Test LGR font encoding definitions

Günter Milde

2020/10/30

The file lgrenc.def provides a comprehensive set of macros to typeset Greek with LGR encoded fonts. It works for both, monotonic and polytonic Greek, independent of the *Babel* package.

The example from `usage.tex` in *babel-greek* input using the LICR macros:

> Τί φήις; Ἰδὼν ἐνθέδε παῖδ' ἐλευθέραν
> τὰς πλησίον Νύμφας στεφανοῦσαν, Σώστρατε,
> ἐρῶν ἀπῆλθες εὐθύς;

Contents

1 Symbols

See the source file lgrenc-test.tex for the macros used to access the symbols.

1.1 Generic text symbols

Direct input:

Latin: + - = < > – — [()]

LGR: + - = ‘ ’ – — [()] Less than and greater than characters are mapped to breathing marks, use LICRs.

Character macros:

Latin:

> < `\textless`
> > `\textgreater`
> { `\textbraceleft`
> } `\textbraceright`
> \ `\textbackslash`
> | `\textbar`
> ‰ `\textperthousand`

‰ \textpertenthousand
␣ \textvisiblespace

LGR: (per-tenthousand sign is missing in LGR, LaTeX selects the global default)

< \textless
> \textgreater
{ \textbraceleft
} \textbraceright
\ \textbackslash
| \textbar
‰ \textperthousand
‰ \textpertenthousand
␣ \textvisiblespace

Quotes:[1] «a» «α», ‘a’ ‘α’, “a” “α” (double quotes wrong with Kerkis fonts)

Single guillemets and base-quotes (‹a› „a” ‚a’) are missing in LGR.

Suppress ligatures and kerning with \textcompwordmark: AY fi AϒϊAY fi AϒϊA

Spacing accent chars: ^a ^α ^ι ~a ~α ~ι ˘a ˘α ˘ι ¯a ¯α ¯ι ¨a ¨α ¨ι ´a ´α ´ι `a `α `ι

Letter schwa and Euro symbol: ə \textschwa, € \texteuro

Some ASCII symbols are replaced by different symbols in LGR encoding other symbols are composed from Latin letters and show Greek letters in LGR. *babel-greek* redefines some affected macros to use a standard font encoding, however this cannot be done in a font encoding definition file.

Beware that " # & ' ; < > ? @ becomes ’ ‟ ‧ ′ · ʿ ʾ ; ˝!

The *textcomp* package[2] provides pre-composed coyright ©, registered ® and trademark ™ symbols that work in all font encodings. Up-to-date LaTeX installations load *textcomp* by default. The legacy composite fallback definitions showed Greek letters if used in LGR: © \textcopyright, ® \textregistered, ™ \texttrademark.

textcomp also provides the upright MICRO SIGN and OHM SIGN for SI units: R = 5 µΩ

In LGR, \textmicro and \textohm are aliases to \textmu and \textOmega that do not change case: Αντίσταση = 5 µΩ, ΑΝΤΊΣΤΑΣΗ = 5 µΩ, αντίσταση = 5 µΩ.

1.2 Greek alphabet

Greek letters via Latin transliteration and LICR macros:

[1] Single quotes need special attention to prevent conversion to accents. Test the input conventions: ‘α’ ‘α’ ‘α’ ‘α’ but not ὰ᾽ ἐ᾽ ϊ᾽

[2] loaded by default in not too old LaTeX

Α Β Γ Δ Ε Ζ Η Θ Ι Κ Λ Μ Ν Ξ Ο Π Ρ Σ Τ Υ Φ Χ 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

𝚨 𝚩 𝚪 𝚫 𝚬 𝚭 𝚮 𝚯 𝚰 𝚱 𝚲 𝚳 𝚴 𝚵 𝚶 𝚷 𝚸 𝚺 𝚻 𝚼 𝚽 𝚾 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

The small sigma is set with a different glyph if it ends a word:

σ `\textsigma`
ς `\textfinalsigma`

In the Latin transliteration, the letter 's' stands for `\textautosigma` which automatically chooses the glyph according to the position.

1.3 additional Greek symbols

ϟ `\textkoppa` (numeral koppa = 90)
Ϟ `\textKoppa` (numeral Koppa = 90)[3]
ϙ `\textqoppa` (archaic koppa)
Ϙ `\textQoppa` (archaic Koppa)
ϛ `\textstigma`
ϲ `\textvarstigma`
Ϛ `\textStigma` (Sigma-Tau-Ligature in CB-fonts)[4]
ϡ `\textsampi`
Ϡ `\textSampi`
ϝ `\textdigamma`
Ϝ `\textDigamma`
΄ `\textdexiakeraia` (dexia keraia)
῾ `\textaristerikeraia` (aristeri keraia)

Up/Downcasing of the additional Symbols from the Greek And Coptic Unicode block:

΄ ῾ ͵ ; ΄ ΅ Ά · Έ Ή Ί Ό Ύ Ώ ΐ Ϊ Ϋ ά έ ή ί ΰ ϊ ϋ ό ύ ώ ϙ ϙ Ϛ ϛ F ϝ ϟ ϟ Ϡ ϡ

MakeUppercase:
΄ ῾ ͵ ; ΅ Ά · Έ Ή Ί Ό Ύ Ώ ΐ Ϊ Ϋ Ά Έ Ή Ί ΰ Ϊ Ϋ Ό Ύ Ώ ϘϘ Ϛ Ϛ F F ϟ ϟ Ϡ Ϡ

MakeLowercase:
΄ ῾ ͵ ; ΄ ΅ ά · έ ή ί ό ύ ώ ΐ ϊ ϋ ά έ ή ί ΰ ϊ ϋ ό ύ ώ ϙ ϙ ϛ ϛ ϝ ϝ ϟ ϟ ϡ ϡ

1.4 aliases

Aliases are defined in the included file greek-fontenc.def.

Names matching mathematical variant symbols:

[3] In LGR, there is no separate code point for uppercase koppa.
[4] the name "stigma" originally applied to a medieval sigma-tau ligature, whose shape was confusingly similar to the cursive digamma

ε \textvarepsilon = ε \textepsilon
φ \textvarphi = φ \textphi
ς \textvarsigma = ς \textfinalsigma

Compatibility aliases for hyperref's puenc.def:

μ \textmugreek = μ \textmu
ϟ \textkoppagreek = ϟ \textkoppa
Ϟ \textKoppagreek = Ϟ \textKoppa
Ϛ \textStigmagreek = Ϛ \textStigma
ϛ \textstigmagreek = ϛ \textstigma
Ϡ \textSampigreek = Ϡ \textSampi
ϡ \textsampigreek = ϡ \textsampi
ϝ \textdigammagreek = ϝ \textdigamma
Ϝ \textDigammagreek = Ϝ \textDigamma
ʹ \textnumeralsigngreek = ʹ \textdexiakeraia
͵ \textnumeralsignlowergreek = ͵ \textaristerikeraia

Two Unicode code points and names for one character:

ʹ \accoxia = ʹ \acctonos
᾽ \acckoronis = ᾽ \accpsili

1.5 symbol variants

Mathematical notation distinguishes variant shapes for pi ($\pi|\varpi$), rho ($\rho|\varrho$), theta ($\theta|\vartheta$), beta ($\beta|\daleth$), and kappa ($\kappa|\varkappa$) (characters for the last two variant symbols are not included in TeX's standard math fonts). These variations have no semantic significance in Greek text and are not given code-points in the LGR encoding. Greek text fonts use the shape variants interchangeably.

2 Diacritics

Capital Greek letters have breathings and accents (except dialytika) to the left (instead of above) and drop them if text is set in UPPERCASE.[5] This is implemented for all combinations that are used in Greek texts (i.e. for which precomposed Unicode character exist), but not for, e.g., $\tilde{\Omega}$). Since 2022, LaTeX's `\MakeUppercase` uses the Greek upcasing rules only if the text language (locale) is set to "greek" with Babel (i.e. not in this document). See char-list.pdf for a comprehensive example using the Greek language.

Different conventions exist for the treatment of the sub-iota with uppercase letters. The CB-Fonts use a capital Iota "index" (A_I, H_I, Ω_I).

LaTeX standard accents[6] (Latin, Greek, Greek Capitals ↦ UPPERCASE)

[5] The word ᾽Η ("or", monotonic Ή), is an exception to this rule because of the need to distinguish it from the nominative feminine article Η.

[6] The ogonek (*little hook*) accent ˛ (\k) does not work in LGR.

à á ã ä â ā ă ȧ å ǎ ǎ a̲ ą̲ ạ ą ↦ À Á Ã Ä Â Ā Ă Ȧ Å Ǎ Ǎ A̲ Ą̲ Ạ Ą

ὰ ά ᾶ ϊ̈ ᾶ ᾱ ᾰ ά̇ ᾰ̊ ᾰ̌ ᾶ ᾱ̲ ᾱ̨ ᾱ̣ ↦ ᾺΑ ΆΑ Ã Ä Â Ā Ă Ȧ Å Ǎ Ǎ A̲ Ą̲ Ạ

ᾺΆÃÄ ÂĀĂȦÅǍǍ A̲ Ą̲ Ạ ↦ ᾺΆÃÄ ÂĀĂȦÅǍǍ A̲ Ą̲ Ạ

Additional Greek diacritics (Greek, Greek Capitals[7] ↦ UPPERCASE)

ά ὲ ῖ ΐ ῗ ῆ ő ŏ ŭ ῶ ᾱ̣ ↦ ΄Α ΄Ε ΅Ι Ί Ὶ Ϊ͂ ΅Η ΅Ο ΅Ο ΅Υ ΅Ω Aɪ

΄Α ΄Ε ΅Ι ΅Η ΅Ο ΅Ο ΅Υ ΅Ω Aɪ ↦ ΄Α ΄Ε ΅Ι ΅Η ΅Ο ΅Ο ΅Υ ΅Ω Aɪ

Input variants and their conversion with `\MakeUppercase`:[8]

ᾶ ᾶ ᾶ, ᾶ ᾶ ᾶ ᾶ ᾶ, ῆ ῆ ῆ ῆ ῆ ῆ̧, ῆ ῆ, ῖ ῖ, ῖ ῖ, ῖ ῖ ῖ →
῞Α ῞Α ῞Α, ῞Α ῞Α ῞Α ῞Α ῞Α, ῞Η ῞Η ῞Η ῞Η ῞Η ῞Η, ῞Η ῞Η, ῞Ι ῞Ι, ῞Ι ῞Ι, ΐ ΐ

ῦ ῦ, ῦ ῦ, ῦ ῦ, ὼ, ώ, ῶ ῶ, ῶ ῶ, ᾱ̣ ᾱ̣ →
῟Υ ΰ, ῟Υ ῟Υ, ΰ ΰ, ΅Ω, ΄Ω, ῞Ω ῞Ω, ῞Ω ῞Ω, Aɪ Aɪ

ά ὲ ΐ ῆ ῆ ő ŭ ῶ ΄Α ΅Ε Ϊ ΅Η ΅Ο ῟Υ ῞Ω, ᾱ̣ ᾱ̣ →
΄Α ΅Ε Ί ΅Η Η ΅Ο ῟Υ ῞Ω ΄Α ΅Ε Ϊ ΅Η ΅Ο ῟Υ ῞Ω, Aɪ Aɪ

῞Α ῞Α ῞Α ῞Α → ῞Α ῞Α ῞Α ῞Α.

Input variants and their conversion with MakeLowercase:

῞Α ῞Α ῞Α, ῞Α ῞Α ῞Α ῞Α, ῞Η ῞Η ῞Η ῞Η ῞Η, ῞Η ῞Η, ῞Ι ῞Ι, ῞Ι ῞Ι
ᾶ ᾶ ᾶ, ᾶ ᾶ ᾶ ᾶ, ῆ ῆ ῆ ῆ ῆ, ῆ ῆ, ῖ ῖ, ῖ ῖ

῟Υ ῟Υ, ΰ, ΰ, ῞Ω ῞Ω, ῞Ω ῞Ω, Aɪ Aɪ Aɪ.
ῦ ῦ, ὼ, ώ, ῶ ῶ, ῶ ῶ, ᾱ̣ ᾱ̣ ᾱ̣.

ά ὲ ΐ ῆ ő ŭ ῶ ΄Α ΅Ε Ϊ ΅Η ΅Ο ῟Υ ῞Ω Aɪ Aɪ Aɪ
ά ὲ ΐ ῆ ő ŭ ῶ ά ὲ ϊ ῆ ő ŭ ῶ ᾱ̣ ᾱ̣ ᾱ̣

῞Α ῞Α ῞Α ῞Α ↦ ᾶ ᾶ ᾶ ᾶ

The tilde character can be used in combined accents. However, in documents not defining the Babel language *greek*, the tilde character produces a no-break space if converted with `\MakeUppercase` or `\MakeLowercase`:

combined accent with tilde character:
ῖ ῖ ŭ ŭ ŭ ↦ Ῐ ΅Ι ῟Υ ῞Υ ῟Υ
Ῐ ΅Ι ῟Υ ῞Υ ῟Υ ↦ ῖ ῖ ŭ ŭ ŭ

combined accent with tilde-accent macro:
ῖ ŭ ↦ Ῐ ῟Υ
Ῐ ῟Υ ↦ ῖ ŭ

[7] The dialytika is not used on initial letters.

[8] With the 2022 `\MakeUppercase` implementation, there are issues like wrongly placed diacritics or leading to a NBSP with letters/diacritics that are never used in an initial position because this document does not set the text language to "greek". However, there should not be any compiling errors.

Accents input via the Latin transliteration are not dropped with MakeUppercase, unless Babel is loaded and the current language is Greek (because the required local re-definitions are done in `greek.ldf` from the *babel-greek* package).

ά ϊ ὰ ἀ ἁ ᾳ ↦ Ά Ϊ Ὰ Ἀ Ἁ A_{I}

Accent macros can start with `\a` instead of `\` when the short form is redefined, e. g. inside a *tabbing* environment. This also works for the locally defined dasia and psili shortcuts `\<` and `\>`:

COL1	COL2	COL3	COL4
COL1		COL3	
Viele	Grüße	ἁ	ὠ

Combinations with named accents: ᾰ ᾱ ᾶ.