Test LGR font encoding definitions

Günter Milde

November 29, 2013

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 ϛτεφανοῦσαν, Σώϛτρατε,
> 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: + - = < > – — { [()] } \ | ‰ ‱ ␣

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

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

> Single quotes need special attention to prevent conversion to accents.
> Test the input conventions: ‘α’ ‘α’ ‘α’ ‘α’ but not ὰ΄ έ΄ ΐ΄’
>
> Single guillemots and base-quotes (‹a› „a” ‚a’) are missing in LGR.

Ligature break up: AY fi AΥ ϊ ↦ AY fi AΥ ¨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 m, 5 kΩ; 5 µm, 5 kΩ

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 coyright, registered, and trademark symbols for use with any font encodings. (Like any other Latin character, the "sharp s" (ß) is not save 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

ϟ textkoppa
ϙ textqoppa (archaic koppa)
Ϙ textQoppa (archaic Koppa)
ϛ textstigma
ϗ textvarstigma
Ϛ textStigma (Sigma-Tau-Ligature in CB-fonts)[1]
ϡ textsampi
Ϡ textSampi
ϝ textdigamma
Ϝ textDigamma
ʹ textnumeralsigngreek (Dexia keraia)
͵ textnumeralsignlowergreek (Aristeri keraia)

Mathematical notation sometimes uses variant shapes for pi (ϖ), kappa (*no TeX symbol available*), rho (ϱ), and theta (ϑ). These variations have no syntactic

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

meaning in Greek text and are not given code-points in the LGR encoding. Some Greek text fonts use variant shapes in place of the "regular" ones.

2 Diacritics

Capital Greek letters have Greek diacritics (except the dialytika) to the left (instead of above) and drop them in UPPERCASE. This is implemented for all combinations for which a pre-composed Unicode character exists (but not, e.g., Ã).

Different conventions exist for the treatment of the sub-iota with uppercase letters. The CB-Fonts use a capital Iota "index".

LaTeX standard accents (Latin, Greek, Greek Capitals) ↦ UPPERCASE

à á ã ä â ā a̋ å ă ǎ a̱ a ạ ą ↦ À Á Ã Ä Â Ā A̋ Å Ă Ǎ A̱ A Ạ Ą

ὰ ά α̃ α̈ α̂ ᾱ α̋ α̇ α̊ ᾰ α̌ α̱ α̧ α̣ ↦ A A A Ä Â Ā A̋ Ȧ Å Ă Ǎ A̱ A̧ Ạ

᾿A ´A Ã Ä Â Ā A̋ Ȧ Å Ă Ǎ A̱ A̧ Ạ ↦ A A A Ä Â Ā A̋ Ȧ Å Ă Ǎ A̱ A̧ Ạ

The comma-below (`\k`) is not defined in LGR.

Additional Greek diacritics

ἁ ἐ ἳ ΐ ῒ ῗ ᾓ ὅ ὄ ὖ ὧ ↦ A E I Ϊ Ϊ Ϊ H O O Υ Ω

Ἁ Ἐ Ἳ Ϊ́ Ϊ̀ Ϊ͂ Ἣ Ὅ Ὄ Ὗ Ὧ ↦ A E I Ϊ Ϊ Ϊ H O O Υ Ω

Input variants and their conversion with MakeUppercase:

ἆ ἆ ἆ, ἇ ἇ ἇ ἇ ἇ, ἦ ἦ ἦ ἦ ἦ ᾖ, ἧ ἧ, ἶ ἶ, ἷ ἷ, ῗ ῗ ῗ,
ὖ ὖ, ὗ ὗ, ῧ ῧ, ὠ, ὡ, ὦ ὦ, ὧ ὧ.

A A A, A A A A A, H H H H H Hɪ, H H, I I, I I, Ϊ Ϊ ’ I,
Υ Υ, Υ Υ, Ϋ Ϋ, Ω, Ω, Ω Ω, Ω Ω.

ἁ ἐ ἴ ᾓ ὅ ὖ ὧ Ἁ Ἐ Ϊ Ἣ Ὅ Ὗ Ὧ
A E Ϊ H O Υ Ω. A E Ϊ H O Υ Ω

Ἅ Ἅ Ἅ Ἅ ↦ A A A A.

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

ῗ ἶ ῧ ὖ ὒ ↦ ’ I I ’ Υ Υ Υ
Use the tilde-accent macro: Ϊ Ϋ.

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 new-defined Dasia and Psili shortcuts: ἁ ὠ.

Combinations with named accents: ᾅ ᾓ ᾇ.

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 ῖ ῗ ϋ ΰ ῧ ῢ ↦ 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, ᾰ̓υλος ↦ ΑΫΛΟΣ, μάινα ↦ ΜΑΪΝΑ, κέικ, ↦ ΚΕΪΚ ἀυπνία ↦ ΑΫΠΝΙΑ

Test the auto-hiatus feature for side-effects:

A B: keep space after A.

AO AΨ AI AΥ PA OA ϒA ΔΥ [
̓AO AΨ AI AΥ PA OA ϒA ΔΥ [
̔AO AΨ AΪ AΫ PA OA ϒA ΔΥ [
̽AO AΨ AΪ AΫ PA OA ϒA ΔΥ [
́AO AΨ AΪ AΫ PA OA ϒA ΔΥ [
̓́AO AΨ AΪ AΫ PA OA ϒA ΔΥ [
̽́AO AΨ AΪ AΫ 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 AΪ AΫ PA OA ϒA ΔΥ [
̔ÄO ÄΨ ÄI ÄΥ ̈PA ÖA ϒÄ ΔΫ [
̓̈ÄO ÄΨ ÄI ÄΥ ̈PA ÖA ϒÄ ΔΫ [

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).

The uppercase of the zero-width space at the place of "v" is the Dasia-Oxia accent ̓́, the glyph at the position of "V". It is suppressed for uppercased accents:

̓ ̔ ̽ ́ ̓́ ̽́ ̀ ̓̀ ̃ ̓̃ ̽̃ ̈́ ̓̈ ↦ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́ ̓́

Downcasing should keep diacritics (of course, it cannot regenerate "manually" dropped ones): ΆΪ Ϋ ΄Α ↦ ά ϊ ϋ ᾰ