

Microsystems - THE document experts  
3025 Highland Parkway, Suite 450  
Downers Grove, IL 60515  
Phone: 630-598-1100  
Fax 630-598-9520  
<http://www.microsystems.com/>



# **Word 2002 from a Document Forensics Perspective**

*Adding a new version of Word to your  
Document Creation, Conversion and  
Collaboration Work*

**August, 2001**

## Contents

<b>Introduction .....</b>	<b>1</b>
<b>Converting WordPerfect Documents into Word 2002.....</b>	<b>1</b>
<i>Document Object Issues .....</i>	<i>3</i>
<i>Section Object Issues.....</i>	<i>3</i>
<i>Paragraph Object Issues.....</i>	<i>4</i>
<i>Character Object Issues.....</i>	<i>4</i>
<i>Special Issues .....</i>	<i>5</i>
<i>Field Code Issues.....</i>	<i>6</i>
<b>Documents from Lower Versions of Word Opened in Word 2002 .....</b>	<b>6</b>
<i>Document Object Issues .....</i>	<i>6</i>
<i>Section Object Issues.....</i>	<i>8</i>
<i>Paragraph Object Issues.....</i>	<i>9</i>
<i>Character Object Issues.....</i>	<i>9</i>
<i>Footnote Issues.....</i>	<i>9</i>
<i>Table Issues.....</i>	<i>10</i>
<i>Field Code Issues.....</i>	<i>10</i>
<b>Word 2002 Documents Opened in Lower Versions of Word .....</b>	<b>10</b>
<i>Document Object Issues .....</i>	<i>11</i>
<i>Section Object Issues.....</i>	<i>11</i>
<i>Paragraph Object Issues.....</i>	<i>11</i>
<i>Character Object Issues.....</i>	<i>13</i>
<i>Footnote Issues.....</i>	<i>13</i>
<i>Table Issues.....</i>	<i>13</i>
<i>Field Code Issues.....</i>	<i>14</i>
<b>New Features in Word 2002 of Use in Document Forensics.....</b>	<b>14</b>
<i>Document Recovery.....</i>	<i>14</i>

<i>Paste Options Button</i> .....	15
<i>Goodbye Jason Tab</i> .....	16
<i>Styles and Formatting Task Pane</i> .....	17
<i>Reveal Formatting Task Pane</i> .....	18
<b>Conclusion</b> .....	<b>18</b>

## Introduction

Word 2002 is here. Law firms, for the most part, seem to be circling it on tiptoe, not willing as yet to make the leap of faith. But many of their clients will have no such reluctance, and will shortly purchase it, create documents with it, and open their Word 97 and Word 2000 (not to mention WordPerfect) documents into it. It won't be long before Word 2002 documents begin appearing in an email in-box near you!

What issues will this create? How will you prepare for this inevitable moment? Let's break it down. From a document forensics perspective:

- What should we be aware of when documents created in lower versions of Word are opened in Word 2002?
- What should we be aware of when Word 2002 documents are opened in lower versions of Word?
- What should we know about Word 2002 documents generally?

We consider it premature to attempt an exhaustive catalogue of all the potential issues, since Word 2002 is so newly released, and not many law firms have had an opportunity to "bang on it." However, we *can* consider the typical potential problem "hit parade." As usual, the simplest and best way to approach understanding Word is to consider it within the framework of the object model. Thus, we'll discuss:

- The definition of the Normal style and Compatibility Options.
- The components of sections.
- Paragraph formatting, either direct or via a style.
- Character formatting.
- Footnotes.
- Tables.
- Fields.

## Converting WordPerfect Documents into Word 2002

Even in law firms that have been in Word for years, there are still those occasions when an attorney will want an old archived WordPerfect document to use as a precedent document for a new deal; or when someone in the firm receives a WordPerfect file as an attachment to email. One way or another, we'll probably be converting WordPerfect documents into Word forever.

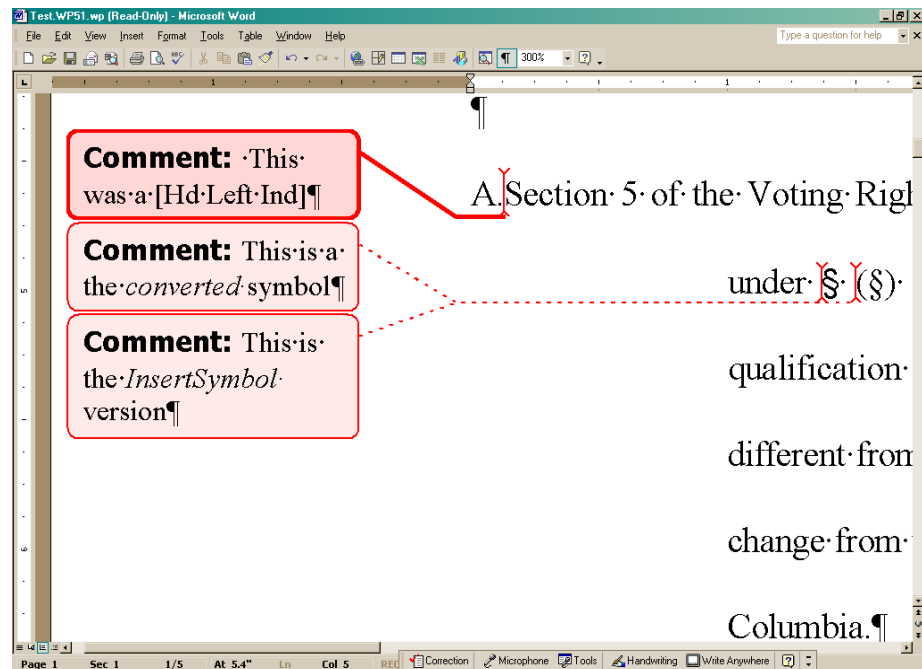
Back in April, 2001, Microsoft released an updated Conversion Filter Pak for use with all versions of Word. These same revised filters are delivered when

you install Office XP/Word 2002 and choose to install the Text Converters for WordPerfect. Two specific changes in the filters may alter results you were able to attain before these changes were made: the method used to preserve WordPerfect's special symbols, and handling of [Hd Left Ind] (Hard Left Indent) codes.

- **New handling of WP Typographic Symbols.** In converted WordPerfect documents, the special characters do not convert properly. The new Word 2002 converters convert special characters to hidden {symbol} fields that contain the font and glyph information for each character.

In this way, special characters are always retained, regardless of the font of the surrounding text. You will not see these symbol fields if you turn on field codes in Word. While this is probably okay as long as the document stays in Word 2002, we suspect conversions of these hidden fields back to RTF (the format widely used by electronic comparison software), or to WordPerfect 5.1, or even to earlier versions of Word may present interesting artifacts somewhere down the line.

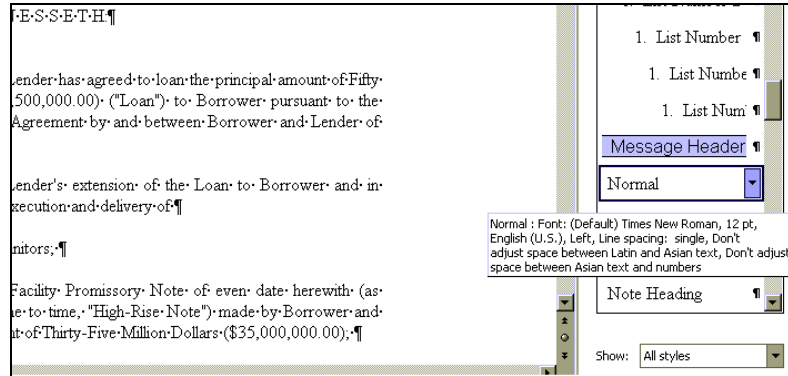
- **Hard Left Indents Stripped from the Document.** Whereas the old filters delivered a tab marker in place of the original [Hd Left Indent] codes, the revised filters appear to be stripping them from the document altogether. When WordPerfect documents where these codes were used are converted to Word, all indents collapse, or conversely, turn into hanging indents (depending on the order of the WordPerfect codes). Since the code is eliminated during conversion, there's no way to automate reinsertion of the tab setting or its marker—the text simply merges together.



Examples of the converter issues results

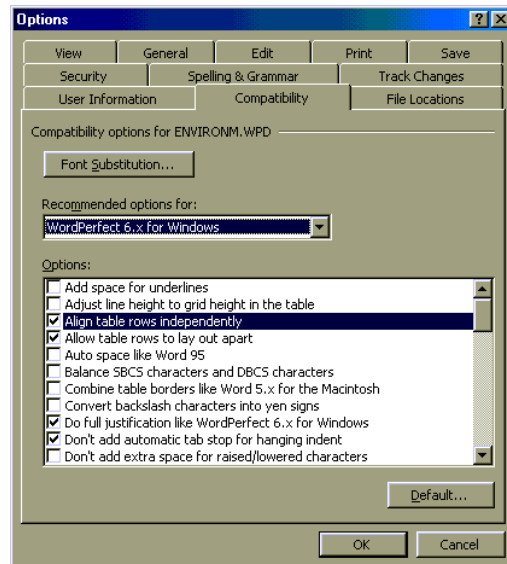
## Document Object Issues

**Poorly defined Normal style.** As always (See our document [Fixing Broken Word Documents](#) for more on the subject.), documents opened into Word 2002 from WordPerfect exhibit a wide variety of problems. Files acquire a Normal style with extremely poor formatting, as defined below:



*The Normal style definition in Word 2002 of a File>Opened WordPerfect document*

**Inappropriate Compatibility Options.** Further, the document's Compatibility options are set as described in the following graphic. This indicates Word's (misguided) attempt to act as if it were WordPerfect in order to display the document.



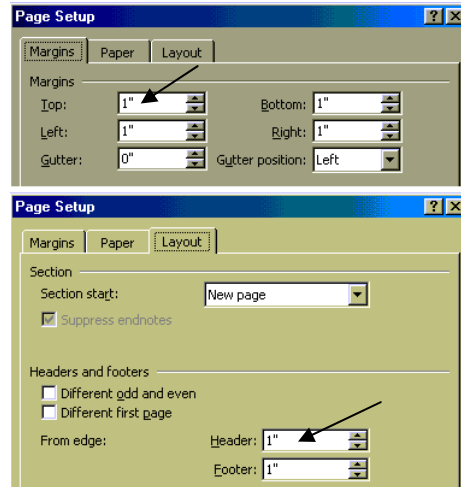
*The Compatibility Options for a File>Opened WordPerfect document in Word 2002*

## Section Object Issues

The good news is we have not uncovered any *new* issues. The bad news? All section issues that we've dealt with for years still remain. For example, documents opened into Word 2002 from WordPerfect:

- Still exhibit multiple unneeded continuous section breaks.

- Still have headers and footers placed at the same position as the document text's top and bottom margins.



*Note the top margin and header settings are identical*

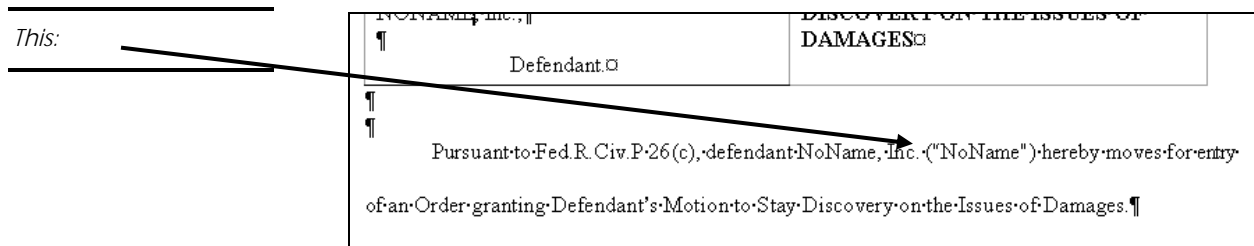
- Still have headers and footers styled with Normal, not Header or Footer.

## Paragraph Object Issues

WordPerfect documents opened into Word 2002 continue to have the issues we've been dealing with since Word 97. When you boil these down at the paragraph level, you're looking at tabs performing inappropriate tasks (like centering), extra empty paragraph marks, directly applied paragraph formatting, and the dreaded {SEQ} field code numbering. The solution? Style your troubles away!

## Character Object Issues

As with WordPerfect documents in earlier versions of Word, watch for printer fonts (CG Times, Courier, etc.), and replace them with their True Type equivalents (Times New Roman, Courier New, etc.). WordPerfect Typographic symbols continue to provide endless hours of entertainment. They remain dangerous because they are difficult to see—until you change the document's font.



*Note "strange-looking" quote marks*

Becomes this:

Defendant.	DAMAGES
Pursuant to Fed.R.Civ.P.26(c), defendant NoName, Inc. (ANoName@) hereby moves for entry of an Order granting Defendant's Motion to Stay Discovery on the Issues of Damages.	

Upon modification of the font, note the quotation marks are revealed actually to be text, rather than symbols, in the new font

## Special Issues

- Word 2002 documents from WordPerfect will still experience occasional difficulties with Footnotes until the document is properly styled.
- Table borders that don't display at the bottoms of pages. When this happens, it is usually because the table is still operating on the basis of four dimensions (top, bottom, left, and right), while Word (starting in Word 97) is operating in *six* (top, bottom, left, right, *inside*, and *outside*). Thus, a border that's defined as a bottom is actually being asked to perform the task of an outside. It can't, and therefore isn't displayed.
- Table shading that crashes comparison software. If you're still using CompareRite, this is an on-going problem. In DeltaView, this appears not to be an issue.
- Extra, empty paragraph marks vertically centering text in a cell. In Word, we don't need to use them, since we can explicitly format text in cells to be centered.

To correct these issues:

- Select the offending table, and define it to have grid borders (this places the appropriate border on each cell). Once this is accomplished, you can redefine the table's borders to whatever custom format you like.
- Similarly, to fix a shading problem, select the shaded cells, row, or column and define them to have No Fill. Then, you can redefine them to have the appropriate shading attributes applied Word's way.
- Replace the extra empty paragraph marks with space before or space after, center, or bottom formatting. Unfortunately, this is not a particularly simple process; since, while you can certainly apply a style for the space before and after, you must manually delete all the unwanted paragraph marks from each cell!

## Field Code Issues

When a document comes from WordPerfect, Word 2002 has essentially the same field code issues as previous versions. The most common one relates to paragraph numbering. Word turns WordPerfect numbering automation into field codes, usually either {SEQ} or {LISTNUM}. Interestingly, neither of these is wrong, per se, so much as not Word's best practice. One *can* do automatic paragraph numbering with field codes, but it is time-consuming and labor intensive, and calls for the user to edit the variable in each field code for each level of the outline. As in earlier versions of Word, it is best to turn the field codes to text using Ctrl+Shift+F9, apply outline numbered Heading styles to those paragraphs, and finally, delete the text numbers.

## Documents from Lower Versions of Word Opened in Word 2002

### Document Object Issues

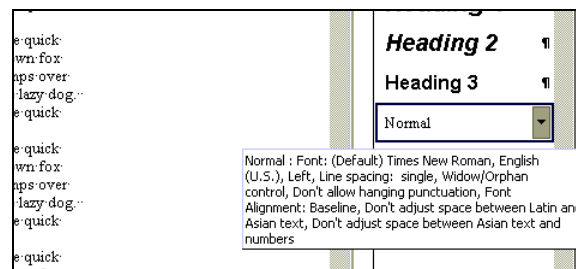
If you've seen our Top-Down Document Design article ([Top-Down Document Design](#)), you know that the document object contains a great many important items. From a "document container" perspective, there are two important issues of which to be aware when Word 97 and Word 2000 documents are opened in Word 2002:

#### Normal Style Definition

We continue to assert that the best Normal style definition is one which describes "plain vanilla" text. The best way to think of Normal is as the foundation upon which the styles you'll actually apply to text are based.

When a Word 97 or a Word 2000 document has a good Normal style definition, it continues to be well defined (and in fact, unchanged) in Word 2002.

Documents coming into Word 2002 from Word 6/95, however, exhibit the type of ugly description you'd expect to see in a WordPerfect file:



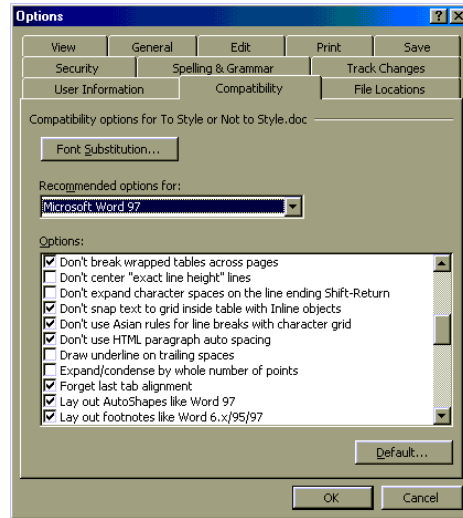
*The Normal style definition of a Word 6/95 file in Word 2002*

Documents with a poorly defined Normal style are disasters waiting to happen. Be sure to use the Organizer (*Format>Style>Organizer*) to copy a good Normal style from another document as soon as you can.

## Compatibility Options

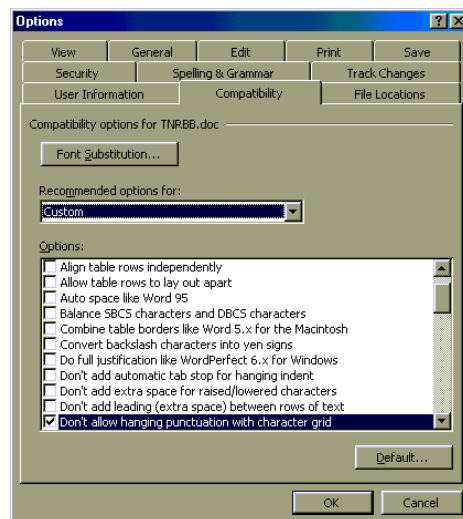
Compatibility options are pretty much as you might expect in documents from Word 97 or Word 2000.

- A Word 97 document opened in Word 2002 reports compatibility options of Word 97. Note particularly the last displayed option—if you leave this option check marked, footnotes separated from their references (a problem that was corrected in Word 2000) will reappear in your Word 2002 document!



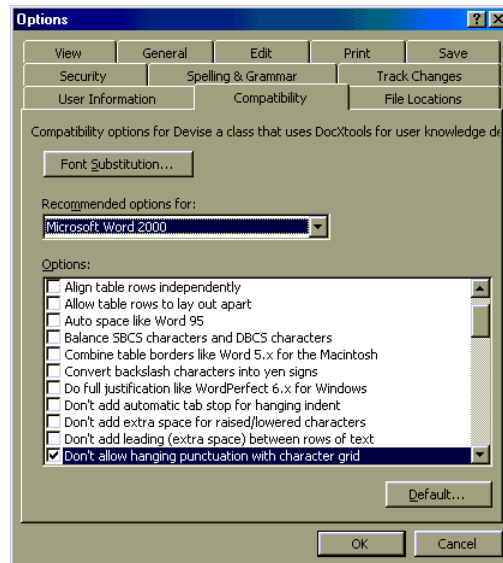
*A Word 97 document in Word 2002*

- A document that has spent time in both Word 97 and Word 2000 reports compatibility options of Custom. While not problematic, per se, this can indicate a document that's not quite sure how it is supposed to look and act.



*A document with both Word 97 and Word 2000 features in Word 2002*

- A Word 2000 document opened into Word 2002 reports compatibility options of Word 2000.



A Word 2000 document in Word 2002

The question, as always, is “why?” If a document is opened in Word 2002, we ought to allow Word 2002 to be itself, rather than to attempt to be some other version of Word. We continue to believe that, in general, a document’s compatibility options should be set to the version of Word in which it will primarily exist. Custom settings are fine as long as the checked options actually perform a needed task, such as “Print Colors as Black on Non-color Printers.” If you’re working in Word 2002, you might not really need custom options like “Lay out Autoshapes like Word 97.” If the checked Custom options seem unnecessary, it is beneficial to change the setting from Custom to the appropriate version of Word. Actually, the only reason to leave these settings alone is if the document actually lives in the lower version and is only “vacationing” in Word 2002!

- If the document’s compatibility options are out of whack, use *Tools>Options>Compatibility Tab* and select the appropriate recommended option.

## Section Object Issues

Documents opened from lower versions of Word can, of course, exhibit any or all WordPerfect behaviors if they *originally* came from WordPerfect. Further, they can have the extremely evil problem of “hidden” sections that potentially contain document corruption. This generally happens when users do not pay attention to a document’s use of “Different first page,” and clone a multiple-page document that contains header or footer content, reducing it to one page, thus hiding the content of the footer.

Properly created and formatted native Word documents, however, do not display any section object problems.

- To remove unwanted section breaks, *Browse* to each section break in turn, and delete those that are unneeded.
- In each section of the document, use *File>Page Setup>Margins Tab* and be sure to set the “From Edge” settings for headers and footers to be some increment *smaller* than the Top and Bottom Margins.
- Visit every section’s header and footer and, if necessary, replace the Normal style with Header or Footer.

## Paragraph Object Issues

As one might expect, directly applied formatting continues to be the most obvious problem at the paragraph level. Word 2002, like its predecessors, will allow users to apply paragraph formatting directly to text; and, like its predecessors, will eventually make you pay for it if you do. We all know this payment is likely to take the form of hours of additional work and missed deadlines or court filings.

It is important to check paragraph formatting—don’t be lulled into a false sense of security simply because a paragraph appears to have a style applied. Use the Reveal Formatting task pane to determine whether (and what!) formatting may be applied *on top* of a paragraph style. Armed with this knowledge, you can determine what changes you might wish to make to the underlying style, or what new style you can create, to accommodate the actual requirements of the paragraph.

## Character Object Issues

At last, some really good news! The odd behavior in outline numbered paragraphs in Word 97, where the trademark symbol (™) and/or the Euro symbol (€) would cause the numbering to lose its sequence, has been corrected. As long as your document will not be spending part of its life in Word 97, you need have no fear of paragraph numbering going bad (at least, not because of these symbols!). Read more about character issues in [The Euro Symbol: Its Use and Known Issues](#) and in [Tricky Times New Roman Bold, Bold](#).

## Footnote Issues

More good news! It appears that the dreaded “footnote reference on one page, footnote text on the next” problem, which was sometimes solved in Word 2000, is truly solved in Word 2002. *Please note* that this fix is only reliable when documents are properly created and styled. This means:

- The footnote should be styled with Footnote Text and Footnote Text More, *not* Normal or any other style.
- Space between paragraphs should be applied using “Space after” attributes in the Footnote Text style.

- The document itself should contain as little direct paragraph formatting as humanly possible—this being the major factor in causing the problem in the first place!

If your footnotes aren't properly formatted, you'll need to

- Visit each one and apply the appropriate style to footnote paragraphs (Footnote Text to the first paragraph of a footnote; Footnote Text More to subsequent paragraphs, if any).
- Delete unneeded, extra paragraph marks.

For more information on footnotes and the issues surrounding them, read [Fractured Footnotes and How to Avoid Them](#).

## Table Issues

When speaking of native Word documents, table issues actually seem to matter more in documents going *from* Word 2002 into lower versions (see below). Word 2002 has no difficulty with tables created and formatted in lower versions of Word.

## Field Code Issues

Word 2002 understands and works with all properly formatted field codes applied to documents created in lower versions of Word. When field codes go bad, it is *usually* because of one or more of these issues:

- A missing opening or closing curly brace [ { or } ].
- A portion of the field name missing ( {STNUM} instead of {LISTNUM}, for example).
- Missing quote marks before and/or after the field variable (i.e., {TC \f c \l \ "2}. In this example, there should be a close quote after the final 2).
- Quote marks surrounding the variable are curly (in proper field codes, quote marks should *always* be straight).

If you would like to learn more about Word's field codes, see [Microsoft Word's Field Codes: Revealing Their {Private} Side](#).

## Word 2002 Documents Opened in Lower Versions of Word

In general, Word 2002 documents are much happier in Word 2000 than they are in Word 97. Many Word 2002 features are supported in Word 2000, but not in Word 97.

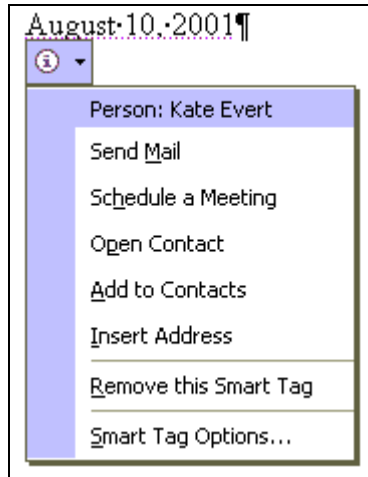
## Document Object Issues

The Normal style is appropriate when a Word 2002 document is opened into either Word 2000 or Word 97. In Word 2000, the Compatibility Options change appropriately to Word 2000, while in Word 97, they become “Custom.” Whether you change this or not depends on where the document will ultimately live. If it will remain in Word 97, change to that Compatibility Option set. Conversely, if it will return to Word 2002, you may wish to maintain the options at Custom.

---

*Note the dotted underlining on the date at right (it shows purple on screen). This denotes an unopened Smart Tag.*

---



*An opened Smart Tag from an inside address*

Smart Tags, a new document object property in Word 2002, do not appear when a document containing them is opened in either Word 2000 or Word 97. If the document was edited in either version, when the file is re-opened in Word 2002, the smart tags, while still present in the document, will no longer appear automatically. However, you can use *Auto Correct > Smart Tags > Recheck document* to re-display the smart tags.

## Section Object Issues

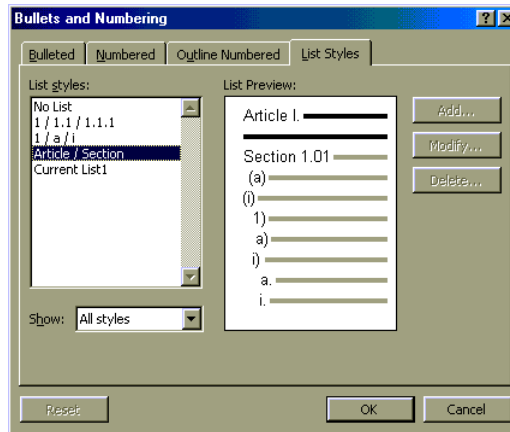
Assuming the sections were created properly in Word 2002, the only feature not supported in Word 97 is a Top gutter setting, which is changed to a side gutter.

## Paragraph Object Issues

Two major changes to the Word 2002 feature set are affected when a document is opened in either Word 2000 or Word 97: List styles and the paragraph separator.

### List Styles

This is a numbering choice in addition to Numbered, Bulleted, and Outline Numbered. When a Word 2002 document has paragraphs formatted using this feature, the styles are changed to outline numbering in both Word 2000 and Word 97.



List styles work best on single level numbers or bullets

## Style separator

This feature is Word's solution to the "hidden paragraph mark" work-around from Word 97 and Word 2000. The Style separator must be added to your toolbar—it isn't part of Word's default installation. Use *View>Toolbars>Customize*, choose the Commands Tab, and then in the Categories list, click All Commands. Scroll through the list in the right-hand pane to find "Insert Style Separator." Drag the icon to the toolbar.

To use the Style separator, you must download and install a patch (there's a bug in the original release). Click here to go to the TechNet article describing the problem; when there, follow the link to the download page.

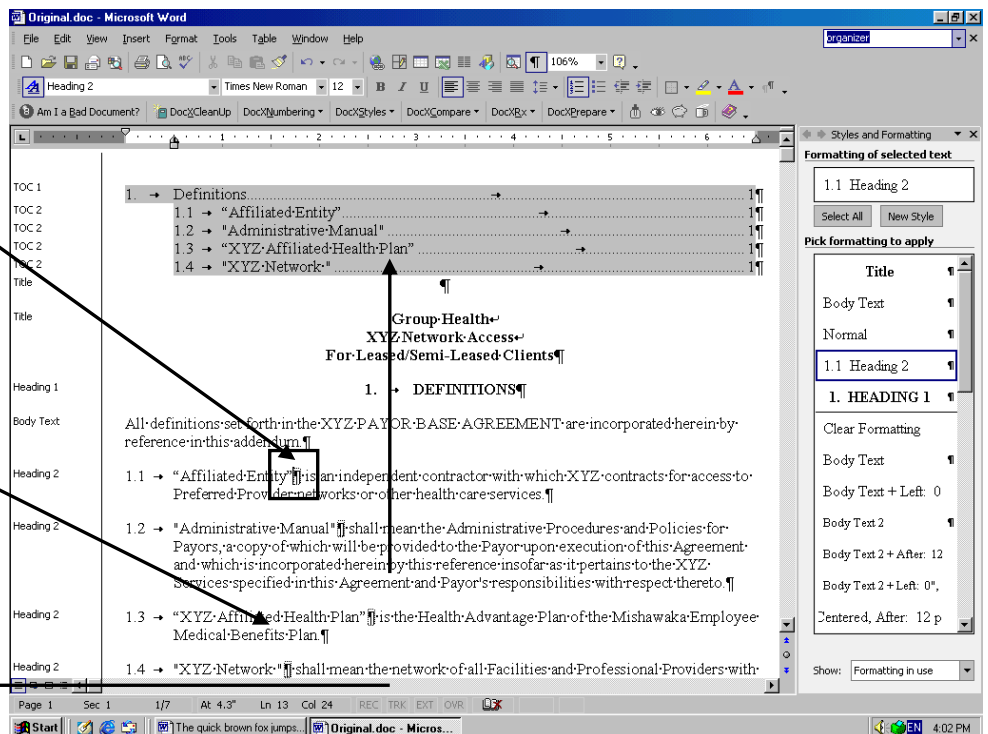
<http://support.microsoft.com/support/kb/articles/Q300/4/28.ASP>.

### The finished product.

The hidden paragraph mark separates the front portion of the paragraph formatted with the Heading 2 style and destined for the TOC.

The remainder of the paragraph is styled with Body Text, facilitating the hanging indent and other paragraph formatting.

When inserted, the TOC correctly displays the desired portion of the paragraph.

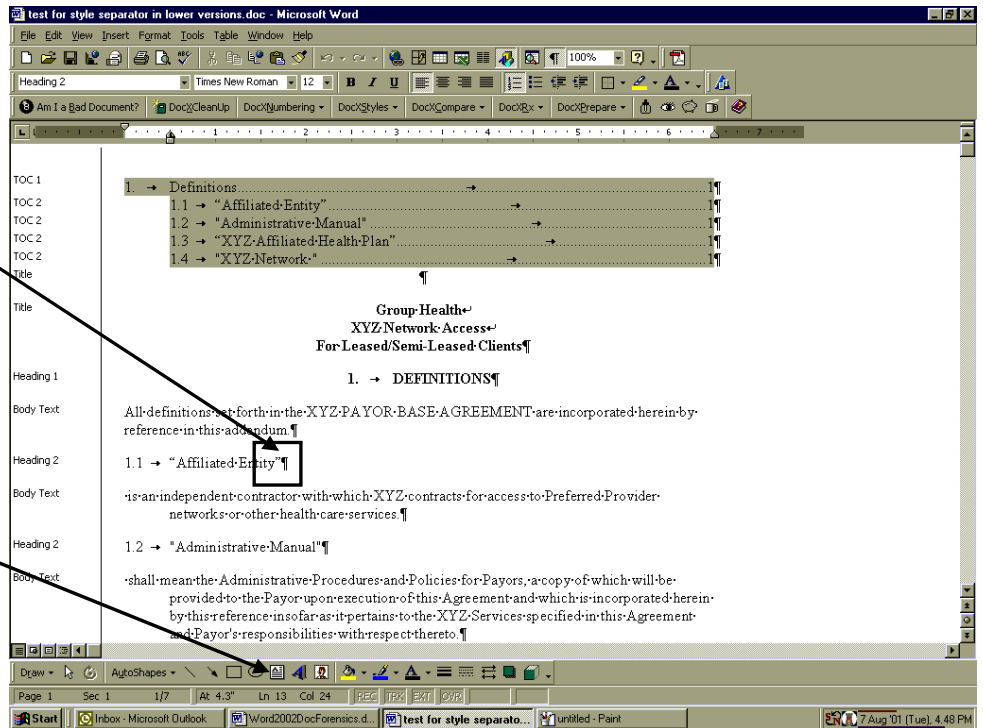


When opened in either Word 97 or Word 2000, the feature is changed to a hidden paragraph mark.

*In Word 2000 (appears the same in Word 97).*

*The hidden paragraph mark separates the front portion of the paragraph formatted with the Heading 2 style and destined for the TOC.*

*The remainder of the paragraph is styled with Body Text, facilitating the hanging indent and other paragraph formatting.*



## Character Object Issues

If you've used Word 2002's Font color (which is 24-bit), or its Underline colors or Underline styles, they're supported in Word 2000, but not in Word 97. There, the Font color is modified to the closest match in 16-bit, Underline colors are changed to black, and several Underline styles, such as Dotted Heavy, are lost.

## Footnote Issues

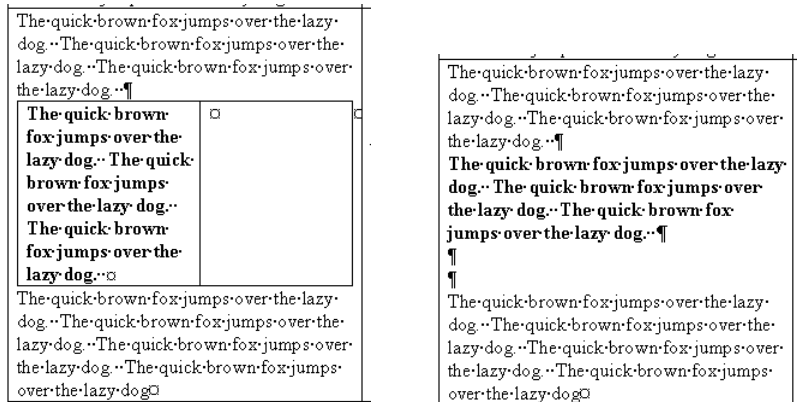
Properly created and styled Word 2002 footnotes move seamlessly into both Word 97 and Word 2000.

## Table Issues

Here, there are several issues to be addressed. First, Word 2000 supports all but one Word 2002 table feature: Table styles. If you've used them in a Word 2002 table, when the document is opened into Word 2000, the table text will appear in the Normal style, although any border or shading attributes defined in the style will be applied.

Between Word 2002 and Word 97, however, there are more difficulties. Table features that were introduced in Word 2000, and are still available in Word 2002, such as nested tables, cell margins, spacing between cells, and others are completely lost when the table reaches Word 97. Word 97 treats

Table styles from Word 2002 the same as Word 2000, above—border or shading attributes are applied, but the table text appears in the Normal style.



*This, in Word 2002 and Word 2000...*

*...becomes this in Word 97*

**Field Code Issues**

We've not been able to uncover any issues with field codes. Providing they are properly created in Word 2002, they operate perfectly in either Word 2000 or Word 97.

**New Features in Word 2002 of Use in Document Forensics**

From a document forensics perspective, here are a few new features that can make document analysis and repair faster and simpler.

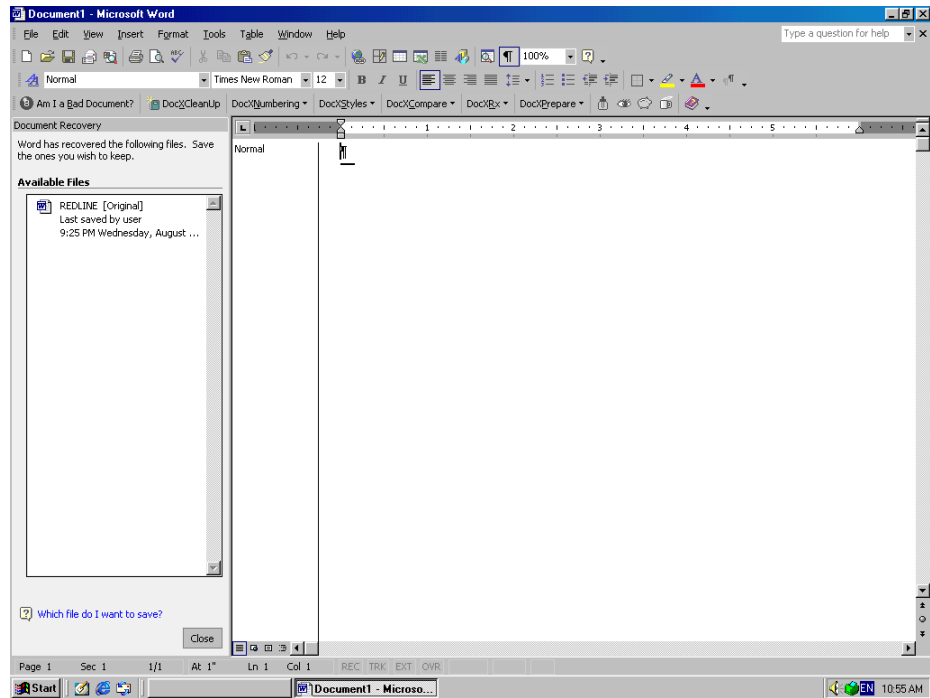
**Document Recovery**

If (or when!) Word encounters a problem and stops responding, this new recovery program can help. Files you were working on are analyzed for errors, and if possible, information in them is recovered. The Document Recovery task pane lists all the files that were recovered when the program stopped responding. Following the name of the file will be a status indicator, which shows what was done to the file during recovery. [Original] denotes the original file based on last manual save, and [Recovered] indicates a file recovered during the recovery process or a file saved during an Auto Recover save process.

---

*We have had significant success using Word 2002's Document Recovery feature to salvage Word 97 and 2000 files which we couldn't open in those versions. Once the file is in Word 2002, it can be saved, reopened in the original version, and its compatibility options changed appropriately.*

---

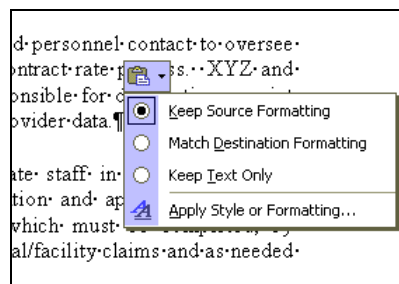


The Document Recovery task pane allows you to open the files that were open when Word crashed, view what repairs were made, and compare the recovered versions. You can then save the best version and delete the other versions, or save all of the open files to review later.

## Paste Options Button

For years, our battle cry has been, “Use *Paste Special>Unformatted Text*” instead of Ctrl+V. We’ve often felt like we were crying in the wilderness, since it was difficult to convince folks that performing four steps instead of one was worth it.

Now, in Word 2002, when you paste text, a button appears where the pasted text is located.

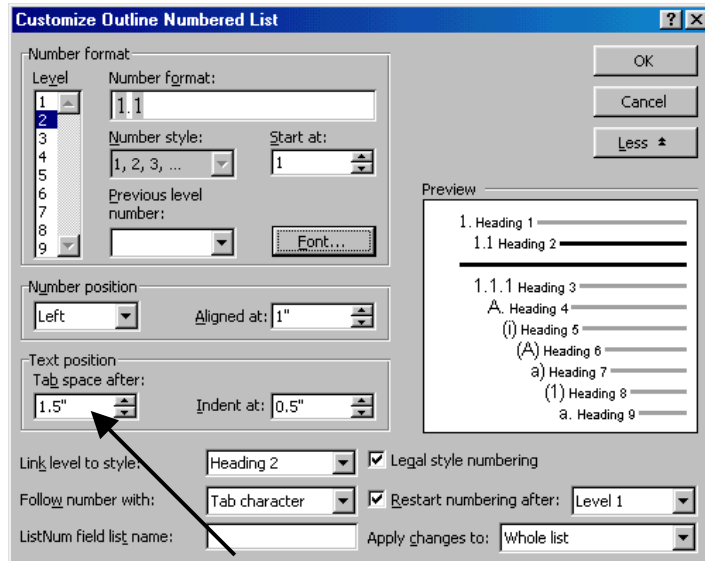


*An opened Paste Options button*

Right-click it, and you’re offered a choice of how this text will be formatted when it’s pasted. You can keep the original formatting (usually the worst idea!), have the text acquire the formatting of the target document, or just paste the text in unformatted. Thus, in one step, you can save yourself endless hassles.

## Goodbye Jason Tab

Unless you've been living on Mars, you've come to loathe the Jason tab, set .25" after the paragraph number, which gets created whenever you visit the customize numbered list dialog box of a numbered Heading style. Microsoft has apparently finally heard our pitiful cries, and has added new functionality.



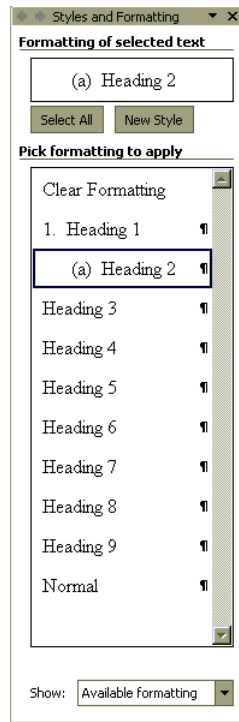
The arrow points out the new *Text Position* area in the Customize Outline Numbered List dialog box (the dialog box you reach when you “dig to China”). By default, Word still wants to position the tab .25" from the number position; however, you can now override that setting here, as we've done for Level 2 in this screen shot.

When these Word 2002 documents are opened in Word 97 and Word 2000, a left tab is set .5" from the number position, facilitating an appropriate distance between the number and text. However, in both Word 97 and Word 2000, any change made in the Customize Numbering dialog box re-introduces the .25" Jason tab to the numbered paragraphs.

You can read more about Word outline numbering in [The Seven Laws of Outline Numbering](#).

## Styles and Formatting Task Pane

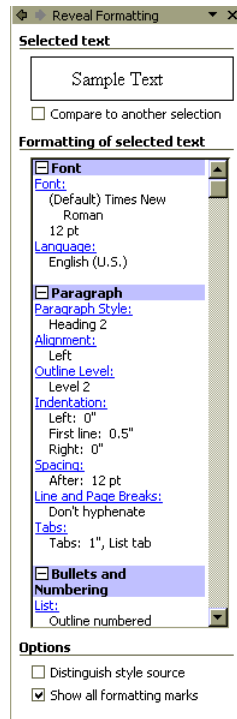
To be honest, our first reaction to this task pane was “Yikes! How do I modify my style?” However, it didn’t take us long to fall deeply in love with task panes in general, but particularly with the Styles and Formatting Task Pane.



In Word 2002, each instance of formatting is “tagged” distinctively, and the Styles and Formatting Task Pane lists each of these distinctive formats individually. You are able, then, to apply, modify, or remove any formatting combination in the document. Here’s an example: say you’re cleaning up a document that appears to be double-spaced, but the Task Pane lists not only double-spacing, but also a format of exactly 24 points. You can select all the paragraphs in the document formatted with exactly 24 points and reformat them with a double-spaced body text style. No need to scroll through the document, moving from paragraph to paragraph, pressing F1, searching for every instance of some particular format! Word automatically finds, selects, and modifies each occurrence.

## Reveal Formatting Task Pane

This task pane more or less takes the place of the “What’s This?” button on the toolbar, which we used to access using Shift+F1. As you can see, it reveals everything you might want to know about the how’s and why’s of a paragraph’s format.



## Conclusion

Because Word 2002 has the same binary file structure as Word 97 and Word 2000, there are few document level issues to be dealt with. This means fewer document corruption issues moving between the three versions.

Many of the same old issues we’ve dealt with since we starting bringing WordPerfect documents into Word are still with us in Word 2002. As long as differences at the binary level exist, conversion from other applications will continue to be a problem. Microsystems’ recommendations haven’t changed: never *File>Open* a WordPerfect document into Word; and if you do, prepare to spend long hours repairing the damage!

Meanwhile, with Word 2002, for all its improvements and additions (and there are many), document forensics skills will still be required. When a document’s numbers are all turning to 1’s, or you see empty square boxes where section marks used to be, or the document won’t compare, you still need to have the knowledge to analyze the problems. We hope this document gives you a head start!

## About Microsystems:

Microsystems is a leader in providing complete, “first draft to final delivery” document lifecycle solutions to the life sciences and legal markets and other document-intensive businesses. The company provides more than 650 client firms and companies with an innovative combination of software products, consulting services and training designed to improve the total document creation, quality control and delivery processes. Microsystems is headquartered near Chicago in Downers Grove, IL.

For additional information please call 630.598.1100 or visit [www.microsystems.com](http://www.microsystems.com)