Greek support for Babel with XeTeX/LuaTeX

Günter Milde

2014/09/02

The babel option «greek» activates the support for the Greek language defined in the file «greek.ldf» (source «greek.dtx»).

Typesetting Greek texts requires a font containing Greek letters. With the XeTeX or LuaTeX engines, the user must ensure that the selected font contains the required glyphs (the default Latin Modern fonts miss them). Examples for suitable fonts are the «Deja Vu», «Linux Libertine», or «Free Serif» OpenType fonts.

1 Language Switch

The declaration `\selectlanguage` switches between languages.

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

The command `\foreignlanguage` sets its second argument in the language specified as first argument. This is intended for short text parts like Βιβλιοθήκη.

2 Font Encoding

Switching to a font encoding supporting the Greek script is possible without switching the Babel language using the declarations `\greekscript` (no switch if the current encoding supports Greek script) or `\greektext` (always switch to LGR). The corresponding macros `\ensuregreek` and `\textgreek` typeset their argument in a Greek-supporting font encoding.

With XeTeX or LuaTeX, declaring the LGR font encoding is optional. The macros `\greektext` and `\textgreek` are only defined, if the LGR font encoding is declared via the `fontenc` package *before* loading babel (see test-unicode-lgr.tex).

Every language switch to `greek` calls the `\extrasgreek` command which in turn calls `\greekscript` to ensure a Greek-supporting font encoding (LGR, EU1, or EU2). (For customization, you can add to or redefine the `\extrasgreek` command.)

The LGR font encoding does not support Latin characters. Therefore, the Babel core defines the declaration `\latintext` and the command `\textlatin` to switch to the T1 or OT1 font encoding or typeset the argument using this encoding. `greek-fontenc` adds a test for EU1 and EU2. At this point, the «latinencoding» is EU1.

Every language switch from `greek` calls the `\noextrasgreek` command which in turn calls `\latintext`. (For customization, you can add to or redefine the `\noextrasgreek` command.)

With the Unicode font encodings EU1 (XeTeX) or EU2 (LuaTeX), Latin characters can be used in Greek text parts and input via the «CB Latin transcription» is not possible.[1]

Φίλων τοῦ TeX (ΕΦΤ) – Friends (Fílwn) of TeX.[2]

3 MakeUppercase, MakeLowercase

Capital Greek letters have diacritics (except the dialytika and sub-iota) to the left (instead of above) and drop them in uppercase, e.g. μαΐστρος ↦ ΜΑΪΣΤΡΟΣ.

Tonos and dasia mark a *hiatus* (break-up of a diphthong) if placed on the first vowel of a diphtong ((άι, άυ, έι, ἄι, ἄυ, ἔι). A dialytika must be placed on the second vowel if they are dropped. This does not work (yet?) with Unicode literals (AI, AY, EI, AI, AY, EI). (See section «hiatus» below for a test with LICR macros.)

The following subsections test MakeUppercase and MakeLowercase with all characters defined in lgrenc.dfu:

3.1 Greek and Coptic

Characters of the Greek and Coptic Unicode Block:

΄΅·΄ ΅Α·ΈΉΊΟΥΩΊΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩΪŸϘϚϜϠ
άέήίϋαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώϼςϝϟϡ

MakeUppercase:

΄΅·΄ ΅Α·ΕΗΙΟΥΩΪΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩΪŸϘϚϜΚ
ΑΕΗΙΫΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΣΤΥΦΧΨΩΪΫΟΥΩΩϘϚϜϟΡ

Letters upcased, diacritics except dialytika dropped. OK.

Ϡ (Sampi) ↦ K (Kappa) and ϡ (sampi) ↦ P (Rho) is a bug in xgreek-fixes.def from the polyglossia package.

MakeLowercase:

΄΅·΄ ΅ά·έάίόύώϊαβγδεζηθικλμνξοπρστυφχψωϊϋόϼςϝϡ
άέήίϋαβγδεζηθικλμνξοπρςστυφχψωϊϋόύώϼςϝϟϡ

The lowercase of Σ is σ (GREEK SMALL LETTER SIGMA).[3] The lowercase of Ϛ (GREEK LETTER STIGMA) is ϛ.

---

[1] The *xunicode* package provides with the `tipa` emulation an example how this could be achieved also for Unicode fonts. Alternatively, LGR encoded fonts can be used (see test-unicode-lgr.tex).

[2] Compare the printout to the similar example in test-greek.pdf.

[3] With LICRs, it is `\textautosigma`.

3.2 Greek extended

Characters of the Greek extended Unicode block:

ἀ ἁ ἂ ἃ ἄ ἅ ἆ ἇ Ἀ Ἁ Ἂ Ἃ Ἄ Ἅ Ἆ Ἇ
ἐ ἑ ἒ ἓ ἔ ἕ Ἐ Ἑ Ἒ Ἓ Ἔ Ἕ
ἠ ἡ ἢ ἣ ἤ ἥ ἦ ἧ Ἠ Ἡ Ἢ Ἣ Ἤ Ἥ Ἦ Ἧ
ἰ ἱ ἲ ἳ ἴ ἵ ἶ ἷ Ἰ Ἱ Ἲ Ἳ Ἴ Ἵ Ἶ Ἷ
ὀ ὁ ὂ ὃ ὄ ὅ Ὀ Ὁ Ὂ Ὃ Ὄ Ὅ
ὐ ὑ ὒ ὓ ὔ ὕ ὖ ὗ Ὑ Ὓ Ὕ Ὗ
ὠ ὡ ὢ ὣ ὤ ὥ ὦ ὧ Ὠ Ὡ Ὢ Ὣ Ὤ Ὥ Ὦ Ὧ
ὰ ά ὲ έ ὴ ή ὶ ί ὸ ό ὺ ύ ὼ ώ
ᾀ ᾁ ᾂ ᾃ ᾄ ᾅ ᾆ ᾇ ᾈ ᾉ ᾊ ᾋ ᾌ ᾍ ᾎ ᾏ
ᾐ ᾑ ᾒ ᾓ ᾔ ᾕ ᾖ ᾗ ᾘ ᾙ ᾚ ᾛ ᾜ ᾝ ᾞ ᾟ
ᾠ ᾡ ᾢ ᾣ ᾤ ᾥ ᾦ ᾧ ᾨ ᾩ ᾪ ᾫ ᾬ ᾭ ᾮ ᾯ
ᾰ ᾱ ᾲ ᾳ ᾴ ᾶ ᾷ Ᾰ Ᾱ Ὰ Ά ᾼ ᾽ ι ᾿
῀ ῁ ῂ ῃ ῄ ῆ ῇ Ὲ Έ Ὴ Ή ῌ ῍ ῎ ῏
ῐ ῑ ῒ ΐ ῖ ῗ Ῐ Ῑ Ὶ Ί ῝ ῞ ῟
ῠ ῡ ῢ ΰ ῤ ῥ ῦ ῧ Ῠ Ῡ Ὺ Ύ Ῥ ῭ ΅ `
ῲ ῳ ῴ ῶ ῷ Ὸ Ό Ὼ Ώ ῼ ´ ῾

MakeUppercase:

A A A A A A A A A A A A A A A A
E E E E E E E E E E E E
H H H H H H H H H H H H H H H H
I I I I I I I I I I I I I I I I
O O O O O O O O O O O O
Y Y Y Y Y Y Y Y Y Y Y Y
Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω Ω
A A E E H H I I O O Y Y Ω Ω
Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι Αι
Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι Ηι
Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι Ωι
Ᾰ Ᾱ Αι Αι Αι A A Αι Ᾰ Ᾱ A A Αι ᾽ I ᾿
῀ ῁ Ηι Ηι Ηι H H Ηι E E H H Ηι ῍ ῎ ῏
Ῐ Ῑ Ϊ I I Ϊ Ῑ I I ῝ ῞ ῟
Ῠ Ῡ Ϋ Ρ Ρ Y Y Ϋ Ῡ Y Y Ρ ῭ ΅ `
Ωι Ωι Ωι Ω Ωι O O Ω Ω Ωι ´ ῾

MakeLowercase:

ἀ ἁ ἂ ἃ ἄ ἅ ἆ ἇ ἀ ἁ ἂ ἃ ἄ ἅ ἆ ἇ
ἐ ἑ ἒ ἓ ἔ ἕ ἐ ἑ ἒ ἓ ἔ ἕ
ἠ ἡ ἢ ἣ ἤ ἥ ἦ ἧ ἠ ἡ ἢ ἣ ἤ ἥ ἦ ἧ
ἰ ἱ ἲ ἳ ἴ ἵ ἶ ἷ ἰ ἱ ἲ ἳ ἴ ἵ ἶ ἷ
ὀ ὁ ὂ ὃ ὄ ὅ ὀ ὁ ὂ ὃ ὄ ὅ
ὐ ὑ ὒ ὓ ὔ ὕ ὖ ὗ ὑ ὓ ὕ ὗ
ὠ ὡ ὢ ὣ ὤ ὥ ὦ ὧ ὠ ὡ ὢ ὣ ὤ ὥ ὦ ὧ
ὰ ά ὲ έ ὴ ή ὶ ί ὸ ό ὺ ύ ὼ ώ
ᾀ ᾁ ᾂ ᾃ ᾄ ᾅ ᾆ ᾇ ᾀ ᾁ ᾂ ᾃ ᾄ ᾅ ᾆ ᾇ
ᾐ ᾑ ᾒ ᾓ ᾔ ᾕ ᾖ ᾗ ᾐ ᾑ ᾒ ᾓ ᾔ ᾕ ᾖ ᾗ

ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ ῲ
ᾰ ᾱ ᾰ ᾱ ᾱ ᾱ ᾰ ᾱ ᾰ ᾰ ᾱ ᾿ ι ᾿
῀ ῁ ῂ ῃ ῄ ῆ ῇ ὲ έ ῂ ή ῃ ῍ ῎ ῏
ῐ ῑ ῒ ΐ ῖ ῗ ῐ ῑ ὶ ί ῍ ῎ ῏
ῠ ῡ ῢ ΰ ῤ ῥ ῦ ῧ ῠ ῡ ὺ ύ ῥ ῝ ῞ ῀
ῲ ῳ ῴ ῶ ῷ ὸ ό ὼ ώ ῳ ΄ `

4 LICR Macros

Babel defines macros for several autogenerated strings so that they may appear in the choosen language. *babel-greek* uses LICR macros in order to let the string macros work independent of the font encoding.

If `fontspec` is loaded before babel, babel-greek loads Greek LICR for EU1 or EU2 from the file `greek-euenc.def` provided with *greek-fontenc* since version 0.10.

4.1 Hiatus

The «hiatus» feature works with macro input:

άυλος ↦ ΑΫΛΟΣ, ἄυλος ↦ ΑΫΛΟΣ, mάινα ↦ ΜΑΪΝΑ, κέικ↦ ΚΕΪΚ, ἀυπνία ↦ Α-ΫΠΝΙΑ.

4.2 Captions

Προοίμιον, Ἀναφοραὶ, Περίληψις, Βιβλιογραφία, Κεφάλαιον, Παράρτημα, Περιεχό-μενα, Κατάλογος σχημάτων, Κατάλογος πινάκων, Εὑρετήριον, Σχῆμα, Πίναξ, Μέρος, Συνημμένως, Κοινοποίησις, Πρὸς, Σελὶς, ὅρα, ὅρα ὡσαύτως, Ἀπόδειξις, Γλωσσάριον

Test correct upcasing (dropping of accents):

ΠΡΟΟΙΜΙΟΝ, ΑΝΑΦΟΡΑΙ, ΠΕΡΙΛΗΨΙΣ, ΒΙΒΛΙΟΓΡΑΦΙΑ, ΚΕΦΑΛΑΙΟΝ, ΠΑΡΑΡΤΗ-ΜΑ, ΠΕΡΙΕΧΟΜΕΝΑ, ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ, ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ, ΕΥΡΕ-ΤΗΡΙΟΝ, ΣΧΗΜΑ, ΠΙΝΑΞ, ΜΕΡΟΣ, ΣΥΝΗΜΜΕΝΩΣ, ΚΟΙΝΟΠΟΙΗΣΙΣ, ΠΡΟΣ, ΣΕ-ΛΙΣ, ΟΡΑ, ΟΡΑ ΩΣΑΥΤΩΣ, ΑΠΟΔΕΙΞΙΣ, ΓΛΩΣΣΑΡΙΟΝ

4.3 Months

18 Ἰανουαρίου 2014
18 Φεβρουαρίου 2014
18 Μαρτίου 2014
18 Ἀπριλίου 2014
18 Μαΐου 2014
18 Ἰουνίου 2014
18 Ἰουλίου 2014
18 Αὐγούστου 2014
18 Σεπτεμβρίου 2014

18 Ὀκτωβρίου 2014
18 Νοεμβρίου 2014
18 Δεκεμβρίου 2014

5 Greek Numerals

See greek.pdf for the formation rules of Greek numerals. Some examples:

αʹ, βʹ, γʹ, δʹ, εʹ, ϛʹ, ζʹ, ηʹ, θʹ, ιʹ, ιαʹ, ιβʹ, κʹ, τμεʹ, φʹ, ͵αϠϙζʹ, ͵βιγʹ,

Αʹ, Βʹ, Γʹ, Δʹ, Εʹ, Ϛʹ, Ζʹ, Ηʹ, Θʹ, Ιʹ, ΙΑʹ, ΙΒʹ, Κʹ, ΤΜΕʹ, Φʹ, ͵ΑϠΘϘΖʹ, ͵ΒΙΓʹ,

Enumerated lists use Greek characters/numerals in the second and fourth level:

1. item 1
    (αʹ) item 1.1
        i. item 1.1.1
            Αʹ. item 1.1.1.1
            Βʹ. item 1.1.1.2
        ii. item 1.1.2

This may be problematic with fonts that only partially support Greek and miss the numeral signs (dexiakeraia and aristerikeraia). You may redefine the commands \textdexiakeraia and \textaristerikeraia to some substitute characters. Or, if you prefer the "normal" enumeration, write in the preamble after loading babel:

```
\makeatletter
\addto\extrasgreek{\let\@alph\latin@alph
      \let\@Alph\latin@Alph}
\makeatother
```