

**NAME**

uni2beta – Convert a polytonic Greek document from UTF-8 Unicode to Beta Code

**SYNOPSIS**

**uni2beta** [{-l|-u}] [-i *input\_file.utf8*] [-o *output\_file.utf8*]

**DESCRIPTION**

**uni2beta**(1) reads a document encoded using Unicode and converts it to Beta Code. Beta Code is an ASCII-only encoding scheme most commonly used for digital representation of polytonic Greek.

Beta Code has become a widely-adopted standard for encoding classical Greek. It was developed by David Packard in the 1970s and adopted by the Thesaurus Linguae Graecae (TLG) Project at the University of California, Irvine shortly thereafter. This encoding was later adopted by the Perseus Project in the 1980s (originally at Harvard University, now at Tufts University) and by many other collections of classical and Koine Greek. Today, the TLG corpus alone contains over 100 million words from classical to Byzantine Greek.

The TLG uses uppercase Latin letters for encoding; the Perseus Project uses lowercase. **uni2beta**(1) defaults to lowercase output. This can be overridden with the ‘-u’ (generate uppercase) option.

Many classicists who use Beta Code have been actively involved in The Unicode Standard, with evolving recommendations for mapping between Beta Code and Unicode. **uni2beta**(1) provides a capability for GNU/Linux users who wish to convert Unicode texts to Beta Code.

Any encountered UTF-8 Unicode Byte Order Mark (BOM), U+FFFE, is ignored.

Note: Thesaurus Linguae Graecae and TLG are registered trademarks of the University of California.

**OPTIONS**

- l Create Beta Code output file using lowercase Latin letters. This is the convention that the Perseus Project uses. This is the default.
- u Create Beta Code output file using uppercase Latin letters. This is the convention that the TLG Project uses.
- i Specify the input file. The default is STDIN.
- o Specify the output file. The default is STDOUT.

Sample usage:

```
uni2beta -i my_input_file.utf8 -o my_output_file.beta
```

**FILES**

Polytonic Greek files in UTF-8 Unicode.

**SEE ALSO**

**unibetaprep**(1), **beta2uni**(1), **libunibetacode**(3), **unibetacode**(5)

**AUTHOR**

**uni2beta** was written by Paul Hardy.

**LICENSE**

**uni2beta** is Copyright © 2018, 2019, 2020 Paul Hardy.

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

**BUGS**

No known bugs exist.