

# Greek Unicode with *inputenc*

Günter Milde

July 18, 2013

## Abstract

The definitions in `lgrenc.dfu` provide UTF-8 support for Greek based on `inputenc` and the LICRs (LaTeX internal character representations) defined in `lgrenc.def` from the `greek-fontec` package.

## 1 Usage

If the *greek-fontenc* package is installed, the line

```
\usepackage[utf8]{inputenc}
```

in the document preamble allows the use of Greek Unicode literals in the document body, if the current font encoding is LGR.<sup>1</sup> (The font encoding is automatically set to LGR by babel for the *greek* language.)

In combination with the *textalpha* package from `lgrx`, Greek Unicode literals can be used in text with any font encoding just like the symbols provided by package “textcomp” (i.e. with some limitations to kerning, MakeUppercase/MakeLowercase and multiple accents, described in `textalpha-test` from the “lgrx” package):

```
\usepackage[utf8]{inputenc}
\usepackage{textalpha}
```

In combination with the *alphabet* package (also from `lgrx`), Greek Unicode literals can also be used in math mode:

```
\usepackage[utf8]{inputenc}
\usepackage{alphabet}
```

## 2 Supported Characters

LGR is no “standard font encoding”. Latin characters are mapped to Greek if LGR is the active font encoding. This means you need an explicit font-encoding switch for Latin words and abbreviations in Greek text.

Unicode definitions exist for all non-ASCII characters that can be rendered with an LGR-encoded font.

---

<sup>1</sup>All *inputenc* Unicode definitions only work for text in the correct font encoding.

## 2.1 Greek and Coptic

legend: \* glyph missing in LGR, [space] Unicode point not defined

	0	1	2	3	4	5	6	7	8	9	A	B	C	Δ	E	Φ
370	*	*	*	*	´		*	*				*	*	*	;	
380					´	´	Α	·	Ε	Η	Ι		Ο		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	*	*	*	*	*	*	*	*	ϙ	ϛ	ϝ	ϟ	ϡ	ϣ	*	ι
3E0	λ	ϝ	*	*	*	*	*	*	*	*	*	*	*	*	*	*
3F0	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

## 2.2 Greek Extended

	0	1	2	3	4	5	6	7	8	9	A	B	C	Δ	E	Φ
1F00	ά	ά	ᾀ	ᾀ	ᾀ	ᾀ	ᾀ	ᾀ	Α	Α	Α	Α	Α	Α	Α	Α
1F10	έ	έ	ῃ	ῃ	ῃ	ῃ	ῃ	ῃ	Ε	Ε	Ε	Ε	Ε	Ε	Ε	Ε
1F20	ή	ή	ῆ	ῆ	ῆ	ῆ	ῆ	ῆ	Η	Η	Η	Η	Η	Η	Η	Η
1F30	ι	ι	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι
1F40	ό	ό	ὀ	ὀ	ὀ	ὀ			Ο	Ο	Ο	Ο	Ο	Ο	Ο	
1F50	ύ	ύ	ϋ	ϋ	ϋ	ϋ	ϋ	ϋ		Υ		Υ		Υ		Υ
1F60	ώ	ώ	ὠ	ὠ	ὠ	ὠ	ὠ	ὠ	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω
1F70	ᾶ	ᾶ	ῆ	ῆ	ῆ	ῆ	ῆ	ῆ	ὀ	ὀ	ὀ	ὀ	ὀ	ὀ	ὀ	
1F80	ᾷ	ᾷ	ῇ	ῇ	ῇ	ῇ	ῇ	ῇ	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>	Α <sub>1</sub>
1F90	ῆ	ῆ	ῆ	ῆ	ῆ	ῆ	ῆ	ῆ	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>	Η <sub>1</sub>
1FA0	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>	Ι <sub>1</sub>
1FB0	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Α	Α	Α	Α	Α	Α	Α	Α
1FC0	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Ε	Ε	Ε	Ε	Ε	Ε	Ε	Ε
1FD0	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι
1FE0	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Υ	Υ	Υ	Υ	Υ	Υ	Υ	Υ
1FF0			ῖ	ῖ	ῖ	ῖ	ῖ	ῖ	Ο	Ο	Ο	Ο	Ο	Ο	Ο	Ο

## 2.3 Other Unicode Blocks

Latin-1 Supplement: “ « ’ ’ »

IPA Extensions: ə

Spacing Modifier Letters: ˘α (here followed by letter alpha)

General Punctuation: — ‘ ’ ‰ ZWNJ (zero width no joiner, prevents kerning and ligatures, e.g. ΑΥ vs. ΑΥ and ’α vs. ά)

Currency Symbols: €

Letterlike Symbols: Ω

Ancient Greek Numbers: ͵Ͷͷ͸͹

## 3 MakeUppercase, MakeLowercase

Capital Greek letters have diacritics (except the dialytika) to the left (instead of above) and drop them in uppercase, e.g. μαῖστρος ↦ ΜΑΪΣΤΡΟΣ.

The auto-hiatus feature in `lgrxenc.def` works nice with the Latin transcription and with character-macros ( $\text{\AA}$ ,  $\text{\A}$ ,  $\text{\E}$ ) and also if the first character is wrapped in `\ensuregreek` (as done by the `lgrnc.dfu` definition for accented characters) or a literal Unicode character ( $\text{\AA}$ ,  $\text{\A}$ ,  $\text{\A}$ ) but not if the second character of the diphthong is a Unicode literal ( $\text{\AE}$ ,  $\text{\AE}$ ,  $\text{\EI}$ ).

Fixing this shortcoming requires knowledge of what the `\LGR@ifnextchar` “sees” when the next character is an upcased Unicode literal.

The following subsections test MakeUppercase and MakeLowercase with all characters defined in lgrenc.dfu:

### Characters of the Greek and Coptic Unicode Block:

MakeUppercase:

Letters and ypogegrammeni upcased, tonos dropped, dialytika kept.

MakeLowercase:

The lowercase of  $\Sigma$  is the «auto-sigma» (`\textautosigma`):  $\Sigma\Sigma \mapsto \sigma\varsigma$ .

3

### 3.2 Greek extended

MakeUppercase:

Α Α Α Α Α Α Α Α Α Α Α Α Α Α Α  
Ε Ε Ε Ε Ε Ε Ε Ε Ε Ε Ε  
Η Η Η Η Η Η Η Η Η Η Η Η Η Η Η  
Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι Ι  
Ο Ο Ο Ο Ο Ο Ο Ο Ο Ο Ο Ο  
Υ Υ Υ Υ Υ Υ Υ Υ Υ Υ Υ Υ  
Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω  
Α Α Ε Ε Η Η Ι Ι Ο Ο Υ Υ Ω Ω  
Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub>  
Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub>  
Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub>  
Ά Ά Α<sub>1</sub> Α<sub>1</sub> Α<sub>1</sub> Α Α<sub>1</sub> Ά Ά Α Α Α<sub>1</sub> ,  
“ Η<sub>1</sub> Η<sub>1</sub> Η<sub>1</sub> Η Η<sub>1</sub> Ε Ε Η Η Η<sub>1</sub>  
Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ  
Υ̂ Υ̂ Υ̂ Υ̂ Ρ Ρ Υ̂ Υ̂ Υ̂ Υ̂ Υ̂ Ρ ...  
Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ω<sub>1</sub> Ο Ο Ω Ω<sub>1</sub>

MakeLowercase:

ά ά ἀ ἀ ἀ ἀ ἀ ἀ ἀ ἀ ἀ ἀ ἀ  
έ έ ε ε ε ε ε ε ε ε ε ε ε  
ή ή η η η η η η η η η η η  
ι ι ι ι ι ι ι ι ι ι ι ι ι ι ι  
ο ό ο ο ο ο ο ο ο ο ο ο ο  
ύ ύ υ υ υ υ υ υ υ υ υ υ υ  
ώ ώ ω ω ω ω ω ω ω ω ω ω ω  
α α ε ε η η ι ι ò ó ù ú ω ω  
ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ  
ή ή ή ή ή ή ή ή ή ή ή ή ή ή ή  
φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ  
ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ  
~ η η ή ή έ έ η ή η ~  
ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ ϛ  
ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ ϣ  
φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ φ

### 3.3 Other Unicode Blocks

MakeUppercase works without effect on non-letter symbols and the letter shwa:

“ « - ’ . » ə ˘ A — ‘ ’ %₀ AΥ € ☒ ☒ ☒ ☒

MakeLowercase works without effect on non-letter symbols, too:

“ « - ’ . » ə ˘ α — ‘ ’ %₀ αυ € ☒ ☒ ☒ ☒