

Rendering issues

Ken Zook

October 30, 2008

Contents

1	Introduction	1
2	Uniscribe.....	1
3	Graphite	1
4	Unicode fonts.....	2
5	Doulos SIL or Charis SIL font installation problems.....	3
6	Saltillo problems in FieldWorks 5.4.....	3

1 Introduction

As discussed in Unicode introduction.doc, the Unicode data stored in FieldWorks reflects the underlying characters in the language, but requires special rendering engines to display the data properly either on the screen or for printing. FieldWorks uses two engines for rendering data: Uniscribe and Graphite. As of FieldWorks 5.4 FieldWorks also uses MultiScribe which adds most Graphite capabilities to standard Windows controls used in FieldWorks applications.

2 Uniscribe

Microsoft provides the Uniscribe engine that supports most major languages in the world using standard fonts. This engine works in most current Windows software including Microsoft Office, Open Office, Toolbox, and Paratext. Uniscribe is a DLL that contains knowledge about the different scripts and how they need to be rendered, based on context. It generally works well for scripts it knows about. One thing we have learned is that data for some languages (e.g., Korean) needs to be encoded with normalization NFC. As a result, FieldWorks converts all data to NFC before giving it to the Uniscribe engine.

3 Graphite

The Graphite rendering engine was developed by SIL to support languages and scripts that Uniscribe does not support. In order to use Graphite, an application program must be designed to use Graphite. Currently Graphite support is in FieldWorks, as well as a version of Open Office and support is nearing completion for InDesign. A version also runs on Linux.

Graphite provides the capability for sophisticated users to develop a Graphite table for a given font. The table controls glyphs and positioning for code points based on context. This information is then compiled into the font, making it a Graphite font. The Graphite tables added to the font are ignored by programs not designed to use Graphite. For information or help in developing Graphite fonts, contact the SIL NonRoman Script Initiative (NRSI) at nrsi@sil.org. You can also check out the NRSI Web site at http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&cat_id=RenderingGraphite.

Graphite fonts contain font features that allow glyph customizations of certain code points. For example, a font feature in Doulos SIL and Charis SIL allows a user to choose one of four glyphs when displaying an uppercase Eng (U+014A). FieldWorks can make use of Font Features in Graphite fonts. When users are in the Writing System Properties page, a Font Features button appears next to the font names. These are enabled for Graphite fonts that contain features. Each FieldWorks writing system stores selected features and uses them for displaying that writing system. Read the Graphite font documentation for details on the features it supports (e.g., c:\Program Files\SIL\Fonts\DoulosSIL\DoulosSIL4.100FontDocumentation.pdf).

4 Unicode fonts

To display a writing system in FieldWorks, a single Unicode font must contain all the glyphs needed for that writing system. If a glyph is not present in a font, Microsoft products will search other fonts to try to find and use a glyph to display. FieldWorks will not do this at this time. It displays a box for any glyph it cannot render from the current font. (FieldWorks also displays italic Arabic text in Times New Roman as square boxes. This may be a FieldWorks bug or some deficiency in the Times New Roman font.)

If a font contains Graphite tables, FieldWorks will use Graphite to render the text. There are a few places in current FieldWorks applications where the display has not been Graphite-enabled, but the majority of places provide Graphite support. In these limited places, FieldWorks will use Uniscribe instead of Graphite.

FieldWorks installs two Graphite-enabled fonts developed by NRSI:

- Doulos SIL is a serif font (similar to Times New Roman) that contains a comprehensive inventory of glyphs needed for almost any Roman- or Cyrillic-based writing system, whether used for phonetic or orthographic needs. Doulos SIL only comes in a Regular typeface. It contains the current corporate PUA characters that have been approved by NRSI. The definitions for these PUA characters are preinstalled by FieldWorks, so they can be used in any writing system that needs them. See c:\Program Files\SIL\Fonts\DoulosSIL\DoulosSIL4.100FontDocumentation.pdf for details on the Doulos SIL font, including the font features it supports.
Note: PUA character definitions are only supported by FieldWorks or any other applications using our version of ICU. Other programs can display the glyphs, but cannot process the code point semantics.
- Charis SIL is a serif font (similar to Bitstream Charter) designed specifically for quality printing on laser printers. It contains a repertoire of glyphs similar to Doulos SIL, but adds some extra glyphs for italic and bold italic typefaces. It comes in regular, bold, italic, and bold italic faces. See c:\Program Files\SIL\Fonts\CharisSIL\CharisSIL4.100FontDocumentation.pdf for details on the Charis SIL font, including the font features it supports.

Users can check the NRSI Web site at

http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&cat_id=FontDownloads for additional Unicode fonts as they are developed. FieldWorks also installs the Ezra SIL font for Hebrew and the Galatia SIL font for Greek. NRSI is working on a font that will display tone in a variety of ways.

Code2000 is a shareware font that covers most of the basic plane covering Unicode 5.1. According to Bob Hallissy in the NRSI Resource Library, “SIL International has a site license, which means that the Code2000 font may be copied and installed on any SIL computer, but this license does not include permission to redistribute or modify it.” This font can be downloaded from http://www.code2000.net/code2000_page.htm. This site also provides Beta test versions of Code2001 that covers parts of Plane One and Code2002 that covers parts of Plane Two.

Users can freely download many Unicode fonts at www.unifont.org/fontguide/. Many other Unicode fonts are also available commercially.

Note: FieldWorks can actually use non-Unicode “hacked” fonts, but this is highly discouraged! It means users must store hacked data rather than real Unicode data.

5 Doulos SIL or Charis SIL font installation problems

The FieldWorks 5.2 installer embeds the latest NRSI font installers for Doulos SIL and Charis SIL. If the font installer detects the same or newer version of the font, it will not overwrite the font when run from the FieldWorks installer. Under some circumstances this can result in the font not being installed. For example, if a user deletes the font directly from the Windows\Fonts folder, the font installer doesn’t currently detect this and thinks the font is still there.

If you find that Doulos SIL or Charis SIL are not available after a FieldWorks installation, you can force the installation by downloading the font installer(s) from the NRSI site ([Doulos SIL](#), [Charis SIL](#)) and run them manually. The installer will probably give a message “A newer or same version of DoulosSIL is already installed. Do you want me to force the installation of this font package?” If you click Yes to this dialog, it will force installation of the new font. Even after this, one Vista machine required a reboot before the new font was available to programs.

6 Saltillo problems in FieldWorks 5.4

Unicode 5.1 added new code points for
A78B LATIN CAPITAL LETTER SALTILLO
A78C LATIN SMALL LETTER SALTILLO

Since FieldWorks 5.4 is still using ICU 3.6, FieldWorks does not have any of the semantics for these new code points. We hope to update to ICU 4.0 for the next release which does include this. FieldWorks 5.4 does include these two code points in WordFormingCharOverrides.xml as wordforming, so interlinearization should work with these code points. However, without semantics, FieldWorks has no way of recognizing these two code points as lowercase/uppercase equivalents, and our current UI does not allow the user to set the semantics for these two characters since they are not in the PUA area. So if you are using both cases, you may bump into a problem.

FieldWorks 5.4 installs Doulos SIL 4.102 which does not include these code points. Times New Roman also does not include these two code points. So if you insert the characters in FieldWorks using either of these fonts, they will be displayed as empty boxes. This can be fixed, however, by downloading and installing the latest Doulos SIL

4.104 from the NRSI site:

http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=DoulosSIL_download. This newer version of Doulos SIL defines these characters, and FieldWorks will display the characters fine once you've installed this font. You'll also want to download and install the equivalent Charis SIL version as well.

There is one additional problem if you plan to use character checks in TE. FieldWorks 5.4 does not allow a person to manually add the Saltillo to valid characters. The easiest way to get the Saltillo into valid characters is to edit the xml language definition file in %ALLUSERSPROFILE%\Application Data\SIL\FieldWorks\Languages. If you add the Saltillo (or any other character) to this file, the next time you open a FieldWorks database, it will load this into the valid characters field in the database.