

# Test Greek LaTeX internal character representations (LICR macros)

Günter Milde

2023/03/08

This document tests the compatibility of “luainputenc” and the Greek font setup for TU and PU. It uses only ASCII input.

See the source `test-inputenc.tex` for the input used in the examples.

TODO: Compiling with LuaTeX fails: `tuenc-greek.def` uses literal Unicode characters which is incompatible with legacy 8-bit encodings defined via "luainputenc".

## Contents

## 1 LICR input

The LaTeX internal character representation (LICR) is a verbose, fail-safe 7-bit ASCII encoding that can be used unaltered under both, 8-bit TeX (with any ASCII-compatible input encoding) and XeTeX/LuaTeX. Use cases are macro definitions and generated text.

## 1.1 Greek alphabet

Greek letters via 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 χ 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:

$\sigma$   $\text{textrm{sigma}}$   
 $\varsigma$   $\text{textrm{finalsigma}}$  or  $\text{textrm{varsigma}}$

With Unicode fonts (Xe/LuaTeX, font encoding TU), the `\textautosigma` macro (which automatically chooses the glyph according to the position) does not work with LaTeX versions older than 2022/06 (requires the new `\MakeLowercase` implementation).

## 1.2 Diacritics

Greek accents are tonos = oxia, varia, psili, dasia, dialytika, and perispomeni.

Greek diacritics can be input by named macro or symbol macro:

Diacritics as spacing characters:

with empty argument:

with space as argument: `//  ...  ~  ~  ~  ~  ~`

with protected space: / / / / o .. ~ \* \* \* \* \* \* \*

### 1.2.1 mute iota

The mute iota is input after the base letter.

- \ypogegrammeni following a Greek letter sets a sub-iota (corresponding to COMBINING GREEK YPOGEGRAMMENI), e.g.  $\alpha$

In Unicode, a GREEK CAPITAL LETTER ... followed by COMBINING GREEK YPOGEGRAMMENI is normalized to GREEK CAPITAL LETTER ... WITH [...] AND] PROSGEGRAMMENI, if a corresponding letter exists in the Unicode standard. In LGR fonts, this is implemented via a

ligature definition (set the Babel language or wrap in `\ensuregreek`):  $A_i$  but  $\Lambda_i$ .

The shape and position of the mute iota with pre-composed capital letters depends on the selected font, both sub-iota and adscript iota are possible.

- `\prosgegrammeni` sets an adscript iota (GREEK PROSGEGRAMMENI), e.g.  $A_i$ . In Unicode fonts the prosgegrammeni is spaced similar to the letter iota. In the CB Greek fonts, the only visible difference to the pre-composed characters is a slightly increased spacing.

Copy/Paste may convert the adscript iota to a small letter iota!

`Ypogegrammeni` and `prosgegrammeni` following matching/not-matching base character (unchanged, lowercase, uppercase):

$A_i A_i \alpha\alpha A_i A_i$   
 $\Lambda_i \Lambda_i \lambda\lambda \Lambda_i \Lambda_i$   
 $\alpha\alpha_i \alpha\alpha A_i A_i$

Using `\ypogegrammeni` for the mute iota with both, small and capital letters usually gives better results.

## 1.3 Additional Greek symbols

### 1.3.1 symbols for Greek numbers

$\text{\textkoppa}$   
 $\text{\textKoppa}$   
 $\text{\textqoppa}$  (archaic koppa)  
 $\text{\textQoppa}$  (archaic Koppa)  
 $\text{\textstigma}$   
 $\text{\textStigma}$  (Sigma-Tau-Ligature in CB-fonts)<sup>1</sup>  
 $\text{\textsampi}$   
 $\text{\textSampi}$   
 $\text{\textdigamma}$   
 $\text{\textDigamma}$   
 $\text{\textdexiakeria}$   
 $\text{\textaristerikeraia}$

### 1.3.2 generic text symbols

LICR macros for some symbols from the 8-bit font encoding LGR that are not confined to Greek but not defined in `tuenc.def` [2018/08/11 v2.0j].

$\text{\textsemicolon}$   
 $\text{\textmicro}$   
 $\text{\textschwa}$

The SI unit prefix MICRO SIGN is not upcased with MakeUppercase:

$\text{\textmu}$   $\mapsto M$  but  $\text{\textmicro}$   $\mapsto \mu$ .

---

<sup>1</sup>the name “stigma” originally applied to a medieval sigma-tau ligature, whose shape was confusingly similar to the cursive digamma

text		mathematics	
macro	output	macro	output
\textpi	$\pi$	\pi	$\pi$
\textvarpi	missing	\varpi	$\varpi$
\textpisymbol	$\pi$		
\textrho	$\rho$	\rho	$\rho$
\textvarrho	missing	\varrho	$\varrho$
\textrhosymbol	$\rho$		
\texttheta	$\vartheta$	\theta	$\theta$
\textvartheta	missing	\vartheta	$\vartheta$
\textthetasymbol	$\vartheta$		
\textepsilon	$\varepsilon$	\epsilon	$\epsilon$
\textvarepsilon	missing	\varepsilon	$\varepsilon$
\textepsilonsymbol	$\varepsilon$		
\textphi	$\varphi$	\phi	$\phi$
\textvarphi	missing	\varphi	$\varphi$
\textphisymbol	$\varphi$		
\textbeta	$\beta$	\beta	$\beta$
\textvarbeta	missing	missing	
\textbetasymbol	$\beta$		
\textkappa	$\kappa$	\kappa	$\kappa$
\textvarkappa	missing	\varkappa	$\varkappa$
\textkappasymbol	$\kappa$		
\textTheta	$\Theta$	\Theta	$\Theta$
\textvarTheta	missing	missing	
\textThetasymbol	$\Theta$		

Table 1: Macros for Greek symbol variants

## 2 Greek in section headings

The packages *textalpha* and *alphabeta* as well as *babel-greek*, add hyperref support for L<sup>I</sup>C<sub>R</sub> input with non-standard accents or combined diacritics.