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;

1 Symbols

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

1.1 Generic text symbols

Latin: + - = < > – — [()] { } \ | ‰ ‰₀ ␣

LGR: + - = – — [()]

> < \textless
> > \textgreater
> { \textbraceleft
> } \textbraceright
> \ \textbackslash
> | \textbar
> ‰ \textperthousand (Per-mille symbol is missing in LGR.)
> ␣ \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.

Ligature break up with \textcompwordmark: ΑΥ fi ΑΥ ï ↦ ΑΥ fi ΑΥ ï

---

[1] Single quotes need special attention to prevent conversion to accents. Test the input conventions: 'α' 'α' 'α' 'α' but not ά έ ï'

Spacing accent chars: ᾶa ᾶα ῖι ~a ~α ~ι ˘a ˘α ˘ι ¯a ¯α ¯ι ¨a ¨α ¨ι ´a ´α ´ι `a `α `ι

Letter schwa and Euro symbol: ǝ \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 with \latintext, however this cannot be done in a font encoding definition file.

Beware that "#&';<>?@ becomes ¨˜´·´``;˜.

The *textcomp* package provides pre-composed coyright ©, registered ® and trademark ™ symbols that work in all font encodings. In LGR (with textcomp), they come out as: © \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, ΑΝΤΙΣΤΑΣΗ = 5 μΩ, αντίσταση = 5 μΩ.

1.2   Greek alphabet

Greek letters via Latin transcription and LICR 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 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 transcription, 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)[2]
    ϙ \textqoppa (archaic koppa)
    Ϙ \textQoppa (archaic Koppa)
    ϛ \textstigma
    ϛ \textvarstigma

---

[2]Modern typographical practice normally does not observe a contrast between uppercase and lowercase forms for numeric koppa. In LGR, there is no separate code point for Koppa.

Ϛ \textStigma (Sigma-Tau-Ligature in CB-fonts)[3]
ϡ \textsampi
Ϡ \textSampi
ϝ \textdigamma
F \textDigamma
ʹ \textdexiakeraia (Dexia keraia)
͵ \textaristerikeraia (Aristeri keraia)

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

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

MakeUppercase:
ʹ ͵ · ; ¨ Α · Ε Η Ι Ο Υ Ω Ϊ Ϊ Ϋ Α Ε Η Ι Ϋ Ϊ Ϋ Ο Υ Ω Ϟ Ϟ Ϛ Ϛ F F ϟ ϟ Ϡ Ϡ

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

1.4 aliases

Aliases are defined in the included file [greek-fontenc.def](#).

Names matching mathematical variant symbols:

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
F \textDigammagreek = F \textDigamma
ʹ \textnumeralsigngreek = ʹ \textdexiakeraia
͵ \textnumeralsignlowergreek = ͵ \textaristerikeraia

Two Unicode code points and names for one character:

ʹ \accoxia = ʹ \acctonos
ʼ \acckoronis = ʼ \accpsili

---

[3] the name "stigma" originally applied to a medieval sigma-tau ligature, whose shape was confusingly similar to the cursive digamma

1.5 symbol variants

Mathematical notation distinguishes variant shapes for pi ($\pi|\varpi$), rho ($\rho|\varrho$), theta ($\theta|\vartheta$), beta, and kappa (characters for the last two symbols are not included in TeX's standard 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" (Aᵢ, Hᵢ, Ωᵢ).

LaTeX standard accents[4] (Latin, Greek, Greek Capitals ↦ UPPERCASE)

à á ã ä â ā ā ắ å å ǎ ǎ a̱ ą̊ a̧ ą ↦ À Á Ã Ä Â Ā Ā Ắ Å Å Ǎ Ǎ A̱ Ą A̧ Ą

ὰ ά ᾶ ᾶ ᾶ ᾱ ᾱ ᾱ ᾶ ᾶ ᾶ ᾶ α̱ ᾳ ᾳ ↦ A A A Ä Â Ā Ắ Å Å Ǎ Ǎ A̱ Ą A̧

ʼΑʹΑΑ̃Ä ẪĀ̃Ắ̃ÅÅ̃Ǎ̃ A̱ Ą A̧ ↦ AAAÄ ÂĀÁÅÅ̃ǍǍ A̱ Ą A̧

Additional Greek diacritics (Greek, Greek Capitals[5] ↦ UPPERCASE)

ἀ ἐ ῖ ῗ ῗ ῗ ῇ ὄ ὄ ὔ ὤ ᾳ ↦ A E I Ï Ï H O O Υ Ω Aᵢ

ʼΑ Ἐ Ἰ Ἦ Ὄ Ὂ Ὓ Ὢ Aᵢ ↦ A E I H O O Υ Ω Aᵢ

Input variants and their conversion with MakeUppercase:

ᾰ ᾰ ᾰ, ᾰ ᾰ ᾰ ᾰ ᾰ, ῄ ῄ ῄ ῄ ῄ ῄ, ῄ ῄ, ῗ ῗ, ῗ ῗ, ῗ ῗ ῗ,
ῠ ῠ, ῠ ῠ, ῠ ῠ, ὼ, ώ, ὤ ὤ, ὤ ὤ, ᾳ ᾳ

A A A, A A A A A, H H H H H Hᵢ, H H, I I, I I, Ï Ï
Υ Υ, Υ Υ, Ϋ Ϋ, Ω, Ω, Ω Ω, Ω Ω, Aᵢ Aᵢ

ἀ ἐ ῐ ῄ ῄ ὄ ὔ ὤ Ἀ Ἐ Ϊ Ἦ Ὄ Ὓ Ὢ, ᾳ ᾳ
A E Ï H H O Υ Ω A E Ï H O Υ Ω, Aᵢ Aᵢ

Ἂ Ἂ Ἂ Ἂ ↦ A A A A.

Input variants and their conversion with MakeLowercase:

Ἂ Ἂ Ἂ, Ἂ Ἂ Ἂ Ἂ, Ἦ Ἦ Ἦ Ἦ Ἦᵢ, Ἦ Ἦ, Ἲ Ἲ, Ἲ Ἲ
ᾰ ᾰ ᾰ, ᾰ ᾰ ᾰ ᾰ, ῄ ῄ ῄ ῄ ῄ, ῄ ῄ, ῗ ῗ, ῗ ῗ

---

[4]The ogonek (*little hook*) accent ˛ (\k) is not defined in LGR.
[5]The dialytika is not used on Initial letters.

῏Υ ῎Υ, Ὠ, Ὠ, ῎Ω ῞Ω, ῏Ω ῏Ω, Αᵢ Αᵢ Αᵢ.

ὒ ὒ, ὼ, ὡ, ὦ ὧ, ὦ ὧ, ᾳ ᾳ ᾳ.

ἀ ἐ ἴ ἤ ὄ ὒ ὦ Ἁ Ἐ Ϊ Ἡ Ὅ ῎Υ ῏Ω Αᵢ Αᵢ Αᵢ

ἀ ἐ ἴ ἤ ὄ ὒ ὦ ἀ ἐ ϊ ἤ ὄ ὒ ὦ ᾳ ᾳ ᾳ

῎Α ῎Α ῎Α ῎Α ↦ ἄ ἄ ἄ ἄ

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:
> ῗ ῗ ὒ ὒ ὒ ↦ ¨ Ι Ι ¨ Υ Υ Υ
> Ϊ ᾿Ι ῧ ῎Υ ῧ ↦ ¨ ι ´ ι ¨ υ ΄ υ ὒ

> combined accent with tilde-accent macro:
> ῗ ὒ ↦ Ϊ Ϋ
> Ϊ Ϋ ↦ ῗ ὒ

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).

> ά ϊ ὰ ἀ ἀ ᾳ ↦ Ά Ϊ Ὰ Ἀ Ἀ Αᵢ

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 | ἀ | ὡ |

Combinations with named accents: ἄ ἃ ἆ.

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

> μαΐστρος ↦ ΜΑΪΣΤΡΟΣ or εὐζωΐα ↦ ΕΥΖΩΪΑ.

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

> ϊ ΐ ΐ ῗ ϋ ΰ ὒ ῧ ↦ Ϊ Ϊ Ϊ Ϊ Ϋ Ϋ Ϋ Ϋ,

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: (ΑΪ, ΑΫ, ΕΪ).

> άυλος ↦ ΑΫΛΟΣ, ἄυλος ↦ ΑΫΛΟΣ, μάινα ↦ ΜΑΪΝΑ, κέικ, ↦ ΚΕΪΚ
> ἀυπνία ↦ ΑΫΠΝΙΑ

Test the auto-hiatus feature for side-effects:

A B (must keep space after A).

Kerning (see the input):

```
 AO AΨ AI AΥ PA OA ΥA ΔΥ [
˘ AO AΨ AI AΥ PA OA ΥA ΔΥ [
᾿ AO AΨ AÏ AῨ PA OA ΥA ΔΥ [
῎ AO AΨ AÏ AῨ PA OA ΥA ΔΥ [
´ AO AΨ AÏ AῨ PA OA ΥA ΔΥ [
῀ AO AΨ AÏ AῨ PA OA ΥA ΔΥ [
῎ AO AΨ AÏ AῨ PA OA ΥA ΔΥ [
` AO AΨ AI AΥ PA OA ΥA ΔΥ [
῾ AO AΨ AI AΥ PA OA ΥA ΔΥ [
˜ AO AΨ AI AΥ PA OA ΥA ΔΥ [
῀ AO AΨ AI AΥ PA OA ΥA ΔΥ [
῀ AO AΨ AI AΥ PA OA ΥA ΔΥ [
῀ AO AΨ AI AΥ PA OA ΥA ΔΥ [
¨ AO AΨ AÏ AῨ PA OA ῨA ΔΥ [
¨ ÄO ÄΨ ÄI ÄΥ � PA ÖA ΥÄ ΔῨ [
῀ ÄO ÄΨ ÄI ÄΥ ῾PA ÖA ΥÄ ΔῨ [
```

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): Ά Ï Ῠ ῎A ↦ ά ϊ ῠ ᾶ