

Test LGR font encoding definitions

Günter Milde

February 5, 2016

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 φήεις; Ἰδὼν ἐνθῆδε παῖδ' ἐλευθέραν
τὰς πλησίον Νύμφας στεφανοῦσαν, Σώστρατε,
ἔρωϊν ἀπῆλθεες εὐθύς;

1 Symbols

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

1.1 Generic text symbols

Latin: + - = < > - - { [()] } \ | %₀ %₀₀ □

LGR: + - = < > - - { [()] } \ | %₀ ‰ (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: \hat{a} $\hat{\alpha}$ $\hat{\iota}$ $\sim a$ $\sim \alpha$ $\sim \iota$ \check{a} $\check{\alpha}$ $\check{\iota}$ \bar{a} $\bar{\alpha}$ $\bar{\iota}$ \grave{a} $\grave{\alpha}$ $\grave{\iota}$ \acute{a} $\acute{\alpha}$ $\acute{\iota}$ \grave{a} $\acute{\alpha}$ $\acute{\iota}$

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

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

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. (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 LICR macros:

```
A B Γ Δ E Z H Θ I K Λ M N Ξ O Π P Σ T Υ Φ X 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
A B Γ Δ E Z H Θ I K Λ M N Ξ O Π P Σ T Υ Φ X 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
Ϟ textqoppa (archaic koppa)
ϙ textQoppa (archaic Koppa)
Ϛ textstigma
ϛ textvarstigma
Ϝ textStigma (Sigma-Tau-Ligature in CB-fonts)2
ϝ textsampi
Ϟ textSampi
ϟ textdigamma
Ϡ textDigamma
' textdexiakeraia | ' textnumeralsigngreek (Dexia keraia)
, textaristerikeraia | , textnumeralsignlowergreek (Aristeri keraia)
```

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

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 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)

$\grave{\alpha} \acute{\alpha} \tilde{\alpha} \grave{\alpha} \bar{\alpha} \acute{\alpha} \grave{\alpha} \grave{\alpha} \grave{\alpha} \grave{\alpha} \grave{\alpha} \mapsto \grave{A} \acute{A} \tilde{A} \grave{A} \bar{A} \acute{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A}$
 $\grave{\alpha} \acute{\alpha} \tilde{\alpha} \grave{\alpha} \bar{\alpha} \acute{\alpha} \grave{\alpha} \grave{\alpha} \grave{\alpha} \grave{\alpha} \acute{\alpha} \acute{\alpha} \mapsto A A A \tilde{A} \bar{A} \acute{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A}$
 $\grave{A} \acute{A} \tilde{A} \grave{A} \bar{A} \acute{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \mapsto AAA \tilde{A} \bar{A} \acute{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A} \grave{A}$

Additional Greek diacritics (Greek, Greek Capitals⁴ \mapsto UPPERCASE)

$\acute{\alpha} \acute{\epsilon} \acute{\iota} \acute{\iota} \acute{\iota} \acute{\eta} \acute{\delta} \acute{\delta} \acute{\delta} \acute{\delta} \acute{\delta} \acute{\delta} \mapsto A E I \tilde{I} \tilde{I} \tilde{I} H O O \Upsilon \Omega A_I$
 $\acute{A} \acute{E} \acute{I} \acute{H} \acute{O} \acute{O} \acute{O} \acute{\Upsilon} \acute{\Omega} A_I \mapsto A E I H O O \Upsilon \Omega A_I$

Input variants and their conversion with `MakeUppercase`:

$\acute{\alpha} \acute{\alpha} \acute{\alpha}, \acute{\alpha} \acute{\alpha} \acute{\alpha} \acute{\alpha} \acute{\alpha}, \acute{\eta} \acute{\eta} \acute{\eta} \acute{\eta} \acute{\eta}, \acute{\eta} \acute{\eta}, \acute{\iota} \acute{\iota}, \acute{\iota} \acute{\iota}, \acute{\iota} \acute{\iota} \acute{\iota},$
 $\acute{\upsilon} \acute{\upsilon}, \acute{\upsilon} \acute{\upsilon}, \acute{\omega} \acute{\omega}, \acute{\omega} \acute{\omega}, \acute{\omega} \acute{\omega}, \acute{\omega} \acute{\omega}, \acute{\omega} \acute{\omega} A_I A_I.$

$A A A, A A A A A, H H H H H H_I, H H, I I, I I, \tilde{I} \tilde{I}$
 $\Upsilon \Upsilon, \Upsilon \Upsilon, \Upsilon \Upsilon, \Omega, \Omega, \Omega \Omega, \Omega \Omega, A_I A_I A_I A_I.$

$\acute{\alpha} \acute{\epsilon} \acute{\iota} \acute{\eta} \acute{\delta} \acute{\delta} \acute{\delta} \acute{A} \acute{E} \acute{I} \acute{H} \acute{O} \acute{O} \acute{\Upsilon} \acute{\Omega}$
 $A E \tilde{I} H O \Upsilon \Omega A E \tilde{I} H O \Upsilon \Omega$

$\acute{A} \acute{E} \acute{I} \acute{A} \acute{A} \mapsto A A A A.$

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`:

³The ogonek (*little hook*) accent $\acute{\cdot}$ ($\backslash k$) is not defined in LGR.

⁴The dialytika is not used on Initial letters.

combined accent with tilde character: $\text{ĩ ĩ ũ ũ} \mapsto \text{" I I" \Upsilon \Upsilon \Upsilon}$

combined accent with tilde-accent macro: $\text{ĩ ũ} \mapsto \text{Ĩ Ŷ}$

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).

$\acute{\alpha} \grave{\alpha} \acute{\alpha} \acute{\alpha} \alpha \mapsto \text{'A \u00c7 \u00c7 \u00c7 \u00c7 A}$

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 Gr\u00fc\u00dfe \u00c7 \u00f6
```

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

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

$\mu\acute{\alpha}\iota\sigma\tau\rho\omicron\varsigma \mapsto \text{MA\u00cd\Sigma\Upsilon\Upsilon\Upsilon\Upsilon}$ or $\epsilon\upsilon\zeta\omega\acute{\iota}\alpha \mapsto \text{E\u00c7Z\u00d9\u00c7A}$.

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

$\text{ĩ ĩ ĩ ũ ũ ũ} \mapsto \text{Ĩ Ĩ Ĩ \u00c7 \u00c7 \u00c7}$,

Tonos and dasia mark a *hiatus* (break-up of a diphthong) if placed on the first vowel of a diphthong ($\acute{\alpha}\iota$, $\acute{\alpha}\upsilon$, $\acute{\epsilon}\iota$). A dialytika must be placed on the second vowel if they are dropped: (A\u00cd, A\u00c7, E\u00cd).

$\acute{\alpha}\upsilon\lambda\omicron\varsigma \mapsto \text{A\u00c7\u00c7\u00c7\u00c7}$, $\acute{\alpha}\upsilon\lambda\omicron\varsigma \mapsto \text{A\u00c7\u00c7\u00c7\u00c7}$, $\mu\acute{\alpha}\iota\nu\alpha \mapsto \text{MA\u00cdNA}$, $\kappa\acute{\epsilon}\iota\kappa \mapsto \text{KE\u00cdK}$
 $\acute{\alpha}\upsilon\pi\nu\acute{\iota}\alpha \mapsto \text{A\u00c7\u00c7\u00c7\u00c7}$

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 AĪ AŸ PA OA TA ΔΥ [
˚AO AΨ AĪ AŸ PA OA TA ΔΥ [
´AO AΨ AĪ AŸ PA OA TA ΔΥ [
ˆAO AΨ AĪ AŸ PA OA TA ΔΥ [
˜AO AΨ AĪ AŸ PA OA TA ΔΥ [
`AO AΨ AI AY PA OA TA ΔΥ [
ˆAO 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 AI AY PA OA TA ΔΥ [
˚AO AΨ AĪ 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 Ī Ÿ ˚A ↦ á ĩ ü ž