

# Greek Unicode with `lgrenc.dfu`

The definitions in `lgrenc.dfu` provide standards-compatible UTF-8 support for Greek based on the LICRs (LaTeX internal character representations) defined in `lgrxenc.def`.

Like other *inputenc* Unicode definitions, they only work for text in the correct font encoding (in this case LGR). The font encoding is automatically set to LGR by `babel` for the languages *greek* and *polutonikogreek*.

With the combination

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

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`).

## 1 Supported Characters

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

### 1.1 Greek and Coptic

legend: \* glyph missing in LGR, · Unicode point not defined

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
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	ό	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	*	*	*	*	*	*	*	*	Ω	ι	Γ	τ	Φ	·	*	ι
3E0	λ	λ	*	*	*	*	*	*	*	*	*	*	*	*	*	*
3F0	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

## 1.2 Greek Extended

	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	Ἀ	Ἀ	Ἀ	Ἀ	Ἀ	Ἀ	Ἀ	Ἀ
1F10	ἒ	ἓ	ἔ	ἕ	἖	἗	·	·	Ἐ	Ἐ	Ἐ	Ἐ	Ἐ	Ἐ	·	·
1F20	ἡ	ἢ	ἣ	ἣ	ἥ	ἥ	ἥ	ἥ	Ἠ	Ἠ	Ἠ	Ἠ	Ἠ	Ἠ	Ἠ	Ἠ
1F30	ἰ	ἱ	ἲ	ἳ	ἴ	ἵ	ἶ	ἷ	Ἰ	Ἰ	Ἰ	Ἰ	Ἰ	Ἰ	Ἰ	Ἰ
1F40	ὀ	ὁ	ὂ	ὃ	ὄ	ὅ	·	·	Ὀ	Ὀ	Ὀ	Ὀ	Ὀ	Ὀ	·	·
1F50	ὐ	ὕ	ὖ	ὗ	὘	Ὑ	὚	Ὓ	·	Υ	·	Υ	·	Υ	·	Υ
1F60	ὠ	ὡ	ὢ	ὣ	ὤ	ὥ	ὦ	ὧ	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω
1F70	ᾐ	ᾑ	ᾒ	ᾓ	ᾔ	ᾕ	ᾖ	ᾗ	ὀ	ὁ	ὐ	ὕ	ὖ	ὗ	·	·
1F80	ᾘ	ᾙ	ᾚ	ᾛ	ᾜ	ᾝ	ᾞ	ᾟ	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	ἡ	ἢ	ἣ	ἣ	ἥ	ἥ	ἥ	ἥ	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	ὀ	ὁ	ὂ	ὃ	ὄ	ὅ	὆	὇	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>
1FB0	ᾘ	ᾙ	ᾚ	ᾛ	ᾜ	·	ᾞ	ᾟ	Α	Α	Α	Α	Α <sub>1</sub>	·	·	·
1FC0	~	~	ἡ	ἢ	ἣ	·	ἥ	ἥ	Ἐ	Ἐ	Ἠ	Ἠ	Η <sub>1</sub>	·	·	·
1FD0	ἰ	ἱ	ἲ	ἳ	·	·	ἶ	ἷ	Ἰ	Ἰ	Ἰ	Ἰ	·	·	·	·
1FE0	ὐ	ὕ	ὖ	ὗ	὘	Ὑ	὚	Ὓ	Υ	Υ	Υ	Υ	Ρ	·	·	·
1FF0	·	·	ὀ	ὁ	ὂ	·	ὦ	ὧ	Ὀ	Ὀ	Ὀ	Ὀ	Ω <sub>1</sub>	·	·	·

## 1.3 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. ᾱ)

Currency Symbols: €

Letterlike Symbols: Ω

Ancient Greek Numbers: ͵Ͷ ͵ͷ ͵͸ ͵͹

## 2 MakeUppercase, MakeLowercase

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

Tonos and dasia on the first vowel of a diphthong (αῖ, αὖ, εῖ) imply a *hiatus*. A ditytika must be placed on the second vowel if they are dropped (Αἶ, Αῦ, Εἶ).

The auto-hiatus feature in lgrxenc.def works nice with the Latin transcription and with character-macros (Αἶ, Αῦ, Εἶ) and also if the first character is wrapped in \ensuregreek (as done by the lgrenc.dfu definition for accented characters) or a literal Unicode character (Αἶ, Αῦ, Αἶ) but not if the second character of the diphthong is a Unicode literal (ΑΙ, ΑΥ, ΕΙ).

Therefore, the diaeresis is missing in the following examples:  $\acute{\alpha}\upsilon\lambda\omicron\varsigma \mapsto \text{AY}\Lambda\text{O}\Sigma$ ,  $\acute{\alpha}\upsilon\lambda\omicron\varsigma \mapsto \text{AY}\Lambda\text{O}\Sigma$ ,  $\mu\acute{\alpha}\iota\nu\alpha \mapsto \text{MAIN}\alpha$ ,  $\kappa\acute{\epsilon}\iota\kappa, \mapsto \text{KEIK}$ ,  $\acute{\alpha}\upsilon\pi\nu\acute{\iota}\alpha \mapsto \text{AY}\Pi\text{NIA}$ .

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

As an ugly workaround, use `\textiota` resp. `\textupsilon` for the character that should get the diaeresis:  $\acute{\alpha}\upsilon\pi\nu\acute{\iota}\alpha \mapsto \text{A}\acute{\text{Y}}\Pi\text{NIA}$ .

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

## 2.1 Greek and Coptic

Characters of the Greek and Coptic Unicode Block:

```
';; 'A·EHTO'YŦABΓΔΕΖΗΘΙΚΑΜΝΞΟΠΡΣΤΥΦΧΨΩĲŦTFλ
άέήίϵαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώıƒıʀ
```

`MakeUppercase`:

```
';; "A·EHIOTYŦABΓΔΕΖΗΘΙΚΑΜΝΞΟΠΡΣΤΥΦΧΨΩĲŦTFλ
ΑΕΗΙŦΑΒΓΔΕΖΗΘΙΚΑΜΝΞΟΠΡΣΣΤΥΦΧΨΩĲŦΟΥΩƒıʀ
```

Letters and ypogegrammeni upcased, tonos dropped, dialytika kept.

There is no capital Koppa in LGR, therefore  $\text{ı}$  is left unchanged with `MakeUppercase`.

`MakeLowercase`:

```
';; 'άέήίόύώıαβγδεζηθικλμνξοπρςστυφχψωϊϋıʀƒıʀ
άέήίϵαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώıƒıʀ
```

The lowercase of  $\Sigma$  is the «auto-sigma» (`\textautosigma`):  $\Sigma\S \mapsto \sigma\varsigma$ . Add a ZWNJ or use the `\noboundary` macro to prevent conversion to final sigma:  $\sigma\sigma$ . The lowercase of GREEK LETTER STIGMA  $\text{Ͳ}$  is  $\text{ͳ}$ .

## 2.2 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  
Η 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 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 Υ  
Ω 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>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub>  
Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub>  
Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub>  
Α̇ Α̇ Α<sub>ι</sub> Α<sub>ι</sub> Α<sub>ι</sub> Α Α<sub>ι</sub> Α̇ Α̇ Α Α Α<sub>ι</sub> ῀  
῀ Η<sub>ι</sub> Η<sub>ι</sub> Η<sub>ι</sub> Η Η<sub>ι</sub> Ε Ε Η Η Η<sub>ι</sub>  
῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀ ῀  
Υ̇ Υ̇ Υ̇ Υ̇ Ρ Ρ Υ̇ Υ̇ Υ̇ Υ̇ Υ̇ Ρ Ρ  
Ω<sub>ι</sub> Ω<sub>ι</sub> Ω<sub>ι</sub> Ω Ω Ω Ο Ο Ω Ω Ω<sub>ι</sub>

MakeLowercase:

ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ  
ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ ἐ  
ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ  
ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ  
ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ  
ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ ὐ  
ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ  
ᾰ ᾰ ᾰ ᾰ ῆ ῆ ῖ ῖ ὀ ὀ ὐ ὐ ὦ ὦ  
ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ ᾱ  
ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ  
ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ  
ᾰ ᾰ ᾱ ᾱ ᾱ ᾱ ᾱ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ ᾰ  
ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ ῆ  
ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ  
ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ ῡ  
ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ

## 2.3 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 € ☒ ☒ ☒ ☒