

Greek in Proper ConT_EXt

T_EX has always had some support for typesetting Greek characters: most (all?) Greek letters are contained in the Computer-Modern math fonts. In ConT_EXt one can use the iso-8859-7 encoding to write modern Greek.

However, if you wanted to typeset ancient (polytonic) Greek, you had to look for specialized solutions. Unlike modern Greek, ancient Greek uses three accents (acute ´, gravis ` , circumflex ~), two breathings (the “smooth” breathing ¨ and the “rough” breathing ͂), the dieresis or “trema” ¨ and one text-like subscript character, the iota subscript (which only occurs with the vowels α η ω). These combine with each other and all seven Greek vowels, and for satisfying typographic results, these combinations have to be precomposed. A number of solutions exists for L^AT_EX; none of them works “out of the box” for ConT_EXt.

When I began to think about support for polytonic Greek in ConT_EXt, a number of requirements was obvious:

- Most users will want to mix shorter Greek passages with text in Roman alphabets. There has to be an easy and consistent way to switch between Roman and Greek passages.
- There should be a simple way to switch between several Greek fonts; every font should follow the same encoding and behave in the same way.
- Since different Roman fonts have different sizes, there has to be a simple way to scale the Greek font to adapt it to the size of the main bodyfont.

The next question was how Greek should be input into the T_EX source file. There are several handcrafted solutions for different operating systems and applications, but none of them is cross-platform portable, none can be called standard. Two ways were obvious as being standardized and portable; both have their advantages and drawbacks:

1. Babel input uses only ASCII-characters; these are mapped to (more or less) corresponding Greek characters. Accents are typed with the characters ` ´ ~, breathings with <>, iota subscript with |; combining these accents with vowels is achieved via T_EX’s internal ligature mechanism. Thus, >~a| would become ᾗ. Advantage: ASCII can be edited on any system, in any editor. This solution is highly backward-compatible. Drawbacks: If you have longer passages in Greek, reading your source file becomes pretty cumbersome. Moreover, since accents like ~ are “active” characters in T_EX by default, babel has to rely on some catcode changes; making this work inside other commands (e.g., inside tables) demands some special considerations and may break in some special circumstances.
2. Unicode (utf-8) input allows you to type and see the Greek characters right in your source file. Advantages: Unicode is undoubtedly the future; it is cross-

platform portable, and it provides one precisely defined slot for every character. Drawbacks: not every system and every editor supports Unicode. But even if your computer does, you will still have to look for a proper keyboard driver that will allow you to write these Greek characters. Most importantly, T_EX is still limited to 256 characters per font, so making Unicode work in T_EX demands some trickery. We all hope that projects like Omega and Aleph will one day provide a solution to this problem, but this may take quite a while.

Note that the advantages that Unicode offers are only relative in one aspect: It is true that in order to use Babel, you have to memorize which ASCII character maps to which Greek character. But as long as you use a normal Western keyboard to write your T_EX-files, the same is true if you use Unicode; the only advantage is that you get immediate visual feedback. On the other hand, making Unicode work with editors like emacs is not trivial, so I find myself still writing Greek in Babel most of the time.

1 Some Preliminary Remarks.

The first Greek fonts, developed in the Italian Renaissance, tried to imitate Greek handwriting as closely as possible. Hence (like early “Gothic” fonts), they contained a large number of ligatures and different letter-forms. In the course of times, Greek fonts became simplified; they now contain glyphs for the 24 lower-case and upper-case letters of the Greek alphabet. Combinations of vowel + breathing/accent/iota subscript are generally not composed out of different glyphs, but precombined. Conventions in different areas vary as to additional letters: traditionally, we differentiate whether the letter sigma appears at the end of a word (ς) or within the word (σ). During the last years, fonts have become more popular that use only one “round” or “lunate” form of sigma (ς). In France, most Greek fonts have a special form for the letter beta within a word, using the “normal” form only at the beginning of words. Moreover, the form of the circumflex accent can either be shaped like a tilde (~; this seems to be the preferred way in German-speaking countries) or round, the so-called “Porsonian” form (after the English philologist Richard Porson, 1759–1808, ^, which seems to be preferred in English-speaking countries).

Moreover, in T_EX, users think of fonts as combined in complete sets, with (at least) serif, sans, and typewriter shapes available in roman, italic and semibold; maybe with a SmallCaps variant thrown in for good measure. There is a family of Greek fonts for use with T_EX that offers precisely such a complete set, CBGreek, produced by Claudio Beccari. However, most professional high-quality fonts that contain Greek offer only one form. For the purpose of this module, I found this restriction of little importance. While users who want to typeset entire books or periodicals in modern Greek will want to have such a complete set, scholarly work usually requires just one form for Greek letters. Moreover, there is a historical reason why “Greek italics” does not exist: At a very early date, typesetters distinguished between an upright (“Roman”) and slanted (“Italic”) form of Latin letters; both forms were mixed for graphical effect and to add emphasis to parts of the text; whence our use of “Italics.”

This distinction has no historical existence for Greek fonts because here, “Roman” and “Italic” forms did not exist. Nevertheless, I can see why users may want to have a form of Greek characters that adds emphasis. When you use the Greek modules, it is possible to mix two Greek fonts. Although I would caution against using these artificial forms, there are three “italic” and three “bold” Greek fonts (and even two “bold italic” variants, which I consider really spurious).

2 Installation

Installation is straightforward: unpack the gzipped archive. On most popular operating systems, you will be able to simply double-click the file, and an appropriate utility (like StuffIt Expander) will take care of this. You might as well use the command line since some of the steps necessitate using it anyway:

```
tar zxvf greek.tar.gz
```

As always, fonts and map files will have to go to directories “where T_EX can find them,” as the sacred formula says in most manuals. Depending on your system, that will either be a local tree (if you don’t know where this is located, type `kpsewhich -expand-var=' $TEXMFLOCAL '` on the command line) or a texmf-tree in your home directory (again, you could type `kpsewhich -expand-var=' $HOMETEXMF '` to find out where that could be). On most systems, you can simply copy the zip-archive “texmf.zip” to your local or home texmf-tree and unzip it.

If you have added these files to a texmf-tree that uses ls-R files, you will need to rebuild this database by running `texhash` (if you don’t know what this means, run it anyway, it won’t hurt).

Finally, add the new map file to the configuration of your postscript-drivers:

```
sudo updmap --enable Map tasgreek.map
```

3 Usage

The package contains two modules: `t-greek` enables `babel-input`, `t-utfgre` (you guessed it) enables `utf-input`. The user interface is identical; choose either one in the preamble of your document:

```
\usemodule[t-greek]
```

Then, the appearance of the Greek font needs to be set up with the command

```
\setupgreek[font=,altfont=,scale=]
```

For the fonts, there are eleven values corresponding to the Greek fonts which the module enables. Because of the unclear license situation, the module includes just a few fonts; the others can be downloaded on the web. Here’s a very brief description of the fonts:

Alkaios This is a very nice-looking truetype-font that can be downloaded at Lucius Hartmann’s homepage <http://www.lucius-hartmann.ch/diverse/greekfonts/#unicode>). It will mix very well with the Roman fonts used

most often such as Times or Garamond. I have included the tfms but not the font itself, so you will need to download it and put it into the fonts/truetype directory of your texmf-tree.

- | | |
|------------|---|
| Cardo | Another truetype-font; it has a very wide range of special characters. The typographical quality is a matter of taste. Download at http://scholars-fonts.net/cardofnt.html . |
| Dioxipe | Yet another truetype-font, somewhat similar to Alkaios, but it uses the “tilde”-form of the circumflex accent. In its normal design-size, Dioxipe is noticeably smaller than most roman fonts, so you will need to scale it a bit. Dioxipe can be downloaded from http://semata.delendis.com/fuentes.html . Again, all the necessary files are included, but you’ll have to download the font yourself. |
| Gentium | This is a wonderful truetype-font designed by Victor Gaultney; it can be downloaded at http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=Gentium_download . There is also an “Italic” version Gentiumit. The special beauty of Gentium lies in the fact that Latin and Greek characters are drawn consistently: the Roman “M” and the uppercase Greek “Mu” will be identical if you use Gentium both for the Latin and Greek text. |
| Greekserif | This truetype font has been developed for Debian Linux by Peter Hawkins. It has both a very nice Roman font and good looking Greek characters and comes in a Regular, Italic (Greekserifit), bold (Greekserifbo) and bold italic (Greekserifboit) form. It can be downloaded at http://packages.debian.org/unstable/x11/ttf-freefont.html . (You will need the tool fontforge in order to use this font. There is another version at http://gainsford.tripod.com/archive/freeserif.zip which has the ttf-files in a form that should be immediately usable, but unfortunately, this version is broken; it is missing a number of glyphs.) |
| Ibycus | This Greek font was created by Pierre McKay for the Ibycus-package for L ^A T _E X. I modified it a bit for using it with ConT _E Xt; the font is included. |
| Leipzig | This font was created by Claudio Beccari and is part of the CBGreek fonts. I have used the type-I version as it looks much nicer on screen; the font is included. |
| Minion | This font comes with the free Adobe Reader software (downloadable at http://www.adobe.com/products/acrobat/readstep2.html) and contains the Greek Extended glyphs; there is also a bold version, called Minionbo. |

- Oxoniensis Another truetype-font that can be downloaded at <http://semata.delendis.com/fuentes.html>. It is a version of the Greek font used by Oxford University Press, and it looks very beautiful, especially at smaller sizes.
- Palatino This font comes with the Windows operating system (it can also be downloaded at <http://www.ellopos.net/elpenor/greek-texts/greek-fonts.asp> and other sites). Although it has been designed by the famous Hermann Zapf, the font has a number of odd-looking characters (see Jeffrey Rusten's review at <http://ccat.sas.upenn.edu/bmcr/1998/98.1.11.html>). It is also available in italic (Palatinoit), bold (Palatinobo) and bold italic (Palatinoboit).
- Teubner This is a font that I created myself after the Greek typeface used by the publishing house Teubner at the beginning of the twentieth century; it is included in the module.

After choosing your font, it will be necessary to adapt the scaling. As you know, roman fonts look very different even at the same size: Lucida Bright at 10 pt is almost as big as Computer Modern at 12 pt. Since our Greek fonts will be mixed with roman text, the module provides the ability to adapt the size of the Greek font: `[scale=1]` would correspond to the design size of the font; `[scale=.95]` would shrink the font to 95%; `[scale=1.05]` magnify to 105%. Just play around until the relation between both fonts looks right.

For eight out of these fonts (viz. Alkaios, Cardo, Dioxipe, Gentium, Gentiumit, Ibycus, Leipzig, Minion, Minionbo, Oxoniensis), an alternative with the “lunar” (rounded) shape of the letter sigma (without distinction between final and intermediate sigma) is available; this can be specified by adding “lun” directly to the name of the font, e.g., `\setupgreek[font=Gentiumlun,scale=1.2]`. The font “Teubner” only has the round form, the other fonts do not include the lunate sigma. After setting these values in the preamble, switching to Greek in the body of your documents is easy: use the command `\localgreek{...}` for shorter passages and the environment `\startgreek \stopgreek` for longer stretches (the corresponding commands for the alternate Greek font are `\localaltgreek{...}` for shorter passages and the environment `\startaltgreek \stopaltgreek` for longer stretches). It is safest not to include commands within these Greek passages: if, e.g., you want to leave a vertical space by writing `\blank[big]`, it is best to end Greek input by saying `\stopgreek` or closing the braces before the command.

Input according to the Babel standard has been around for quite a while, so I give just a brief description (you can find an extended discussion in the documentation to the “Teubner” package for L^AT_EX). First, you will have to memorize the correspondence between Greek letters and their ASCII equivalents:

A	B	G	D	E	Z	H	J	I	K	L	M	N	X	O	P	R	S	T	U	F	Q	Y	W
A	B	Γ	Δ	E	Z	H	Θ	I	K	Λ	M	N	Ξ	O	Π	P	Σ	T	Υ	Φ	X	Ψ	Ω
a	b	g	d	e	z	h	j	i	k	l	m	n	x	o	p	r	s	t	u	f	q	y	w
α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο	π	ρ	σ	τ	υ	φ	χ	ψ	ω

Accents, breathings, and iota subscript are written with ' ` ~ < > |; these can be combined as necessary. I provide just a few examples:

```
>e = ê
"~u = ũ
<~w| = ẁ
```

The original Greek encoding for L^AT_EX took care of the letter sigma: when it occurred at the end of a word, it would automatically be transformed into a final sigma ζ. After some thought, I decided not to follow this approach: there are too many cases where users might want a “normal” sigma even in front of a space or a punctuation mark. In order to obtain a final sigma, type “c.”

Moreover, the modules provide a couple of convenient commands to typeset symbols that are needed for writing ancient Greek or editing papyrological material:

```
\digamma Ϝ
\sampi ϗ
\stigma Ϛ
\koppa ϙ
\lunars ϭ
\lunarS ϭ
\textbraceleft {
\textbraceright }
\halfbraceleft ⏟
\halfbraceright ⏟
\textdoublebracketleft ⌈ (can also be input by typing [ [ in your source file)
\textdoublebracketright ⌋ (can also be input by typing ] ] in your source file)
\crux †
```

In order to get an apostrophe, type ' '. For opening and closing quotes, type ((and)). To obtain the sublinear dot in papyrological or epigraphical editions, type an exclamation mark after the character: >a!n!'h!r! becomes ἀνῆρ.

Unicode input is intuitive: just place `usemodule[t-utfgre]` in the preamble of your document; setting up the Greek fonts is identical. If you have a proper keyboard driver for input of Unicode Greek, just type away, but remember to include Greek passages between `\localgreek{...}` or `\startgreek \stopgreek`. Of course, all the named glyphs are available for Unicode-input as well.

Please be advised that not all fonts contain all symbols or characters. Archaic number symbols, the papyrological double brackets or half brackets or characters that do not occur in normal Attic morphology (like ε̃) are not included in some of the fonts, so if you need those, you will have to experiment which font provides them (Ibycus, Leipzig and Cardo are pretty complete).

Moreover, if you need such special symbols, using Babel input may be easier for you. Since they have not been assigned a Unicode value, they cannot be typed directly with a Unicode keyboard driver. However, if you want to use them in a Unicode file, there is a simple way to obtain them: every character in the Greek fonts can be accessed via a command name. Just open the file `enco-agr.tex` in your ConTeXt directory, and you will find that, e.g., ξ can also be accessed by typing `\greekepsilon` in your source file.

4 Goodies

Since both modules provide the same commands, you cannot mix babel and Unicode Greek *within the same source file* (of course, you can use both methods for different files). However, the modules contain a very handy tool if you change your mind about which input method you want to use. The perl scripts `babeltounicode.pl` and `unicodetobabel.pl` will convert Greek passages either from Babel to Unicode or the other way around. This is one reason why keeping the `\localgreek{...}` business is useful even for Unicode input: that's how these scripts know which parts of your file need translating. That's also the reason why you shouldn't include any commands in your Greek passages: the script `babeltounicode.pl` will inexorably convert them to Unicode Greek, and of course, TeX will then complain about unknown commands when you compile the file.

Put these scripts somewhere in your `$Path`, make them executable and run them like so:

```
babeltounicode FILE_NAME.tex
```

This will give you a new file `FILE_NAME.unicode.tex` where all the Greek passages should now be in Unicode. Please be advised that these are homespun solutions. They work reasonably well for me and get 99% of the Greek right, but expect that they will miss some special character every now and then.