

The *textalpha* package

Günter Milde

2020/10/30

The *textalpha* package enables the use of Greek characters in text independent of font encoding or TeX engine.¹ Input is possible via text commands (`\textalpha` ...`\textOmega`) or Unicode literals².

Contents

1 Usage

Load this package in the preamble of your document with

```
\usepackage[<options>]{textalpha}
```

Now you are ready to use literal Unicode characters^{??} or the `\textalpha` ...`\textOmega` macros anywhere in the text.³ See the source of this document `textalpha-doc.tex` for a setup and usage example and [greek-fontenc-doc](#) for links to additional documentation.

1.1 Options

1.1.1 `normalize-symbols`

Mathematical notation uses variant shapes of some Greek letters as additional symbols. There are separate code points for the symbol variants in Unicode. TeX supports some of the variant shape symbols in mathematical mode ($\theta|\vartheta, \phi|\varphi, \pi|\varpi, \rho|\varrho, \epsilon|\varepsilon$) but not in the LGR font encoding used for Greek text in 8-bit TeX. The variations have no syntactic meaning in Greek text and text fonts may use the variant shapes in place of the “regular” ones as a stylistic choice.

The `normalize-symbols` option merges letters and symbols to Greek letters. This way, text copied from external sources can be compiled without errors even if it contains a GREEK SYMBOL ... in place of a GREEK LETTER ...

¹This document was compiled using Unicode fonts (font encoding TU). For a version using 8-bit fonts, see [textalpha-doc.pdf](#).

² Requires [greek-inputenc](#) or XeTeX/LuaTeX.

³Using the shorter `\alpha` ...`\Omega` macros (known from math mode) is possible with the [alphabet](#) package.

The source of this paragraph uses both variants for beta ($\beta|\beta$), theta ($\theta|\theta$), phi ($\phi|\phi$), pi ($\pi|\pi$), kappa ($\kappa|\kappa$), rho ($\rho|\rho$), Theta ($\Theta|\Theta$), and epsilon ($\epsilon|\epsilon$).

This option is ignored with Unicode fonts.

Attention: Do not use this option in cases where the distinction between the symbol variants may be important (e.g. in a mathematical or scientific context). Use the respective characters in mathematical mode or XeTeX/LuaTeX with Unicode fonts.

1.1.2 keep-semicolon

LGR is no **standard text font encoding**. Latin characters and some other ASCII symbols are mapped to Greek “equivalents” if LGR is the active font encoding. (See **babel-greek** for a description of this Latin-Greek transliteration.)

Special care is required with the question mark characters: The LGR font encoding uses the Latin question mark as input for the *erotimatiko* and maps the semicolon to a middle dot (*ano teleia*). As a result, Unicode-encoded texts that use the semicolon as *erotimatiko* end up with an *ano teleia* in its place! Without special care, only the deprecated character 037E GREEK QUESTION MARK⁴ works with both, Xe/LuaTeX and 8-bit TeX.

The `\textsemicolon` command inserts an *erotimatiko* in LGR and a semicolon else (i.e. always a character that looks like a semicolon):

Latin (TU) a; b, Greek (TU) a; b

With the `keep-semicolon` option, character 003B SEMICOLON can be used for the *erotimatiko* also with LGR encoded fonts:

Latin (TU)	Greek (TU)	question mark character
;	;	037E GREEK QUESTION MARK
;	;	003B SEMICOLON
?	?	003F QUESTION MARK

This option is ignored with Unicode fonts (where the SEMICOLON literal always prints a semicolon character).

Test whether this works as expected in math mode: $ab; a\ b, (a;a;2)$.

1.2 Symbol macros for Breathings

textalpha defines the macros `\<` and `\>` for the **dasia** (rough breathing) and **psili** (smooth breathing) diacritics.

⁴The Unicode standard provides the code point 037E GREEK QUESTION MARK but says character 003B SEMICOLON and not 037E is the preferred character for a ‘Greek question mark’ (*erotimatiko*).

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
370	◦	◦	◦	◦	’	’	◦	◦			◦	◦	◦	◦	;	
380					’	’	Α	·	Έ	Ή	Ί		Ό		Υ	Ω
390	ί	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Μ	Ν	Ξ	Ο
3A0	Π	Ρ		Σ	Τ	Υ	Φ	Χ	Ψ	Ω	Ϊ	Ϋ	ά	έ	ή	ί
3B0	ϐ	α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο
3C0	π	ρ	ς	σ	τ	υ	φ	χ	ψ	ω	ϊ	ϋ	ό	ύ	ώ	
3D0	◦	◦	◦	◦	◦	◦	◦	◦	◦	Q	q	ς	ς	F	ƒ	℥
3E0	ᲀ	ᲁ	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦
3F0	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦	◦

Table 1: Greek and Coptic Unicode Block, input as literal Unicode characters in TU font encoding (legend: ◦ glyph missing in LGR).

3.1.1 Sigma

The lower Sigma comes in two variants: `\textsigma` σ is used inside a word and `\textfinalsigma` ς (or `\textvarsigma` ς) at the end of words.

In LGR, the Latin letter `s` and the command `\textautosigma` print the “normal” sigma if followed by another letter and the final sigma if followed by space or punctuation. This is implemented via the font ligature mechanism in LGR⁷:

$\sigma\sigma$ (TU) vs. $\sigma\sigma$ (TU).

The upper case of both sigma variants is `\textSigma`, the lower case of `\textSigma` is `\textautosigma`.

Test Unicode literal and `\text...` commands:

no change:	$\sigma\sigma$ $\varsigma\varsigma\sigma$ $\Sigma\Sigma$ $\Sigma\Sigma$
MakeUppercase:	$\Sigma\Sigma$ $\Sigma\Sigma\Sigma\Sigma$ $\Sigma\Sigma$ $\Sigma\Sigma$
MakeLowercase (TU):	$\sigma\sigma$ $\varsigma\varsigma\sigma$ $\sigma\varsigma$ $\sigma\sigma$
MakeLowercase (TU):	$\sigma\sigma$ $\varsigma\varsigma\sigma$ $\sigma\varsigma$ $\sigma\sigma$

3.2 Greek literal characters in non-Greek text

With the *textalpha* package, **greek-inputenc** and input encoding `utf8`, Greek Unicode literals can be used in text with any font encoding. See Tables ?? and ??.

Kerning is preserved if the active font encoding supports Greek. This can be secured by wrapping the Greek text part in `\ensuregreek` or setting the text language with Babel: AYA

Combined Diacritics work for pre-composed characters: ῥ .

Diacritics (except diaeresis) are dropped with MakeUppercase with LaTeX versions older than 2022/06 For other versions, set the language of to-be-upcased Greek text with Babel: $\mu\acute{\alpha}\iota\sigma\tau\rho\omicron\varsigma$, $\delta\acute{\upsilon}\omicron \mapsto \text{ΜΑΪΣΤΡΟΣ}$, $\Delta\acute{\upsilon}\omicron$.

⁷TODO: Fix `\textautosigma` with Unicode fonts.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1F00	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ
1F10	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F20	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F30	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F40	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F50	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F60	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F70	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F80	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1F90	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FA0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FB0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FC0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FD0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FE0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ
1FF0	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠬ	ᠭ	ᠠ	ᠡ

Table 2: Greek Extended Unicode Block, input as literal Unicode characters in TU font encoding.

3.3 PDF strings

With *textalpha* and *greek-inputenc*, there are two options to get Greek letters in PDF strings: LICR macros and literal Unicode input.

3.3.1 λογος, λογος, and logos

The subsection title above uses: LICR macros, Unicode input and the LGR transliteration for the Greek word logos.

LICR macros and Unicode literals work fine everywhere, the Latin transliteration remains Latin in the PDF metadata (sidebar table of contents in the PDF viewer) and with Xe/LuaTeX.

3.3.2 α β γ δ ε ζ η θ ι κ λ μ ν ξ ο π ρ σ ς τ υ φ χ ψ ω

3.3.3 ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ

3.3.4 □□ □□ qQ □□ □□ · ; ,

Archaic characters are missing in many fonts, including the “Biolinum” font used in this document. .