

The *textalpha* package

Günter Milde

November 28, 2013

With the *textalpha* package, you can easily write a single Greek symbol (like Ψ or μ) or a $\lambda\omicron\gamma\omicron\varsigma$ in non-Greek text as well as ISO-conforming formulas with upright constants (like π): $A = \pi r^2$ vs. $A = \pi r^2$. Input is possible via LICR macros (`\textalpha ... \textomega`) or (with *greek-inputenc* and the `utf8` option to *inputenc*) Unicode literals.

1 Greek alphabet

Greek letters via Latin transcription in LGR font encoding:

A B Γ Δ E Z H Θ I K Λ M N Ξ O Π P Σ T Υ Φ X Ψ Ω
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

Greek letters via default macros in other font encoding (here T1):

A B Γ Δ E Z H Θ I K Λ M N Ξ O Π P Σ T Υ Φ X Ψ Ω
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

Archaic Greek letters and Greek punctuation

Ϝ ϝ Ϟ ϟ;
Ϡ ϡ ϣ Ϥ ϥ

Diacritics

Symbol macros:¹ `'' ' ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~`

Named macros: `'' ' ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~ ^ ~`

¹Composite diacritics require wrapping in `\ensuregreek`.

2 Greek Unicode characters in non-Greek text

With the *textalpha* package and input encoding `utf8`, Greek Unicode characters can be used in text with any font encoding.

Kerning is preserved if the font encoding is LGR. This holds also for pre-composed accented characters: ΑΨΑ.

All characters of the Greek and Coptic Unicode Block present in LGR via literal Unicode input in T1 font encoding (legend: * glyph missing in LGR):

	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	ι	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	ύ	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	*	*	*	*	*	*	*	*	Ϡ	ϡ	Ϣ	ϣ	Ϥ	*		ι
3E0	λ	ϝ	*	*	*	*	*	*	*	*	*	*	*	*	*	*
3F0	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

All characters of the Greek Extended Unicode Block via literal Unicode input in T1 font encoding:

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1F00	ᾀ	ᾁ	ᾂ	ᾃ	ᾄ	ᾅ	ᾆ	ᾇ	Α	Α	Α	Α	Α	Α	Α	Α
1F10	ᾈ	ᾉ	ᾊ	ᾋ	ᾌ	ᾍ			Ε	Ε	Ε	Ε	Ε	Ε	Ε	
1F20	ᾎ	ᾏ	ᾐ	ᾑ	ᾒ	ᾓ	ᾔ	ᾕ	Η	Η	Η	Η	Η	Η	Η	Η
1F30	ᾖ	ᾗ	ᾘ	ᾙ	ᾚ	ᾛ	ᾜ	ᾝ	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι
1F40	ᾞ	ᾟ	ᾠ	ᾡ	ᾣ	ᾤ			Ο	Ο	Ο	Ο	Ο	Ο		
1F50	ᾥ	ᾆ	ᾧ	ᾨ	ᾩ	ᾪ	ᾫ	ᾬ		Υ		Υ		Υ		Υ
1F60	ᾭ	ᾮ	ᾯ	ᾰ	ᾱ	ᾲ	ᾳ	ᾴ	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω
1F70	ᾶ	ᾷ	Ᾰ	Ᾱ	Ὰ	Ά	ᾼ	᾽	ὀ	ὁ	ὂ	ὃ	ὄ	ὅ		
1F80	᾿	᾿	᾿	᾿	᾿	᾿	᾿	᾿	Α _ι	Α _ι	Α _ι	Α _ι	Α _ι	Α _ι	Α _ι	Α _ι
1F90	῀	῁	ῂ	ῃ	ῄ	῅	ῆ	ῇ	Η _ι	Η _ι	Η _ι	Η _ι	Η _ι	Η _ι	Η _ι	Η _ι
1FA0	ῆ	ῇ	Ὲ	Έ	Ὴ	Ή	ῌ	῍	Ω _ι	Ω _ι	Ω _ι	Ω _ι	Ω _ι	Ω _ι	Ω _ι	Ω _ι
1FB0	᾿	᾿	᾿	᾿	᾿		᾿	᾿	ᾷ	ᾷ	ᾷ	ᾷ	ᾷ		ᾷ	ᾷ
1FC0	ῶ	ῷ	Ὸ	Ό	Ὼ		Ώ	ῼ	Ε	Ε	Η	Η	Η		ῼ	ῼ
1FD0	´	῾	῿	῿			῿	῿	ῼ	ῼ	ῼ	ῼ			ῼ	ῼ
1FE0	῾	῿	῾	῾	῾	῾	῾	῾	Υ	Υ	Υ	Υ	Υ	Υ	῾	῾
1FF0			῾	῾	῾		῾	῾	Ο	Ο	Ω	Ω	Ω	Ω	῾	῾

Combined Diacritics work ᾿, diacritics (except diaeresis) are dropped with Make-Uppercase (μαῖστρος \mapsto ΜΑΪΣΤΡΟΣ).

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.

3.1 λογος, λογος and λογος

The subsection title above uses: LICR macros, Unicode input and the LGR transcription for the Greek word λογος. Check the table of contents in the PDF viewer: LICR macros and Unicode literals work fine, the Latin transcription remains Latin in the PDF metadata.

4 Limitations

Because the internal font encoding switch interferes with other work behind the scenes, kerning, diacritics and up/down-casing show problems if Greek letters are used without explicit change of the font encoding. These problems can be avoided by use of *babel* and the correct language setting (greek) or an explicit font encoding switch (e.g. wrapping the Greek text in `\ensuregreek2`).

4.1 Diacritics

Composition of diacritics (like `\accdasia\acctonos`) fails in other font encodings. Long names (like `\accdasiaoxia`) work, however they do not select pre-composed characters. With LGR, pre-composed glyphs are chosen if available (the difference becomes obvious if you drag-and-drop text from the PDF version of this document): Ᾱ Ὰ ᾽ ᾿ ᾿ (LGR) vs. Ᾱ (T1).

According to Greek typographical tradition, diacritics (except the dialytika) are placed before capital letters in titlecase and dropped in all-caps:

Ᾱ Ὰ ᾽ ᾿ ᾿ ᾿
Ᾱ Ὰ ᾽ ᾿ ᾿ ᾿
A E Ī H O Υ Ω.

This fails if the active font encoding is not LGR: Ᾱ (LGR) vs. Ᾱ(T1). To overcome this, the `\ensuregreek` macro is used in composite Unicode character definitions: Ᾱ (LGR) = Ᾱ (T1).

The dialytika marks a *hiatus* (break-up of a diphthong). It must be present in UPPERCASE even where it is redundant in lowercase (the hiatus can also be

²The `\ensuregreek` macro ensures the argument is set in a font encoding supporting Greek. This can be used to fix these problems without adverse side-effects if the active font encoding is already LGR.

marked by an accent on the first character of a diphthong). The auto-hiatus feature works in LGR font encoding only: $\acute{\alpha}\upsilon, \acute{\epsilon}\iota \mapsto \text{A}\acute{\Upsilon}, \text{E}\acute{\text{I}}$ vs. $\text{A}\Upsilon, \text{EI}$.

Currently, the second vowel of the diphthong must be given as macro, not Unicode literal if the auto-hiatus feature should work: $\acute{\alpha}\upsilon\pi\nu\acute{\iota}\alpha \mapsto \text{A}\acute{\Upsilon}\Pi\text{NIA}$ vs. $\text{A}\Upsilon\Pi\text{NIA}$.

4.2 Kerning

No kerning occurs between Greek characters in non-Greek text due to the internal font encoding switch: compare $\text{A}\Upsilon\text{A}$ (LGR) to $\text{A}\Upsilon\text{A}(\text{T1})$.

Compiling with LuaTeX provides kerning also over font encoding boundaries.