

Test LGR font encoding definitions

Günter Milde

February 28, 2020

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:

Τῖ φῆις; Ἰδὼν ἐνθ' ἔδε παῖδ' ἐλευθέραν
τὰς πλησίον Νύμφας στεφανοῦσαν, Σώστρατε,
ἔρῳν ἀπῆλθες εὐθύς;

1 Symbols

See the source file [lgrenc-test.tex](#) for the macros used to access the symbols.

1.1 Generic text symbols

Latin: + - = < > - — { [()] } \ | ‰ ‰ ‰ □

LGR: + - = < > - — { [()] } \ | ‰ ‰ (Per-mille symbol is missing in LGR.)

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

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

Ligature break up: AY fi AY ι ↦ AY fi AY ι

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

Symbols for SI-units: 5 μm, 5 kΩ; 5 μm, 5 kΩ; but never italic: 5 μm

Letter schwa and Euro symbol: ə, €

Some symbol definitions expect a Latin font. *babel-greek* redefines them with `\latintext`, however this macro is not guaranteed to be defined, so it should not be used in a font encoding definition file. The `textcomp.sty` package provides copyright, registered, and trademark symbols for use with any font encodings.

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

(Like any other Latin character, the “sharp s” (ß) is not safe to use when LGR is the active font encoding.)

Latin: © ® ™ SS (uppercase of ß).

LGR (with textcomp): © ® ™ ΣΣ (uppercase of ß).

1.2 Greek alphabet

Greek letters via Latin transcription and LIGR macros:

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 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 τ υ φ χ ψ ω

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

σ textsigma

ς textfinalsigma or textvarsigma

In the Latin transcription, the letter ‘s’ stands for `\textautosigma` which automatically chooses the glyph according to the position.

1.3 Additional Greek symbols

Ϝ textkoppa

ϝ textqoppa (archaic koppa)

Ϟ textQoppa (archaic Koppa)

ϟ textstigma

Ϡ textvarstigma

ϡ textStigma (Sigma-Tau-Ligature in CB-fonts)²

Ϣ textsampi

ϣ textSampi

Ϟ textdigamma

ϟ textDigamma

Ϡ textdexiakeraia | ϡ textnumeralsigngreek (Dexia keraia)

Ϣ textaristerikeraia | ϣ textnumeralsignlowergreek (Aristeri keraia)

1.4 symbol variants

Mathematical notation distinguishes variant shapes for pi ($\pi|\varpi$), rho ($\rho|\varrho$), theta ($\theta|\vartheta$), beta, and kappa (characters for the latter two symbols are not included in TeX’s math fonts). These variations have no syntactic meaning in Greek text

²the name “stigma” originally applied to a medieval sigma-tau ligature, whose shape was confusingly similar to the cursive digamma

and are not given code-points in the LGR encoding. Greek text fonts use the shape variants interchangeably.

2 Diacritics

Capital Greek letters have Greek diacritics (except the dialytika and sub-iota) to the left (instead of above) and drop them if text is set in UPPERCASE. This is implemented for all combinations that are used in Greek texts (i.e. for which pre-composed Unicode character exist), but not for, e.g., $\tilde{\Omega}$).

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 accents³ (Latin, Greek, Greek Capitals \mapsto UPPERCASE)

[illegible]Additional Greek diacritics (Greek, Greek Capitals⁴ \mapsto UPPERCASE)
$$\begin{array}{l} \alpha \varepsilon \imath \check{\imath} \check{\imath} \check{\imath} \check{\imath} \check{\eta} \check{o} \check{o} \check{u} \check{\omega} \varphi \mapsto A E I \check{I} \check{I} \check{I} H O O \Upsilon \Omega A_i \\ A E \imath I H O O \check{O} \check{\Upsilon} \check{\Omega} A_i \mapsto A E I H O O \Upsilon \Omega A_i \end{array}$$

Input variants and their conversion with MakeUppercase:

$\alpha \, \tilde{\alpha} \, \tilde{\alpha}, \, \tilde{\alpha} \, \tilde{\alpha} \, \tilde{\alpha} \, \tilde{\alpha} \, \tilde{\alpha}, \, \tilde{\eta} \, \tilde{\eta} \, \tilde{\eta} \, \tilde{\eta} \, \tilde{\eta} \, \tilde{\eta}, \, \tilde{\eta} \, \tilde{\eta}, \, \tilde{\iota} \, \tilde{\iota}, \, \tilde{\iota} \, \tilde{\iota}, \, \tilde{\iota} \, \tilde{\iota} \, \tilde{\iota},$
 $\tilde{\upsilon} \, \tilde{\upsilon}, \, \tilde{\upsilon} \, \tilde{\upsilon}, \, \tilde{\omega}, \, \tilde{\omega}, \, \tilde{\omega} \, \tilde{\omega}, \, \tilde{\omega} \, \tilde{\omega}, \, \varphi \, \varphi$
 $A \, A \, A, \, A \, A \, \, A \, A \, A, \, H \, H \, \, H \, H \, H \, H, \, H \, H, \, I \, I, \, I \, I, \, \tilde{I} \, \tilde{I}$
 $\Upsilon \, \Upsilon, \, \Upsilon \, \Upsilon, \, \tilde{\Upsilon} \, \tilde{\Upsilon}, \, \Omega, \, \Omega, \, \Omega \, \Omega, \, \Omega \, \Omega, \, A_i \, A_i$
 $\acute{\alpha} \, \acute{\epsilon} \, \acute{\eta} \, \acute{\theta} \, \acute{\upsilon} \, \acute{\omega} \, A \, E \, \tilde{I} \, H \, O \, T \, \Omega, \, \varphi \, \varphi$
 $A \, E \, \tilde{I} \, H \, O \, \Upsilon \, \Omega \, A \, E \, \tilde{I} \, H \, O \, \Upsilon \, \Omega, \, A_i \, A_i$
 $\textit{A} \, \textit{A} \, \textit{A} \, \textit{A} \mapsto A \, A \, A \, A.$

Input variants and their conversion with MakeLowercase:

ʼA ʼA ʼA, ʼA ʼA ʼA ʼA, ʼH ʼH ʼH ʼH ʼH, ʼH ʼH, ʼI ʼI, ʼI ʼI
 ǎ ǎ ǎ, ǎ ǎ ǎ ǎ, ǎ ǎ ǎ ǎ ǎ, ǎ ǎ, ǎ ǎ, ǎ ǎ
 ʼT ʼT, ʼΩ, ʼΩ, ʼΩ ʼΩ, ʼΩ ʼΩ, A_I A_I A_I.
 ʼU ʼU, ʼω, ʼω, ʼω ʼω, ʼω ʼω, ʼφ ʼφ ʼφ.
 ǎ ǎ ǎ ǎ ǎ ʼA ʼE ʼI ʼH ʼO ʼT ʼΩ A_I A_I A_I
 ǎ
 ʼA ʼA ʼA ʼA \mapsto ǎ ǎ ǎ ǎ

³The ogonek (*little hook*) accent, ($\backslash k$) is not defined in LGR.

⁴The dialytika is not used on Initial letters.

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

combined accent with tilde character:

$$\begin{array}{l} \tilde{I} \tilde{I} \tilde{U} \tilde{U} \tilde{U} \mapsto \cdot I \ I \cdot Y \ Y \ Y \\ \tilde{I} \tilde{I} \tilde{Y} \tilde{Y} \tilde{Y} \mapsto \cdot I' \ I' \cdot U' \ U' \ U' \end{array}$$

combined accent with tilde-accent macro:

$$\begin{array}{l} \tilde{\mathfrak{I}} \tilde{\mathfrak{U}} \mapsto \mathfrak{I} \mathfrak{Y} \\ \mathfrak{I} \mathfrak{Y} \mapsto \tilde{\mathfrak{I}} \tilde{\mathfrak{U}} \end{array}$$

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 of the `uccode` are done in `greek.ldf` from the *babel-greek* package).

$$\acute{\alpha} \ddot{\alpha} \grave{\alpha} \tilde{\alpha} \check{\alpha} \alpha \mapsto 'A \ddot{I} `A 'A 'A A,$$

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	Griße	α	ω

Combinations with named accents: $\check{\alpha}$ $\grave{\alpha}$ $\tilde{\alpha}$.

The dialytika must be kept in UPPERCASE, e. g.

$$\mu\alpha\acute{\iota}\sigma\tau\rho\omicron\varsigma \mapsto \text{MAÏΣΤΡΟΣ or } \epsilon\upsilon\zeta\omega\acute{\iota}\alpha \mapsto \text{EΥΖΩΪΑ.}$$

This is implemented for all input variants of diacritics with dialytika:

$$\text{ı ı ı ı ı ı ı ı} \mapsto \text{İ İ İ İ İ İ İ İ},$$

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

ἄυλος \mapsto AÿΛΟΣ, ἄυλος \mapsto AÿΛΟΣ, μᾰίνα \mapsto MAÏNA, κείκ, \mapsto KEÏK
 ἄυπνία \mapsto AÿΠΝΙΑ

Test the auto-hiatus feature for side-effects:

A B (must keep space after A).

Kerning (see the input):

```

AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AI AY PA OA TA ΔΥ [
˘ AO AΨ AĬ AŸ PA OA ŸA ΔΥ [
˘ ÄO ÄΨ ÄI ÄŸ PÄ ÖA TÄ ΔŸ [
˘ ÄO ÄΨ ÄI ÄŸ PÄ ÖA TÄ ΔŸ [

```

Rows 3 . . . 7: Look-ahead (to check for a hiatus) breaks kerning before A with Tonos or Psili.

Rows 15 and 16: Like in any font encoding, there is no kerning for non-defined accent-letter-combinations (dialytica on A O Δ).

Downcasing should keep diacritics (of course, it cannot regenerate “manually” dropped ones): ‘A Ĭ Ÿ ˘A ↦ á ĭ ü ž