' ; ' " Α·ΕΗΤΟ·Τ·Ω·ΑΒΓΔΕΖΗΘΙΚΛΜΝ	4
4.2 ΞΟΗΡΣΤΥΦΧΨΩΪΫ?ΠΓΦΛ	4
4.3 αέήίύφγεζηθιαλμν	4
4.4 ξοπρεστυφχψωΪϋόύώβθφπΩ?ΓΦΛ	4
4.5 αάδδάάάάΑ·Α·Α·Α·Α·Α·Α	5
4.6 εεεεεε·Ε·Ε·Ε·Ε·Ε	5
4.7 ήήήήήήή·Η·Η·Η·Η·Η·Η·Η	5
4.8 ιιιιιιιι·Ι·Ι·Ι·Ι·Ι·Ι	5
4.9 οοοοοο·Ο·Ο·Ο·Ο·Ο·Ο	5
4.10 ύύύύύύ·Τ·Τ·Τ·Τ·Τ	5
4.11 ώωωωω·Ω·Ω·Ω·Ω·Ω	5
4.12 δάεέήήηηηη·Δ·Δ·Δ·Δ·Δ·Δ	5
4.13 αάδδαάάάάΑ·Α·Α·Α·Α·Α·Α·Α	5
4.14 ηήηηηηηηη·Η·Η·Η·Η·Η·Η·Η·Η	5
4.15 φφφφφφφφ·Ω·Ω·Ω·Ω·Ω·Ω·Ω·Ω	5
4.16 ςςςςςςς·Α·Α·Α·Α·Α·Α·Α	5
4.17 ~~ηηηηη·Ε·Ε·Η·Η·Η	5
4.18 ιιιιιι·Ι·Ι·Ι·Ι·Ι	5
4.19 υυυυυρρυ·Τ·Τ·Τ·Τ·Τ·Ρ	5
4.20 φφφφφ·Ο·Ο·Ω·Ω·Ω	5
5 LICR command input	5
5.1 Greek and Coptic	6
5.1.1 ' ; ' " Α·ΕΗΤΟ·Τ·Ω·ΑΒΓΔΕΖΗΘΙΚΛΜΝ	6
5.1.2 ΞΟΗΡΣΤΥΦΧΨΩΪΫ	6
5.1.3 αέήίύφγεζηθιαλμν	6

8 Conclusion

On 2010-11-05, Heiko Oberdiek wrote in comp.text.tex:

`\pdfstringdef` (converting TeX code to PDF strings for bookmarks) supports NFSS2 and needs active characters. Encoding stuff based on the internal font machinery of TeX (letters with catcode 11 or 12, ligatures) does not work, because the strings don't reach TeX's stomach.

The *greek-fontenc* package allows input of Greek characters in a way that “reaches TeX’s stomach” and hence works in both, the main document as well as in PDF strings (e.g. bookmarks). Hyperref’s “puenc.def“ font encoding file defines LICR macros for monotonic Greek (Greek characters of the “Greek and Coptic” unicode block).

All utf8-encoded literal Unicode characters work in PDF strings. With *greek-fontenc* and *greek-inputenc*, this enables use of all Greek character in text and PDF strings.

With the *textalpha* package, Greek letters can be used without explicit change of the font encoding or Babel language (with some caveats, see *textalpha-doc.pdf*). For correct hyphenation and other fixes, mark Greek text parts with the Babel language *greek*. There should be no space around a language switch: `|λογος|`.

1 Transcription: λογος, λογος

Text input using the Latin transliteration: In the PDF-bookmark are Latin letters instead of Greek ones.

2 Macros: λογος, λογος, λογος, λογος,

Babel-Greek or *textalpha* package with `\textalpha` ... `\textomega` macros; *alphabeta* package with `\alpha` ... `\Omega` macros. With 8-bit TeX (pdflatex), literal Greek Unicode characters are converted to L^IC_R Macros, too.

Works, if the `unicode` or `pdfencoding=auto` option is given to `hyperref`.¹

3 Kerning: ΑΥΛ ΑΥΛ ΑΥΛ

Kerning is impeded if the font encoding is switched for every single character. To fix this, wrap the Greek part in a command switching to a font encoding supporting Greek, either `\ensuregreek{...}` (with package *textalpha* or *Babel*) or `\foreignlanguage{greek}{...}` (with *Babel*).

4 Literal Unicode input

The following subsection headings contain all characters from the “Greek and Coptic” and “Greek Extended” Unicode Blocks that are supported by the LGR font encoding.

4.1 ΚΑΙ ΤΟΥ ΑΓΓΕΛΟΥ Ο ΒΑΓΔΕΖΗΘΙΚΛΜΝ

4.2 ΕΟΠΡΣΤΥΦΧΨΩΪΫΩΓΦΛ

Greek and Coptic Unicode block: punctuation and uppercase letters

4.3 ሂኝነትና የዚህ በቃል እንደሆነ

4.4 ΞοπρεστυφχψωὕύώβυφπΩΦΤΕΛ

Greek and Coptic Unicode block: lowercase letters

¹With the “xpdf” viewer, Greek letters are not shown in PDF bookmarks.

Greek Extended Unicode block: Input as literal precomposed Unicode character works fine.

5 LICR command input

textalpha loads definitions for L^IC_R input with non-standard accents or combined diacritics characters also for PU (hyperref).

5.1 Greek and Coptic

5.1.1 Α·Ε·Η·Ι·Ο·Υ·Ω·Α·Β·Γ·Δ·Ε·Ζ·Η·Θ·Ι·Κ·Λ·Μ·Ν·

5.1.2 ΕΟΠΡΣΤΥΦΧΨΩΪΫ

5.1.3 άέηίύαβγδεζηθιαλμν

5.1.4 ΞΟΠΡΙΣΤΑΨΧΨΩΪΪÓÙÁΒΘΨΠΩΪΓΓΦΨΗΛ

5.2 Greek Extended

5.2.2 ດົກເຈົ້າທີ່ມີການສະແດງ

5.2.3 ታክክክክክክክክ ’H~H~H~H~H~H~H~H

5.2.4 I₁, I₂, I₃, I₄, I₅

5.2.5 የዕስዕስዕስ ተወስኝ ተወስኝ

5.2.6 ԵՐԵՎԱՆԻ ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅՈՒՆ

5.2.7 የሆኑዕናዕዳዕስት ትርጓሜ

5.2.8 አልተረከኝነበሮሁዎች

5.2.9 $\tilde{A}_I \tilde{A}_I \tilde{A}_I \tilde{A}_I \tilde{A}_I \tilde{A}_I$

5.2.10 កំពង់កំពង់កំពង់កំពង់កំពង់កំពង់កំពង់កំពង់

5.2.11 պահանջման գործառնություններ

5.2.12 $\bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A}$

5.2.14 $I, \bar{I}, \bar{I} \bar{I}, \bar{I} \bar{I} \bar{I}$

5.2.15 ቤትታዥ ቤትታዥ የሚያጠቃል ስም

5.2.16 ብዕራፍ የዕለታዊ ፌዴራል ስንጻዊ አገልግሎት

6 Alias character names

6.1 Θεψιβθπκρ' μΤF_Fλλ

7 Makeuppercase

According to Greek typesetting conventions, diacritics (except the dialytika) are dropped in **UPPERCASE**.

The LaTeX \Makeuppercase implementation changed fundamentally with the release in 06/2022.² Since the change, \Makeuppercase is also supported

²This change cannot be reverted with the rollback mechanism “for technical reasons”.

in PDF-strings. Greek typesetting rules are only applied if the text language is set to `greek` with Babel's `\setlanguage` or `\foreignlanguage`.

The changes broke the support for upcasing of polutonic Greek (combined accents) with LICR input and of Greek with the LGR Latin transliteration. Support is restored as of 2023-09-07, some issues remain with PDF strings.

7.1 Literal Unicode input

7.1.1 Α·ΕΝΤΟΥΡΩΙΑΒΓΔΕΖΗΘΙΚΛΜΝ

Text language English — diacritics not dropped.

7.1.2 Α·ΕΗΙΟΥΓΩΪΑΒΓΔΕΖΗΘΙΚΛΜΝ

Language set to Greek inside the \subsection command – diacritics dropped in the text and ToC but not in the PDF sidebar (sic!).

7.1.3 Α·ΕΗΙΟΥΩΪΑΒΓΔΕΖΗΘΙΚΛΜΝ

Language switched before the \subsection command – diacritics dropped in text, ToC, and PDF sidebar.

7.1.4 ΕΟΠΡΣΤΥΦΧΨΩΪΫΩΔΦΛ

7.1.5 ΑΕΗΙΫΑΒΓΔΕΖΗΘΙΚΛΜΝ

7.1.6 ΞΟΠΡΣΣΤΥΦΧΨΩΪΫΟΤΩΒΘΦΠΩΩΔΦΛ

7.1.7 AAAAAAAA AAAAAAA

7.1.8 EEEEEEEEEE

7-1-10 LLL
LLL

7.1.11 000000000000

7.1.12

7.1.14 ААЕЕННПООУУЗЗ

7115 AAAA

7.1.16 $\text{H}_2\text{H}_2\text{H}_2\text{H}_2\text{H}_2\text{H}_2\text{H}_2\text{H}_2\text{H}_2$

7.1.17 8888888888888888

7.1.18 ĂĀĂĂĂĂĂĂĂĂĂĂĂĂĂĂ

7.1.19 ~"H_iH_iH_iHH_iEEHHH_i" " ~

7.1.20 Ī Ī ī ī I ī ī Ī Ī I I ~~~

7.1.21 $\ddot{\text{Y}}\bar{\text{Y}}\dot{\text{Y}}\ddot{\text{Y}}\text{P}\text{P}\text{Y}\dot{\text{Y}}\ddot{\text{Y}}\bar{\text{Y}}\text{Y}\text{Y}\text{P}$

7.1.22 $\Omega_i\Omega_i\Omega_i\Omega\Omega_iOO\Omega\Omega\Omega_i$

7.2 LICR input

7.2.1 Α·ΕΗΙΟΥΩΪΑΒΓΔΕΖΗΘΙΚΛΜΝ

7.2.2 ΕΟΠΡΣΤΥΦΧΨΩΪΫ

7.2.3 ΑΕΗΙΫΑΒΓΔΕΖΗΘΙΚΛΜΝ

7.2.4 ΞΟΠΡΣΣΤΥΦΧΨΩΪΫΟΤΩΒΘΦΠΩΓΤΑFFFΛΛ

7.2.5 AAAA AAAAAAAA AAAAAAAA

7.2.6 EEEEEEE EEEEEEE

7.2.7 НННННННН НННННННН

7.2.8

7.2.9 0000000 0000000

7.3.10 

7.2.11 $\Omega\Omega\Omega\Omega\Omega\Omega\Omega\Omega$ $\Omega\Omega\Omega\Omega\Omega\Omega\Omega\Omega$

7.2.12 ААЕЕИИИООУУУУ

7 2 13 A A A A A A A A A A A A A A A A

7.2.14 НННННННННННННННННН

7.2.16 ÄÄA A-A AA ÄÄAAA-

7.3.10 $\ddot{\chi}\bar{\chi}\dot{\chi}\ddot{\chi}BB\chi\ddot{\chi}\bar{\chi}\chi\chi B\cdots$

7.3.30 888888888

7.3. Human

7.3.1. $\text{A}^{\ddot{\text{E}}}\text{A}^{\ddot{\text{E}}} \text{F}^{\ddot{\text{E}}}\text{A}^{\ddot{\text{E}}}\text{A}^{\ddot{\text{E}}} \text{F}^{\ddot{\text{E}}}\text{A}^{\ddot{\text{E}}}\text{A}^{\ddot{\text{E}}}$

7.3.2. $\hat{\alpha}_1 \hat{\alpha}_2 \hat{\alpha}_3 \hat{\alpha}_4 \hat{\alpha}_5 \hat{\alpha}_6 \hat{\alpha}_7 \hat{\alpha}_8 \hat{\alpha}_9 \hat{\alpha}_{10} \rightarrow \Lambda^{\hat{\alpha}_1} \Lambda^{\hat{\alpha}_2} E^{\hat{\alpha}_3} \Lambda^{\hat{\alpha}_4} \Lambda^{\hat{\alpha}_5} E^{\hat{\alpha}_6} \Lambda^{\hat{\alpha}_7}$

In RDF strings, the *hiatus* feature only works with literal input.

8 Conclusion

For Greek text parts in section headers use either literal Unicode characters³ or macros. For proper kerning und upcasing in the main document, set the text language of Greek text parts to `greek`. If you use polytonic Greek, set the `polutoniko` language attribute.

³Combining Unicode characters do not work with `inputenc` and 8-bit LaTeX. (This is a general restriction.) Use pre-composed Unicode characters or accent macros for letters with diacritics.