

Greek Unicode with `lgrenc.dfu`

The definitions in `lgrenc.dfu` provide standards-compatible UTF-8 support for Greek based on the extended font encoding definitions 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 (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	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	*	*	*	*	*	*	*	*	*	Ω	ϙ	ϒ	Ϝ	ϝ	*	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	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ	ǎ
1F10	ê	ê	ê	ê	ê	ê	.	.	Ē	Ē	Ē	Ē	Ē	Ē	.	.
1F20	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ
1F30	ï	ï	ï	ï	ï	ï	ï	ï	İ	İ	İ	İ	İ	İ	İ	İ
1F40	ó	ó	ó	ó	ó	ó	.	.	Ō	Ō	Ō	Ō	Ō	Ō	.	.
1F50	ù	ù	ù	ù	ù	ù	ù	ù	.	Ÿ	.	Ÿ	.	Ÿ	.	Ÿ
1F60	ô	ô	ô	ô	ô	ô	ô	ô	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω
1F70	ǻ	ǻ	ǻ	ǻ	ǻ	ǻ	ǻ	ǻ	ò	ó	ù	ú	ò	ó	.	.
1F80	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ
1F90	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ	Ĥ
1FA0	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	ǿ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ	Ǻ
1FB0	ǻ	ǻ	ǻ	ǻ	ǻ	.	ǻ	ǻ	Ǻ	Ǻ	Ǻ	Ǻ	A	.	.	.
1Fc0	~	~	ĥ	η	η	.	η	η	Ē	Ē	Ĥ	Ĥ	H	.	~	~
1FD0	ı	ı	ı	ı	.	.	ı	ı	İ	İ	İ	İ	.	ı	ı	ı
1FE0	ü	ü	ü	ü	ü	ü	ü	ü	Ÿ	Ÿ	Ÿ	Ÿ	Ÿ	.	.	.
1FF0	.	.	ò	ò	ò	.	ò	ò	Ō	Ō	Ō	Ō	Ō	.	.	.

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. AŸ vs. A Y and ‘α vs. α)

Currency Symbols: €

Letterlike Symbols: Ω

Ancient Greek Numbers: Δ H X M

2 MakeUppercase, MakeLowercase

Capital Greek letters have diacritics (except the dialytika) to the left (instead of above) and drop them in uppercase, e.g. $\mu\alpha\acute{\iota}\sigma\tau\rho\omicron\varsigma \mapsto \text{ΜΑΪΣΤΡΟΣ}$.

Tonos and dasia mark a *hiatus* (break-up of a diphthong) if placed on the first vowel of the diphthong (άι, άυ, έι). A dialytika must be placed on the second vowel if they are dropped (Αἶ, Αῦ, Εἶ).

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

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: $\acute{\alpha}\upsilon\pi\nu\acute{\iota}\alpha \mapsto A\Upsilon\Pi N\acute{I}A$.

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'T'ΩiABΓΔEZHΘIKAMNΞOΠPΣT'TΦXΨΩi'Y'QTFΛ
άέήίύαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώϘϙϒϓϔϕϖϗϘϙϚϛϜϝϞϟϠϡϢϣϤϥϦϧϨϩϪϫϬϭϮϯϰϱϲϳϴϵ϶ϷϸϹϺϻϼϽϾϿϠϡϢϣϤϥϦϧϨϩϪϫϬϭϮϯϰϱϲϳϴϵ϶ϷϸϹϺϻϼϽϾϿ
```

Tonos dropped with `MakeUppercase`:

```
';; "A·EHIOYΩiABΓΔEZHΘIKAMNΞOΠPΣT'TΦXΨΩi'Y'QTFΛ
AEHIYABΓΔEZHΘIKAMNΞOΠPΣΣT'TΦXΨΩi'Y'OTYQϘϙϒϓϔϕϖϗϘϙϚϛϜϝϞϟϠϡϢϣϤϥϦϧϨϩϪϫϬϭϮϯϰϱϲϳϴϵ϶ϷϸϹϺϻϼϽϾϿ
```

`MakeLowercase`:

```
';; 'A·έήίόύώίαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώϘϙϒϓϔϕϖϗϘϙϚϛϜϝϞϟϠϡϢϣϤϥϦϧϨϩϪϫϬϭϮϯϰϱϲϳϴϵ϶ϷϸϹϺϻϼϽϾϿϠϡϢϣϤϥϦϧϨϩϪϫϬϭϮϯϰϱϲϳϴϵ϶ϷϸϹϺϻϼϽϾϿ
```

The lowercase of Σ is the «auto-sigma»: $T\Sigma \mapsto \tau\varsigma$, $\Sigma T \mapsto \sigma\tau$. Add a `ZWNJ` or use the `\noboundary` macro to prevent conversion to final sigma: $\tau\sigma$. The lowercase of Υ is Υ .

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
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
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
Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁ Α₁
Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁ Η₁
Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁ Ω₁
Ά Ά Α₁ Α₁ Α₁ Α Α₁ Ά Ά Α Α Α₁ ,
~Η₁ Η₁ Η₁ Η Η₁ Ε Ε Η Η Η₁
Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ
Υ̂ Υ̂ Υ̂ Υ̂ Ρ Ρ Υ̂ Υ̂ Υ̂ Υ̂ Υ̂ Ρ ...
Ω₁ Ω₁ Ω₁ Ω Ω₁ Ο Ο Ω Ω Ω₁

ΜακεΛωερςασε:

ά ά ἀ ᾀ ᾶ ᾷ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ Ᾱ
έ ε ῔ ῥ Ὶ Ί ῞ ῟ ῠ ῡ ῢ ΰ ῤ ῦ ῖ ῗ
ή ή η ή ῥ ΐ ῥ ΐ ῥ ΐ ῥ ΐ ῥ ΐ ῥ ΐ
ι ι ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ ῖ
ο ο ὀ ὄ ὅ ὅ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ ὀ
ύ ύ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ ὕ
ώ ώ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ ὦ
ὰ ἀ ῆ ῆ ῖ ῖ ὀ ὀ ὕ ὕ ὦ ὦ
ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ ᾷ
ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ ῥ
ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ ᾶ
~ή η ή η ή έ έ ή ή ~~~~~
Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ Ϊ
ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ ϋ
ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ ϖ

2.3 Other Unicode Blocks

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

“ « - ’ . » ə ˘ A — ‘ ’ ‰ A Y € ☒ ☒ ☒ ☒

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

“ « - ’ . » ə ˘ α — ‘ ’ ‰ α υ € ☒ ☒ ☒ ☒