

LaTeX source file for: XeLaTeX_World Script Examples

```
1 \documentclass[11pt,letterpaper]{article}
2 \usepackage[quiet]{fontspec}
3 \ExplSyntaxOn
4 \cs_new_protected:Npn \fp_div:Nn #1 #2
5   {
6     \fp_set:Nn #1 { #1/#2 }
7   }
8 \ExplSyntaxOff
9 \usepackage{color}
10 \usepackage{url}
11 \usepackage{graphicx}
12
13 \usepackage{xltextra}
14 \usepackage{polyglossia}
15
16 \defaultfontfeatures{Mapping=tex-text,Scale=MatchLowercase} % to support
17   TeX conventions like _____
18 \setmainfont{Times New Roman}
19 \setdefaultlanguage[variant=american]{english}
20 \setotherlanguages{greek,hebrew,russian,syriac}
21
22 \title{The World's Writing Systems: A XeLatex document showing
23   typesetting in a variety of writing systems}
24 \author{Jake Warde}
25
26
27 \begin{document}
28 \maketitle
29 \begin{abstract}
30 Sections 1–5 show examples of typesetting in different writing systems.
31 The sections are organized according to recognized categories of
32 writing systems. The XeTeX Companion (Section 3.1.3) contains an
33 overview of types of writing systems. \footnote{http://xml.web.cern.
34 ch/XML/lgc2/xetexmain.pdf} Another recommended resource on writing
35 systems is \emph{Omniglot: The online encycloppedia of writing
36 systems and languages} \footnote{http://www.omniglot.com/index.htm}
37 Sections 6–7 are a summary of the technologies used to typeset text
38 for all the world's langagues. The document may also be a resource
39 to someone learning to use XeLatex. XeLatex source is here: \url{
40 www.wardepub.com}. No claims to being a XeLatex expert are made. In
41 fact the document was created while learning to use XeLatex.
42 \end{abstract}
43
44 \section{Abjads—Consonant Alphabets}
45 \paragraph{} Abjads, or consonant alphabets, represent consonants only,
46 or consonants plus some vowels. Full vowel indication (vocalisation)
47 can be added, usually by means of diacritics, but this is not
48 usually done. Most of abjads, with the exception of Ugaritic, are
49 written from right to left. There are three abjads currently in use
50 : Arabic, Hebrew and Syriac. Abjads, or consonant alphabets,
51 represent consonants only, or consonants plus some vowels. For more
52 about abjads visit: \url{http://www.omniglot.com/writing/
53 alphabetic.htm}
54
55
56
57 \subsection{Arabic}
58 \newfontfamily\Arabic[Script=Arabic]{Andalus}
59
60 \begin{Arabic}
```

```

43
\end{Arabic}
45 \paragraph {Note:} The sample text in Arabic is from the following
    website offering Modern Standard Classical Arabic courses.\url {
        http://www.dalilusa.com}
47 \subsection {Hebrew}
49
51 \paragraph {Note:}The web page from which the above text was copied
    provides information about the characteristics of the script used
    to write Hebrew.
    \url {http://rishida.net/scripts/samples/hebrew}
53
\subsection {Syriac}
55 \newfontfamily\syriac [Script=Syriac] {Estrangelo Edessa}
\begin {syriac}
57
\end {syriac}
59
\subsection {Phoenician}
61 \paragraph {} The Phoenician alphabet, called by convention the Proto-
    Canaanite alphabet for inscriptions older than around 1200 BCE, is
    also an abjad. It was used for the writing of Phoenician, a
    Northern Semitic language, used by the civilization of Phoenicia.
63 Phoenician became one of the most widely used writing systems, spread by
    Phoenician merchants across the Mediterranean world, where it was
    assimilated by many other cultures and evolved. The Aramaic
    alphabet, a modified form of Phoenician, was the ancestor of modern
    Arabic script, while Hebrew script is a stylistic variant of the
    Aramaic script. The Greek alphabet (and by extension its
    descendants such as the Latin, the Cyrillic and the Coptic), was a
    direct successor of Phoenician, though certain letter values were
    changed to represent vowels.\footnote{http://en.wikipedia.org/
    Phoenician alphabet}
65 \vspace {.2cm}
67 \fontspec {Noto Sans Phoenician}
69 \fontspec {Times New Roman}
\paragraph {Note:} Finding some Phoenician text ready for cut-and-paste
    into this document was a little difficult (as one might imagine).
    But in figuring out how to input Phoenician text I discovered an
    excellent tool for entering any Unicode-encoded symbol into a
    computer. Learn more about \emph {BabelPad} at \url {http://www.
    babelstone.co.uk/index.html}
71
\section{Alphabets}
73 \paragraph {}Alphabets, or phonemic alphabets, are sets of letters,
    usually arranged in a fixed order, each of which represents one or
    more phonemes, both consonants and vowels, in the language they are
    used to write. In some cases combinations of letters are used to
    represent single phonemes, as in the English sh, ch and th. For a
    history of alphabets, lists of alphabets currently in use,
    alphabets used to a limited extent, and alphabets that are no
    longer used visit: http://www.omniglot.com/writing/alphabets.htm.

```

75 Most specialists believe that the Phoenician alphabet, an abjad, was adopted for Greek during the early 8th century BC, perhaps in Euboea. The earliest known fragmentary Greek inscriptions date from this time, 770–750 BC^{footnote} {The date of the earliest inscribed objects; A.W. Johnston, "The alphabet", in N. Stampolidis and V. Karageorghis, eds, Sea Routes from Sidon to Huelva: Interconnections in the Mediterranean 2003:263–76, summarizes the present scholarship on the dating}, and they match Phoenician letter forms of c. 800–750 BC.^{footnote}{Pierre Swiggers, Transmission of the Phoenician Script to the West, in Daniels and Bright, The World's Writing Systems, 1996}. The innovation of the Greek writing system was that symbols were used to represent vowels. The introduction of vowels into the script (vocalization) "and the abstraction of consonants as a process in speech — meant that the number of signs (or characters) actually used to represent spoken language could be reduced to a relative handful (around 30) because even in complex languages, such as some Inuit dialects, the number of actual consonantal and vocal sounds is finite. \url {http://people.ucalgary.ca/~mwolfe/Syllabary,alphabet.html}

```
77 \subsection {Greek}
79 \newfontfamily\greekfont[Script=Greek] {Calibri}
81
83
85
87
\paragraph {Note:} The above text is the first 8 lines of the Illiad by Homer. See: \url {http://www.sacred-texts.com/cla/homer/greek/ili01.htm}
```

```
91 \subsection {Cryllic Script Languages}
93 \newfontfamily\cyrillicfont[Script=Cyrillic] {Arial Unicode MS}
: c
(NomC)
(InsC).
, NomC
InsC
NomC InsC
1950
```

```

NomC,
InsC.
InsC , NomC NomC
NomC, InsC
InsC.\footnote{Sample text from: Kuznetzova,
J;Diachronic distribution of predicate nouns in Russian: Russian
Linguistics 2013 http://link.springer.com/article/10.1007/s11185
-012-9107-1}
95 \section {Syllabic Alphabets}
\paragraph {} Syllabic alphabets, alphasyllabaries or abugidas consist
of symbols for consonants and vowels. The consonants each have an
inherent vowel which can be changed to another vowel or muted by
means of diacritics. Vowels can also be written with separate
97 letters when they occur at the beginning of a word or on their own.
When two or more consonants occur together, special conjunct symbols are
often used which add the essential parts of first letter or
letters in the sequence to the final letter. Sample syllabic
99 alphabets are shown at: \url {http://www.omniglot.com/writing/
syllabic.htm}. Two syllabic alphabets are shown.
\subsection {Devan gar }
101 \paragraph {}Devanagari, also called Nagari (N gar , the name of its
parent writing system, is an abugida alphabet of India and Nepal.
Because of its use to write the Hindi language, devanagari is one
of the most used and adopted writing systems in the world.The use
of the name devan gar is relatively recent, and the older term
n gar is still common.The rapid spread of the term devan gar
may be related to the almost exclusive use of this script to
publish sacred Sanskrit texts. This has led to such a close
connection between devanagari and Sanskrit that devanagari is now
widely thought to be the Sanskrit script; however, before the
colonial period there was no standard script for Sanskrit, which
was written in whatever script was familiar to the local populace.\
footnote {http://en.wikipedia.org/wiki/Devanagari}
103 \vspace {.5cm}
105 \fontspec {Mangal}
107
109

```

111

-

113

`\fontspec {Times New Roman}`

115

`\paragraph {Note:}`The web page from which the above text was copied supports the use of Sanskrit in daily life in the cultural life of India. The text is a welcome to Sanscrutam.com `\url {http:www.sanscrutam.com}`

117

`\subsection {Bengali}`

The Bengali alphabet is the writing system for the Bengali language. The script with variations is shared by Assamese and is basis for Meitei, Bishnupriya Manipuri, Kokborok, Garo and Mundari alphabets.

All these languages are spoken in the eastern region of South Asia. Historically, the script has also been used to write the Sanskrit language in the same region. It is written from left to right and lacks distinct letter cases. It is recognizable by a distinctive horizontal line running along the tops of the letters that links them together, a property it shares with two other popular Indian scripts: Devanagari (used for Hindi, Marathi and Nepali) and Gurumukhi (used for Punjabi). The Bengali script is, however, less blocky and presents a more sinuous shape. Because of the large population of literate Bengali speakers, Bengali script is also one of the more widely used writing systems in the world.

119

`\fontspec {Bangla}`

121

:

;

123

`\fontspec {Times New Roman}`

`\paragraph {Note:}` The sample text in Bengali is from `\emph {WikiTravel: The Free Travel Guide}` `\url http://wikitravel.org/en/Bengali_phrasebook`. The text is Article 1 of the Universal Declaration of Human Rights (by the United Nations):

Article 1: All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience. Therefore, they should act towards one another in a spirit of brotherhood.

125

127

`\section {Syllabaries}`

129

`\paragraph {}`A syllabary is a phonetic writing system consisting of symbols representing syllables. A syllable is often made up of a consonant plus a vowel or a single vowel. In a true syllabary there is no systematic graphic similarity between characters that share a common consonant or vowel sound. That is, the characters for 'ke', 'ka', and 'ko' have no similarity to indicate their common "k" sound (e.g. hiragana \

```

131 fontspec {FangSong}          ).\fontspec {Times New Roman}\footnote
    {http://en.wikipedia.org/wiki/Syllabary}

133 \subsection {The Cherokee Syllabary}
\paragraph {}The Cherokee syllabary is a syllabary invented by Sequoyah,
    also known as George Gist, to write the Cherokee language in the
    late 1810s and early 1820s. His creation of the syllabary is
    particularly noteworthy in that he could not previously read any
    script. He first experimented with logograms, but his system later
    developed into a syllabary. In his system, each symbol represents a
    syllable rather than a single phoneme; the 85 (originally 86)[1]
    characters in the Cherokee syllabary provide a suitable method to
    write Cherokee. Some symbols do resemble the Latin, Greek and even
    the Cyrillic scripts' letters, but the sounds are completely
    different (for example, the sound /a/ is written with a letter that
    resembles Latin D). Each of the characters represents one syllable
.
135 \newfontfamily \Cherokee {Plantagenet Cherokee}
\begin{Cherokee}
137 \vspace {.5cm}
139
,
,
,
,
,
,
-
,
,
,
-
.
\end{Cherokee}
141 \section {Semanto-Phonetic Writing Systems}
143 \paragraph {} Semanto-phonetic writing systems often represent both
sound and meaning. As a result, these scripts generally include a
large number of symbols: anything from several hundred to tens of
thousands. In fact there is no theoretical upper limit to the
number of symbols in some scripts, such as Chinese. These scripts
could also be called logophonetic, morphophonemic, logographic or
logosyllabic. Two semanto-phonetic scrips are currently in use,
namely Chinese and Japanese (Nihongo). For more on semanto-phonetic
writing systems visit: \url http://www.omniglot.com/writing/
semanto-phonetic.php
145 \subsection {Chinese}
\newfontfamily \Chinese {FangSong}
147 \begin{Chinese}
149
151 NICHOLAS KULISH
153
155 2 4
157
159 \end {Chinese}
161 \paragraph {Note:}The text excerpt is from the New York Times Chinese

```

```

Edition , February 23, 2013. Nicholas Kulish covers the upcoming
elections in Cyprus.\footnote{http://cn.nytimes.com/}
163 \subsection {Japanese (Nihongo)}
165 \section {Technologies used to set the type in the worlds writing
systems}
167 \subsection {Overview of Technical Issues in Supporting Scripts}
\paragraph {}This section is based on a discussion of technical
challenges in typesetting the worlds scripts presented by SIL
International \url{http:\\scripts.sil.org}
169 "There are a number of challenges in the typesetting of non-Roman
scripts. These include problems of interaction between the font and
typesetting system, problems of very large character sets, and
considerations of typesetting the particular script. Happily, the
advent of new computer technology has considerably lessened some of
these problems."
171 \subsection {Font Basics}
173 \subsection {Overview of Unicode Character Set}
175 \includegraphics [scale=0.20]{Unicode_Map}
177 \subsection {Codepages}
179 \end{document}

```