Introduction to Module 1: Converting Word Docs to FrameMaker

FrameMaker's tools make it easy to globally control formatting and some content. The process of implementing FrameMaker in a publishing environment usually requires that content created in another application, such as Microsoft Word, be converted to FrameMaker. Word's ubiquity virtually ensures that some of your company's content will be converted from Word to FrameMaker. Another common situation is that the technical publications department has FrameMaker licenses but the rest of the company uses Word. Thus content from a Subject Matter Expert (SME) is created in Word, and technical writers must convert, edit, and incorporate that content in their FrameMaker documents.

Even though converting documents is a straightforward process, there are many points worthy of consideration before you simply open your Word documents in FrameMaker.

Objectives

- Learn the general principles of format conversion
- Determine which content should be altered before conversion
- Examine formatting in a Word document to determine the best course of action to convert it to FrameMaker
- Use Import Text Flow by Copy and other options to remove overrides
- Learn processes to apply formats
- Revise content after conversion
- Replace manual formatting with globally controllable formats

Format Translation Processes

In this section, you'll learn about two general processes used to translate content to FrameMaker. The main process used in this module is a Mapping Table.

Format Mapping Table Tools

By "format mapping" we refer to the assignment of one format to another format. There are several different implementations of mapping tables that can be used to convert legacy formats to FrameMaker.
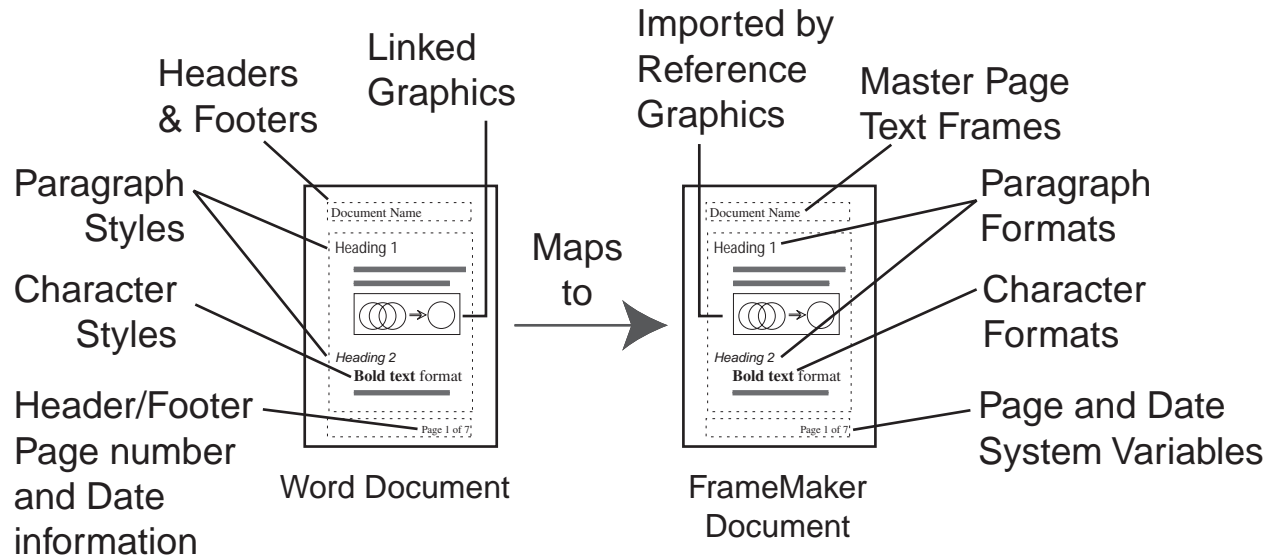
Open Legacy Document in FrameMaker

Opening a legacy document in FrameMaker automatically maps the formats to Paragraph Formats or Character Formats; unfortunately, any format overrides in the source document are retained in the converted document.

For example, you could convert your Word documents simply by opening them in FrameMaker, if the source Word documents met all the following criteria:
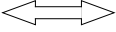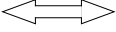
- used Word Styles consistently,
- had very few, if any, format overrides,
- and you didn't have a FrameMaker template

1 Converting Docs to FM

You would probably opt for this conversion process because every Style in Word is converted to a Paragraph Format or Character Format in FrameMaker, Headers and Footers are replicated as Header or Footer background text frames on Master Pages, page number and date information is converted to page and date system variables, and linked graphics are converted to graphics imported by reference and placed in anchored frames.

Headers & Footers

Linked Graphics

Imported by Reference Graphics

Master Page Text Frames

Paragraph Styles

Character Styles

Header/Footer Page number and Date information

Maps to

Paragraph Formats

Character Formats

Page and Date System Variables

Word Document

FrameMaker Document

The table below lists the automatic mappings between FrameMaker and Word and the directions in which the mappings work.

Word	Mapping Direction	FrameMaker
Paragraph Style	⟺	Paragraph Format
Character Style	⟺	Character Format
Tables (Word has no Table Formats)	⟹	TableTag (translated tables are tagged with Table)
Autotext	⟹	Plain text
Header/Footer and any text in them	⟺	Background Text Frame on Right/Left Master Page(s)
Page # in Header/Footer	⟺	CurrentPageNum Variable
Date info in Header/Footer	⟺	Date Variable
Text in Header/Footer	⟺	Text in background text frame(s) on Right/Left Master Page(s)
Index Entry Text	⟺	Index Markers

Word	Mapping Direction	FrameMaker
TOC are separate files in FrameMaker, and thus can't really be converted to Word. You'll have to recreate them in Word	None	TOCs and Indexes in FrameMaker are separate generated files. Word TOCs and Indexes are best removed before conversion and rebuilt in FrameMaker
Text (Word has no Conditional text)	⇐	Conditional Text
No Variables in Word Autotext does not work the same as Variables	None	Variables
Autotext	⇒	plain old text
Pasted graphics	⇐⇒	pasted graphics in anchored frame
Linked graphics	⇐⇒	graphics imported by reference anchored frames
Word Cross-Reference	⇐⇒ up to a point	FrameMaker Cross-Reference
Page #	⇒	Imported Format PageNum
Heading Text	⇒	Imported Format ParaText
Numbered Item	⇒	Imported Format ParaText (numbered Cross-References are not converted to numbers in FrameMaker)
Numbered Paragraphs	⇒	The first converted paragraph may have the counter initialized as n=1 and subsequent counters as n+ or may have each counter set to a specific number, as in n=1, n=2 and so on
Are not reliably translated	⇐	Numbered Paragraphs

Create and Apply Formats

Create and Apply Formats is a FrameMaker utility that examines the current document for overrides. Paragraphs and characters with hard coded formatting or formats with overrides are given new Paragraph or Character format names based on the original Paragraph Format or Character Format name. The Paragraph Catalog of a document processed with Create and Apply Formats displays format names such as Body1, Heading 12, or Normal29.

1 Catalog ✕	**1 Catalog** ✕	**1 Catalog** ✕
Body Text	Body	Normal15
Footer	Body Text	Normal16
Header	Body Text1	Normal17
Heading 1	Body1	Normal18
Heading 2	Footer	Normal19
Heading 3	Footer1	Normal2
Index 1	Header	Normal20
Index 2	Heading 1	Normal21
Index 3	Heading 11	Normal22
Index 4	Heading 12	Normal23
Index 5	Heading 2	Normal24
Index 6	Heading 3	Normal25
Index 7	Index 1	Normal26
Index 8	Index 2	Normal27
Index 9	Index 3	Normal28
Index Heading	Index 4	Normal29
Normal	Index 5	Normal3
TOC 1	Index 6	Normal30
TOC 2	Index 7	Normal31
TOC 3	Index 8	Normal4
TOC 4	Index 9	Normal5
TOC 5	Index Heading	Normal6
TOC 6	Mapping Table Cell	Normal7
TOC 7	Mapping Table Cell1	Normal8
TOC 8	Mapping Table Cell2	Normal9
TOC 9	Mapping Table Cell3	TOC 1
	Mapping Table Title	TOC 11
	Normal	TOC 2
	Normal1	TOC 3
	Normal10	TOC 4
	Normal11	TOC 5
Delete...	Normal12	TOC 6
	Normal13	TOC 7
Paragraph Catalog	Normal14	TOC 8
before using Create	Normal15	TOC 9
and Apply Formats	**Delete...**	**Delete...**

Paragraph Catalog after using Create and Apply Formats

Paragraph Tools and Character Tools

Paragraph Tools and Character Tools from Silicon Prairie Software are third-party APIs for FrameMaker. (For more information, see *Software* on page 281.) Paragraph Tools and Character Tools are pure mapping tools that allow users to specify one or more Paragraph Formats to be mapped to another

Paragraph Format and one or more Character Formats to be mapped to another Character Format. These tools are best used after using Create and Apply Formats to rename all the overridden formats with names based on the original format name.

Using Find/Change to Apply Formats

Even though Find/Change is not strictly a mapping table, by searching for a string of text, you can paste a Paragraph format to the text; this is a big time saver for applying autonumber paragraph formats because a second pass with Find/Change can remove the manually typed numbers or text.

Filtrix

Filtrix is a third-party translation application. Filtrix has extensive conversion filters and can convert many legacy formats to other formats, including FrameMaker MIF (Maker Interchange Format). We recommend that you use Filtrix only to convert Interleaf files to FrameMaker because Filtrix improperly converts Index entries from Word to FrameMaker. As with other applications, Styles are mapped to Paragraph Formats and Character Formats.

Plain Text Markup

The second conversion process is used with plain text. Plain text contains no formatting codes to describe how paragraphs or characters should display. In order to convert plain text to FrameMaker paragraph formats and character formats, the source text must have some formatting codes. To do this, prior to translation, insert MML (Maker Markup Language) codes in the text using PERL or Omnimark. This requires a skilled PERL or Omnimark programmer to add markup to plain text. This process is beyond the scope of this manual and is only applicable when translating large amounts of content because of the extensive amount of work required to create, test, and implement the conversion program.

Other Source File Formats

Microsoft Word is the format that is most commonly translated to FrameMaker. This section provides some pointers about converting source files from one of these other formats; Interleaf, PageMaker, QuarkXPress, InDesign, WordPerfect, Ventura Publisher, HTML, SGML, or XML.

Interleaf. Use the FrameMaker conversion utility (open IAF source text in FrameMaker) for Interleaf to FrameMaker table conversions, because Filtrix does not do as good a job of converting Interleaf tables as FrameMaker does.

Use Filtrix for text conversion because Filtrix converts Interleaf paragraph and character formats to FrameMaker paragraph formats and character formats better than FrameMaker. Using this two-tool approach on one Interleaf file will create two sets of converted files; blend the two converted files into one by replacing FrameMaker converted Tables into Filtrix text conversion.

PageMaker/QuarkXPress/InDesign. PageMaker/QuarkXPress/InDesign are page layout tools and handle content differently than FrameMaker. Many times the content is linked through the entire document and thus can be converted as a whole to a text instance. However, graphics are handled differently in those three applications and cannot be converted to FrameMaker as easily. Quark and InDesign can export content to XML; if so, use the Structured FrameMaker interface and convert to a structured FrameMaker document and remove structure. Then apply paragraph formats.

Note: FrameMaker version 7.1 and later can convert PageMaker and QuarkXPress documents to FrameMaker format.

WordPerfect. WordPerfect is similar to Word, so most of the comments for Word also apply to WordPerfect.

Ventura Publisher. Some of the formatting codes stored in Ventura could be manipulated to resemble MML markup and easily interpreted upon opening the document in FrameMaker. This is beyond the scope of this manual.

HTML. Filtrix does not seem to convert HTML to MIF, even though it is one of the options listed.

Another way to convert HTML to FrameMaker is by changing the HTML instance to XML by changing the declaration line from HTML to well-formed XML and then using the Structured FrameMaker interface.

XML and SGML. To convert XML or SGML instances, you must use the Structured FrameMaker interface.

Converting Word to FrameMaker

In this section, you'll convert a Word document to FrameMaker using a combination of FrameMaker's tools and Silicon Prairie Software's Paragraph Tools.

Analyzing Source Documents

The first step in the overall process is to analyze the formats in your source documents. The easiest method to inspect the formatting of a Word document is to open the document in Word, and take note of the paragraph styles as noted in the Formatting Toolbar.

After you have ascertained the source formatting, you'll do a trial conversion to see what other problems you might have.

| Exercise 1 | **Analyzing the Word Document** |
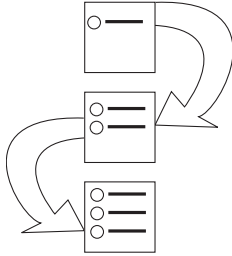
In this exercise, before you actually perform any conversion tasks, you will analyze the original Word document to see how the document was constructed and how consistently Word's Styles were used.

1 Start Microsoft Word. If you don't have Word, skip to step **6**.

2 From within Word, open `../fm201/mod1/test-doc2.doc`.

3 In `test-doc2.doc`, press **Down Arrow** to move your cursor through the source document.

4 As you cursor through the document, look at the Formatting Toolbar and take note of the Styles names.

Many paragraphs are formatted with Normal and yet have dissimilar formatting. Also notice that many words and phrases are italicized or bolded without using a Character Style. Some of the product names have a Character Format ProductName that does not have an override.

5 Close `test-doc2.doc`.

6 Start FrameMaker. If FrameMaker is already running, skip to the next step.

7 From within FrameMaker, open `../fm201/mod1/test-doc2.doc`. In the information dialog box that displays, click **OK**.

8 Go to page 2. At the top of page 2, click on the clouds graphic in the anchored frame next to A.I. Enzymes™.

The object handles are barely visible.

9 From the Graphics menu, select **Object Properties**.

The Object Properties dialog box displays information about the selected graphic.
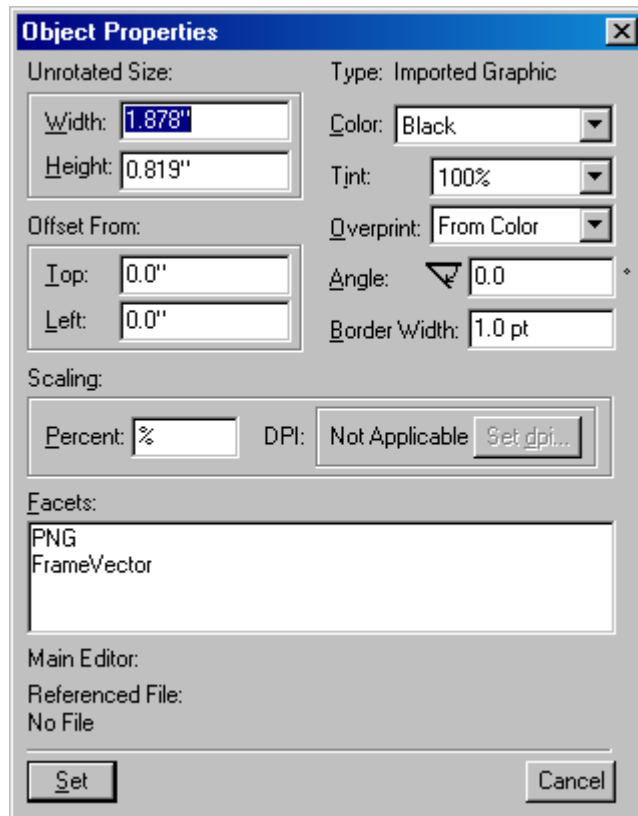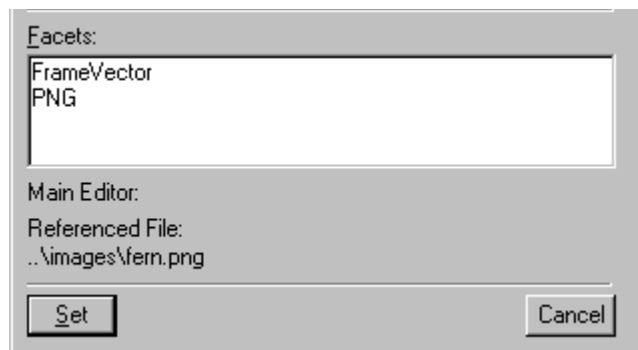

Notice that even though this object is an imported graphic, no file is referenced; this graphic was inserted in the Word source document using the Insert command rather than the Link to File command. Inserted graphics from Word documents are converted to Import by Copy graphics in FrameMaker. This is effectively the same as copy and paste; no link exists back to the original file.

10 Click **Cancel** to close the Object Properties dialog box.

11 At the bottom of page 2, click the fern graphic in the anchored frame next to A.I. Formula.

12 From the Graphics menu, select **Object Properties**.

The Object Properties dialog box displays information about the selected graphic.

```
Facets:
FrameVector
PNG


Main Editor:
Referenced File:
..\images\fern.png
─────────────────────────────────
  Set                      Cancel
```

This graphic was inserted in the Word document using Link to File. When the file was converted to FrameMaker, the linked graphic was converted to a graphic imported by reference.

Note: Graphics in Word documents are converted and placed in anchored frames regardless of the placement or insertion method.

Word can place graphics in documents by pasting or by linking to a file on the system. FrameMaker provides the same placement choices: pasting (Copy and Paste or Import by Copy); and linking (Import by Reference). When you convert a Word doc to FrameMaker, pasted graphics remain pasted in the document and linked graphics are converted to referenced graphics. We recommend that you replace pasted graphics with graphics imported by reference.

The first decision you'll have to make is: should you leave the pasted graphics in the Word documents or remove them prior to conversion? In some ways, it is easier to remove graphics before you convert them, but there is an important consideration: by leaving pasted graphics in, you can more easily identify the replacement graphic and where it belongs.

Even though it is possible to use Find and Replace in Word to replace graphics with a placeholder, you would have to manually type in the graphic pathname because Word does not retain pathnames to pasted graphics.

Remember that while the graphics are in the converted files, your file size is increased. For example, a Word document with a simple PNG graphic (which is 4KB in size) was 56K in Word. When that document was opened in FrameMaker and saved, the filesize went up to 304K. The same file without the graphic, in Word was 51K and was 66K in FrameMaker.

Then when the 304K FrameMaker file has the pasted graphic replaced with an Imported by Reference graphic, the file size reduces to 68K.

13 Close `test2-doc.fm` without saving.

Summary By inspecting the source document, you can estimate the work required to convert it to FrameMaker and what techniques or methods you might use to convert it. You also gain an understanding of what content may not be needed and could be deleted before you begin.

Document Conversion

Now that you have inspected the Word source document, you must consider the best method to convert it to FrameMaker.

The easiest conversion method is to do what you just did: open the Word doc in FrameMaker and allow automatic conversion. However, there are some issues you should consider before choosing this method.

1 Word documents may contain some formats that you don't want in your FrameMaker documents.

   Even if the Word document used Styles consistently throughout, you may still not want to convert all the Styles stored in the Word document. When you open a Word document in FrameMaker, everything in the Word document is converted. By comparison, when you start with a template and import by copy, you can remove any format overrides and have a much simpler conversion process.

2 You'll have to clean up the converted Master Pages or reapply the correct Master Pages from the template.

   If you open the Word document in FrameMaker, any Word Headers and Footers are converted to background text frames on Master Pages. If your template has specific Master Page layouts, you'll have to import them from the template, which adds another step to the process.

   If you create a blank document from your template and import the Word document by copy, anything in the Word Headers and Footers is ignored and the correct Master Pages are used.

The best method we have found for converting Word documents to FrameMaker is to use Import by Copy and Reformat Using Current Document's Formats.