Hyphenator

Fri, 16/05/2008 - 16:20 — webmaster

Utility

Typesetting systems achieve justified paragraph alignment by increasing or decreasing the space between words in each paragraph line. This practice is not applicable to short paragraph lines (e.g. in newspaper columns) or when long words occur at the end of some lines, as excessive word spacing affects the aesthetics of text and part of the printable space is wasted in blanks. In such cases, the traditional practice of hyphenating words at the line edge is irreplaceable. The role of the *Hyphenator* is to indicate all hyphenation points of a word.

Characteristics

The basic characteristics of Neurolingo's Hyphenator for Modern Greek are:

- It utilizes rules and a small exception list. Rules are of two types: the ones that were handcrafted in accordance to linguistic hyphenation rules (of Modern Greek grammar) and the ones that were machine learned from hyphenation patterns found in Neurolingo's Morphological Lexicon [1]. The role of the latter rules is to handle 24 vowel sequences that some times are pronounced as one syllable and other times split into two syllables (phenomenon of synizesis [2]). The exception list, which is accessed before rule application, contains approx. 300 wordforms that can be syllabified (and pronounced) in two ways, but each pronunciation has a different meaning, e.g. ή-πια (I drank) and ή-πι-α (smoothly), λό-για (the words) and λό-γι-α (literary), βιά-ζω (to force) and βι-ά-ζω (to rape), δό-λιος (miserable) and δό-λι-ος (insidious). For such wordforms, the Hyphenator prefers to miss some hyphenation points instead of risking their meaning alteration.
- It hyphenates correctly all the 1.200.000 wordforms of the Morphological Lexicon. For words never seen before, the expected error rate is <0.3%. Wrong hyphenation is expected only for words that contain one of the 24 vowel sequences mentioned above.

Try Hyphenator online. [3]

Applications

The Hyphenator is available for the following text processing environments:

- MS Word 2000/2002(XP)/2003 for Windows and X/2004 for Macintosh.
- OpenOffice Writer 1.1-1.5 & 2.0-2.0.2 for Windows, Linux, Solaris x86 and Solaris Sparc.
- StarOffice Writer 7 & 8 for Windows, Linux, Solaris x86 and Solaris Sparc.
- NeoOffice Writer 1.1-1.5 & 2.0-2.0.3 for Mac OS X.
- QuarkXpress 4.03 for Windows.
- Adobe InDesign 2.0/3.0(CS)/CS2 for Windows and Macintosh.
- Adobe Photoshop 7.0/8.0(CS)/CS2 for Windows and Macintosh.
- Adobe Illustrator CS/CS2 for Windows and Macintosh.

Source URL: http://www.neurolingo.gr/en/technology/application_tools/hyphenator.jsp

©2005-2010 Neurolingo L.P.

Links:
[1] http://www.neurolingo.gr/en/technology/lexica/morpholexicon.jsp
[2] http://www.neurolingo.gr/en/technology/application_tools/hyphenator/synizesis.jsp
[3] http://www.neurolingo.gr/en/online_tools/lexiscope.htm

©2005-2010 Neurolingo L.P.

W3C XHTML 1.0