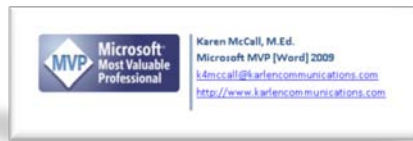




Karlen Communications

Karen McCall, M.Ed.
Microsoft Office Document Imaging Tool

Phone: 1-519-442-2856
E-mail: info@karlencommunications.com
Web: karlencommunications.com



This material copyright 2009 Karen McCall, Karlen Communications.

Table of Contents

Introduction	2
Installing the Document Imaging Tool	2
Upgrading the OCR Software	5
The Document Imaging Tool User Interface	6
Accessibility in the Document Imaging Tool	7
Opening a TIF or TIFF	9
OCR on Scanned Pages	12
Scanning in French or Spanish	17
Microsoft Office Document Imaging Options	18

Introduction

One of the tools often overlooked in the Microsoft Office suites is the "Microsoft Office Document Imaging" tool.

What is this?

The Microsoft Office document Imaging tool lets you either open a TIF or TIFF image of a document or scan a document into the Document Imaging tool and then perform OCR or Optical Character Recognition on the text.

A TIF or TIFF image of a document literally means Tagged Image File Format [TIFF] which is then shortened to "TIF."

The Microsoft Office Document Imaging tool is a free tool that lets someone with a visual or print disability scan pages into the Document Imaging tool and then make the text readable with their adaptive technology such as a screen reader or TTS/Text-to-Speech software. It is not as powerful as OCR scanning software such as OmniPage Pro, TextBridge, ABBYY Fine Reader or tools like the Kurzweil or Freedom Scientific scanning software specifically designed for people with disabilities.

It is, however, a tool at your fingertips for quickly scanning and converting printed pages to text readable using adaptive technology.

Installing the Document Imaging Tool

When you install Office 2007, instead of choosing the "Install" button, choose the "Customize" button. The Microsoft Office Document Imaging tools are not installed as part of the general Office install.

As we will see, there is also the ability to install the OCR tool for three languages: English, Spanish and French. You can choose to install all or whichever ones you need.

After you choose to install a customized version of Office, you will see a dialog with all the components of the Microsoft Office 2007 suite. Items that are shaded have components that will either not be installed at all or will be installed the first time you use the tool.

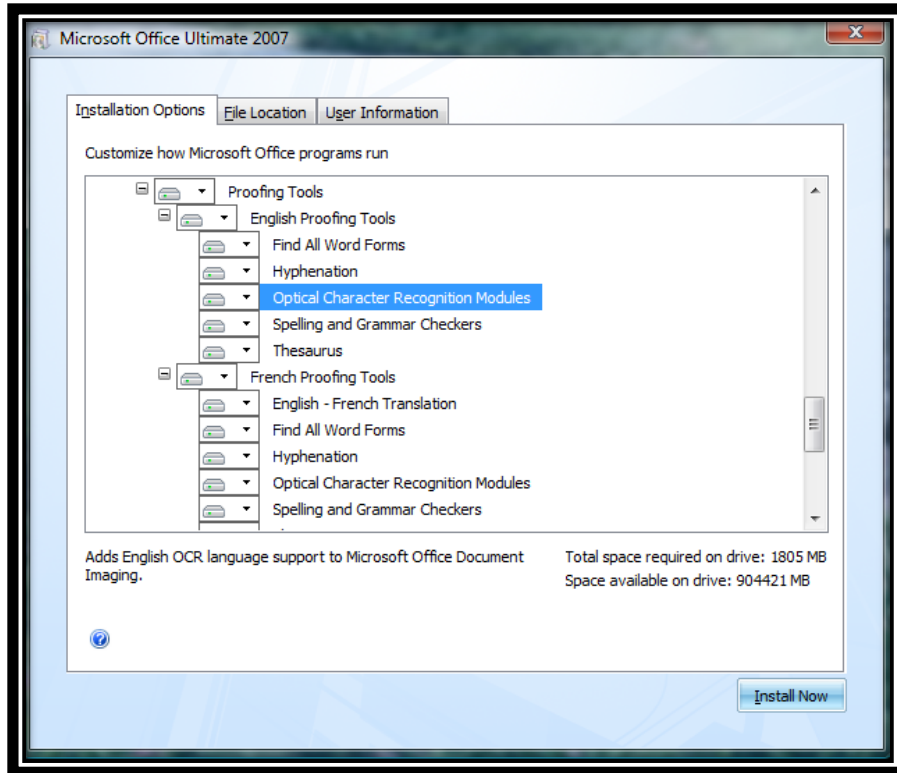


Figure 1 List of components for Microsoft Office install.

If you are using a screen reader, the items will be read to you as you go through them and you will be told that there are sub-items or that the item is collapsed or expanded. Use the Arrow keys to expand or collapse items.

1. Find the "Proofing Tools" components.
2. Expand the list of proofing tools.
3. Locate the "Optical Character Recognition Modules" components use the right mouse button or the AppKey to open the context menu of options for installation.
4. Choose to "Run from my computer." And press Enter to confirm.
5. Move down the list of proofing tools to the French and then the Spanish proofing tools and repeat this the process of running the component from the computer.

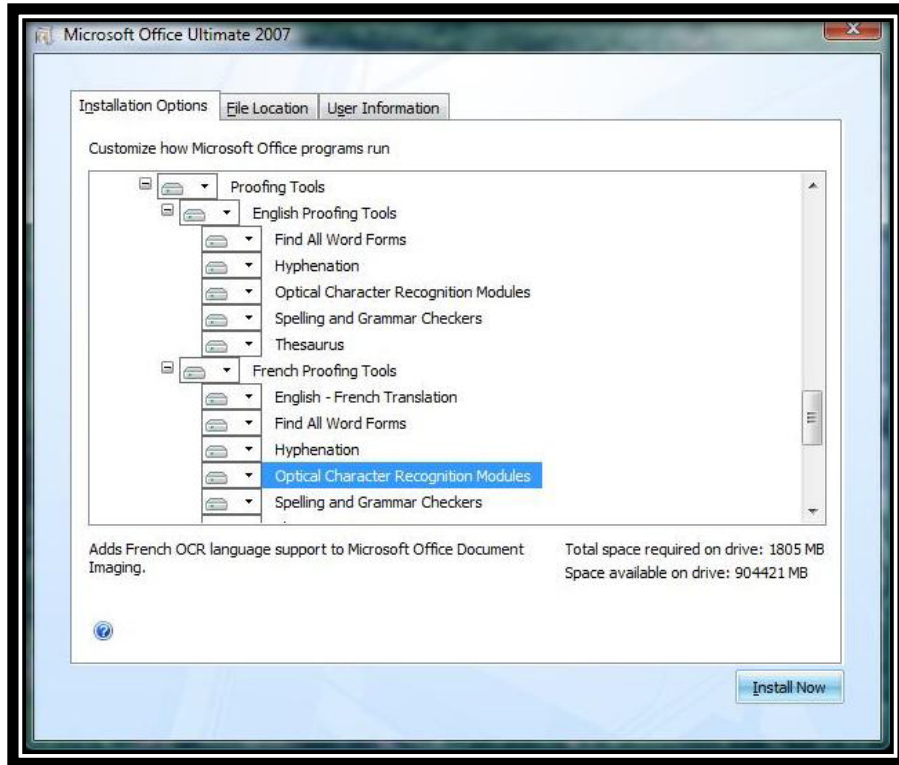


Figure 2 French proofing tools OCR component.

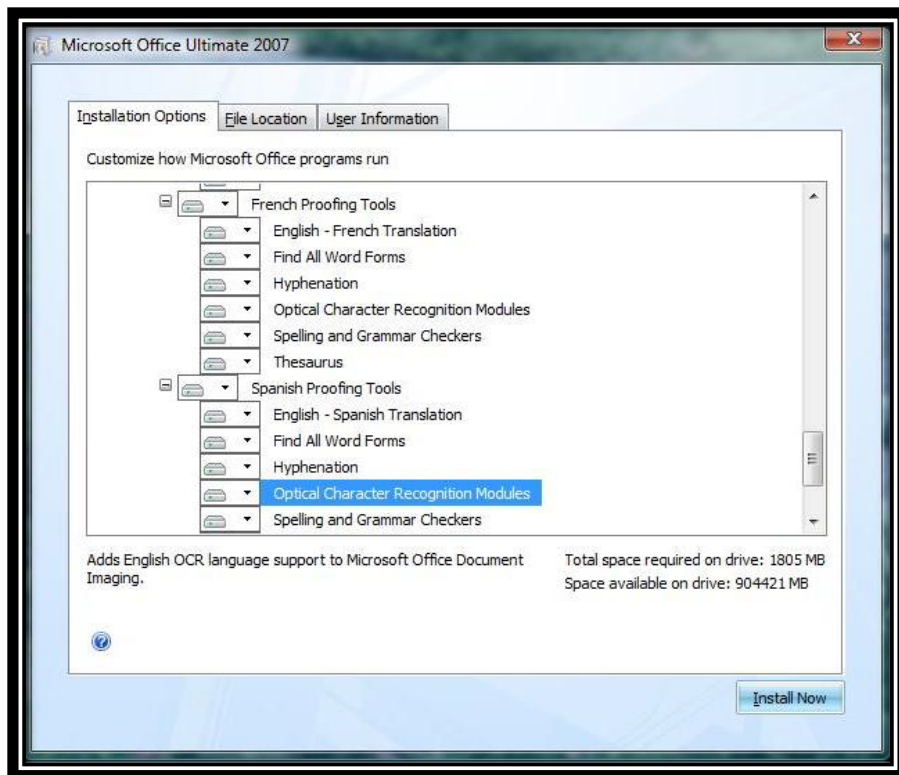


Figure 3 Spanish proofing tool OCR component.

Once you've chosen to install the OCR components you want, move to the first level set of tools called "Office Tools" which is further down the list of components you can install.

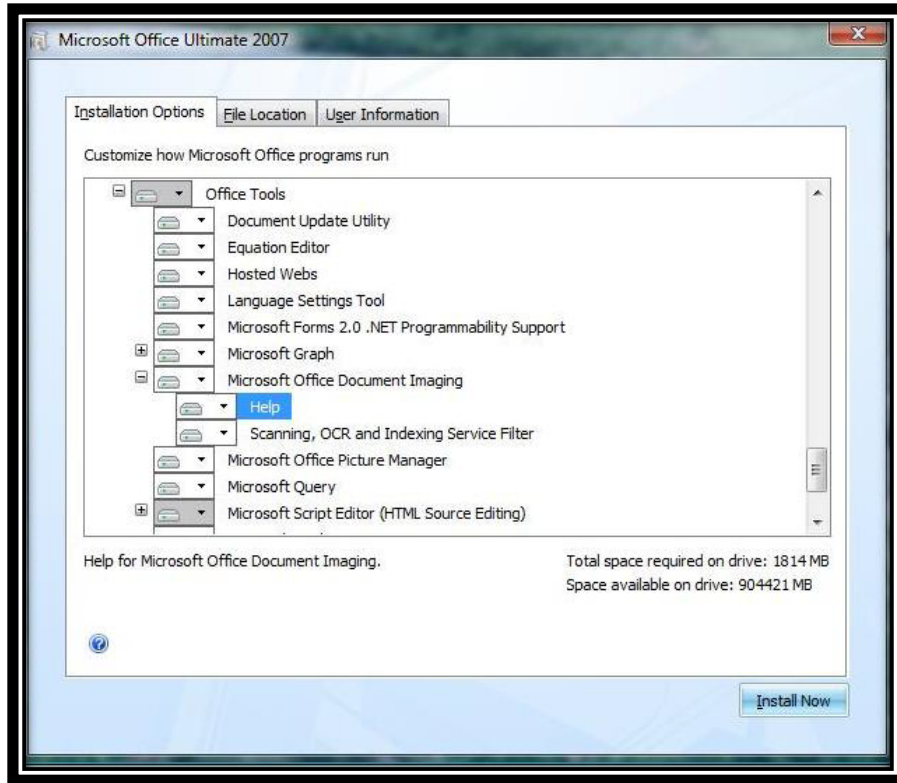


Figure 4 Document Imaging Help component.

You can then move to and activate the "Install Now" button.

If you already have Microsoft Office installed on your computer, you can add the Document Imaging component by going to Programs and Features or Add or Remove Programs and choosing to modify or change the installation of Office.

Upgrading the OCR Software

If you launch the Microsoft Office Document Imaging tool you can go to the Help Menu item where you'll find a topic related to upgrading the OCR.

This lets you upgrade the optical character recognition engine to Nuance's OmniPage Pro OCR scanning software. While the on-board Document Imaging tool lets you read text it will not retain layout and most images. You will need to move to a fuller version of an OCR scanning software such as OmniPage Pro or ABBEY Fine Reader to retain formatting.

However, the Document Imaging tool will work for you in a pinch.

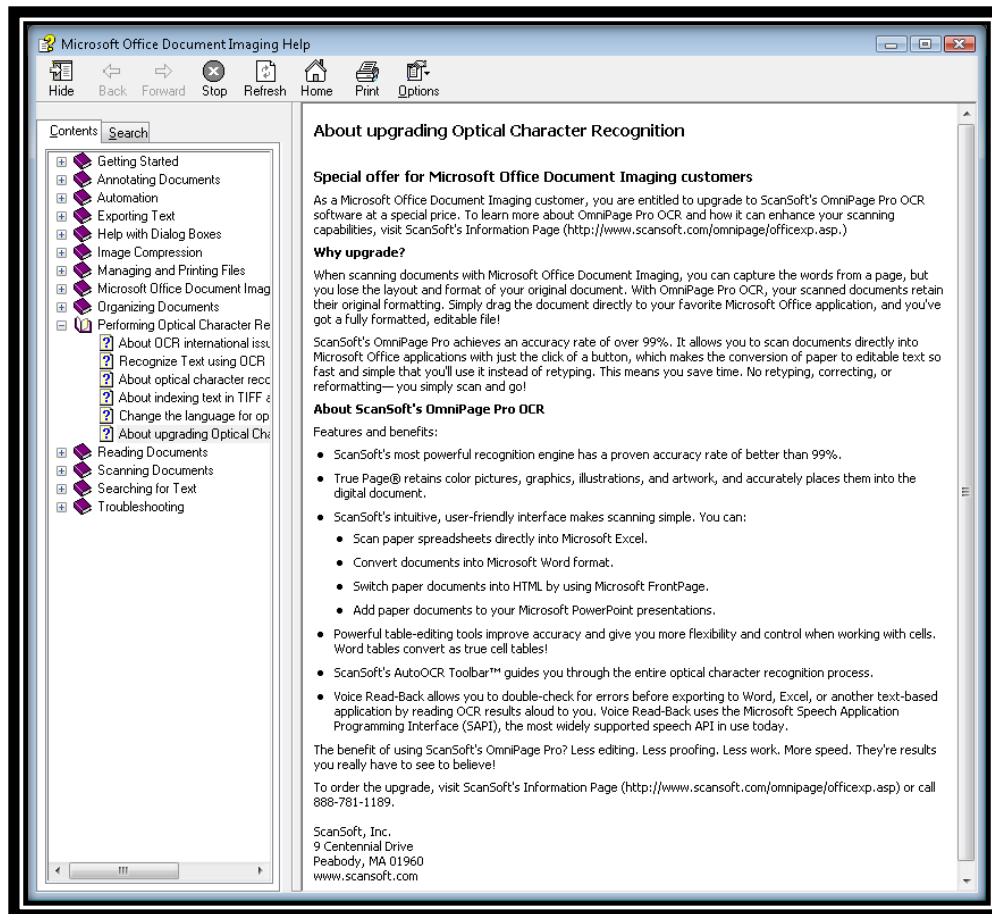


Figure 5 Upgrade information in Help documentation for the Document Imaging tool.

The Document Imaging Tool User Interface

I use the Document Imaging tool on a regular basis and have it on my Desktop. You can find it in Windows Vista or Windows 7 by pressing the Windows Key and typing Document Imaging.

When you launch the Microsoft Office Document Imaging tool, you will see a Menu bar, Toolbar and Status bar. The main document area is divided into two parts: a narrow frame type element on the left and a larger frame like element on the right. These are called the "Thumbnail" and "Page" panes respectively.

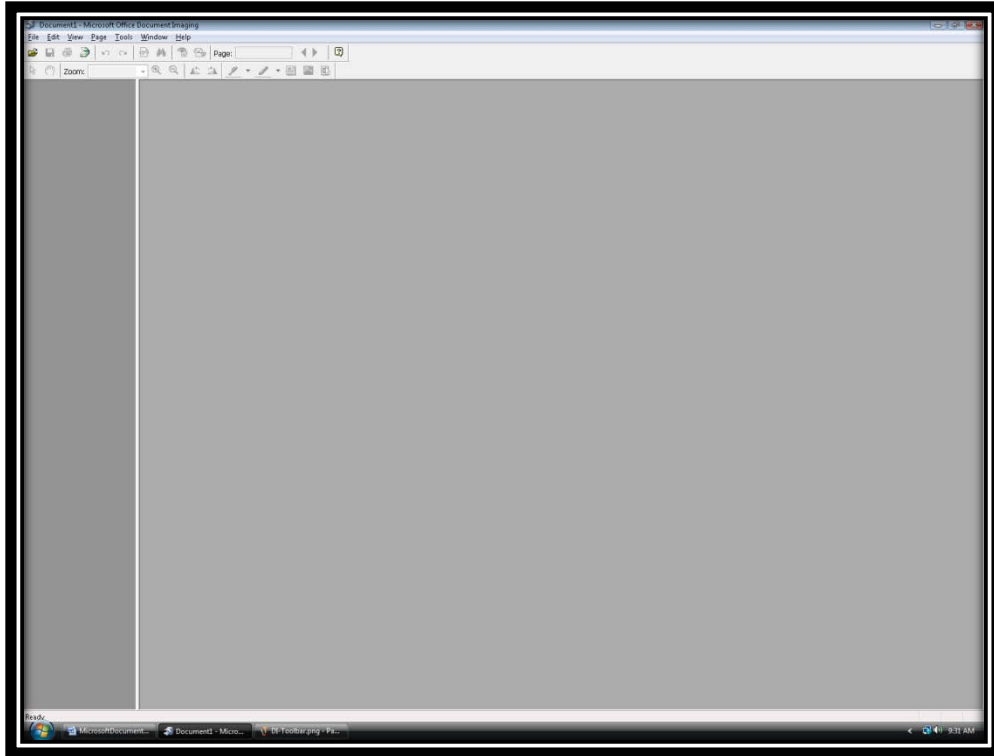


Figure 6 Document Imaging tool user interface.

The following image shows a closer look at the Menu bar and Toolbar.

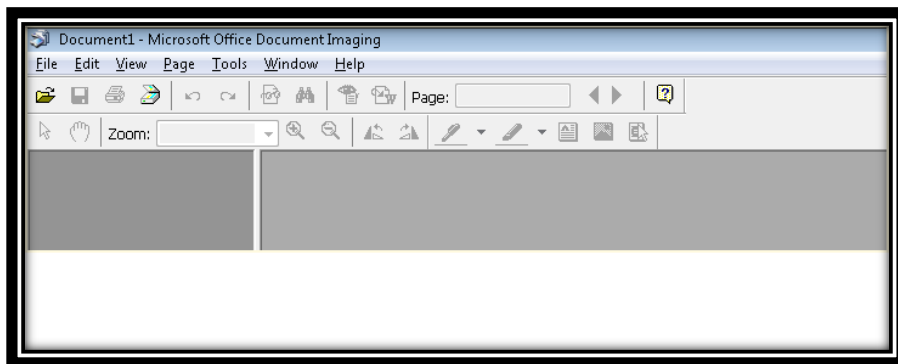


Figure 7 Document Imaging Toolbars.

Accessibility in the Document Imaging Tool

There are two Help documents related to accessibility within the Help documentation for the Microsoft Document Imaging tool. One contains keyboard commands for working with the Document Imaging tool and the other contains tips and techniques for scrolling, zooming, panning and annotating text in the Document Imaging tool.

You can use some standard keyboard commands such as Ctrl + C to Copy, Ctrl + X to Cut, Ctrl + V to Paste or Ctrl + the letter O to launch the Open dialog.

In addition, some of the standard keyboard commands have different functions. Below is a list of keyboard commands you'll find useful when working with the document Imaging tool. For the complete list, consult the Help documentation for the Microsoft Office Document Imaging tool.

To carry out a task in the Document Imaging tool, press:

- Ctrl + A to select the current page.
- Ctrl + Shift + A to select all pages.
- Ctrl + letter I to copy the selected area as a bitmap image.
- Ctrl + C to copy selected annotations if annotations are selected. Use this keyboard command to copy selected text as well.
- Ctrl + X to cut selected annotations if one or more annotations are selected. This will also cut selected text.
- Ctrl + Shift + X to cut the selected page.
- Ctrl + Shift + V to "Paste a copied or cut page. If an image was copied, the image is pasted on a new page." [copied directly from the Microsoft Help documentation].
- Shift + Delete to remove the selected pages.
- Delete to remove the current page.
- Ctrl + T to send all text in the current document to Microsoft Word.
- Tab to move between the Thumbnail and Page panes.
- Ctrl + E to move to the Page box in the Toolbar.
 - Press F6 to move back to the Thumbnail Pane.
- Ctrl + M to move to the Zoom box in the Toolbar.
 - Press F6 to move back to the Thumbnail Pane.
- Spacebar to move to the next page in the Thumbnail Pane.
- Backspace to move to the previous page in the Thumbnail Pane.
- Arrow keys in the Thumbnail Pane to scroll in the direction of the Arrow key you press.

Note: You cannot use standard navigation keyboard commands such as the Arrow keys for moving through words or characters. In the Document Imaging tool these are used to move through pages or sections of content. To navigate the document character by character or word by word, send the document to Microsoft Word.

The page used in this How To guide comes from the book " Microsoft Word 2007 from the Keyboard" by Karen McCall. It was chosen because the page has an image, a caption, a list, a heading, and paragraph text. It represents a typical page.

Opening a TIF or TIFF

Sometimes printed pages are scanned in as TIF or TIFF documents. Using Microsoft Office Document Imaging we can open them and perform OCR on them so that we can read the text using either a screen reader or TTS/Text-to-Speech tool.

To read a TIF or TIFF document:

1. Locate the TIFF OR TIF file.
2. Launch Microsoft Office Document Imaging.
3. Press Ctrl + letter O to show the Open dialog.
4. Locate the file and move to and activate the Open button.

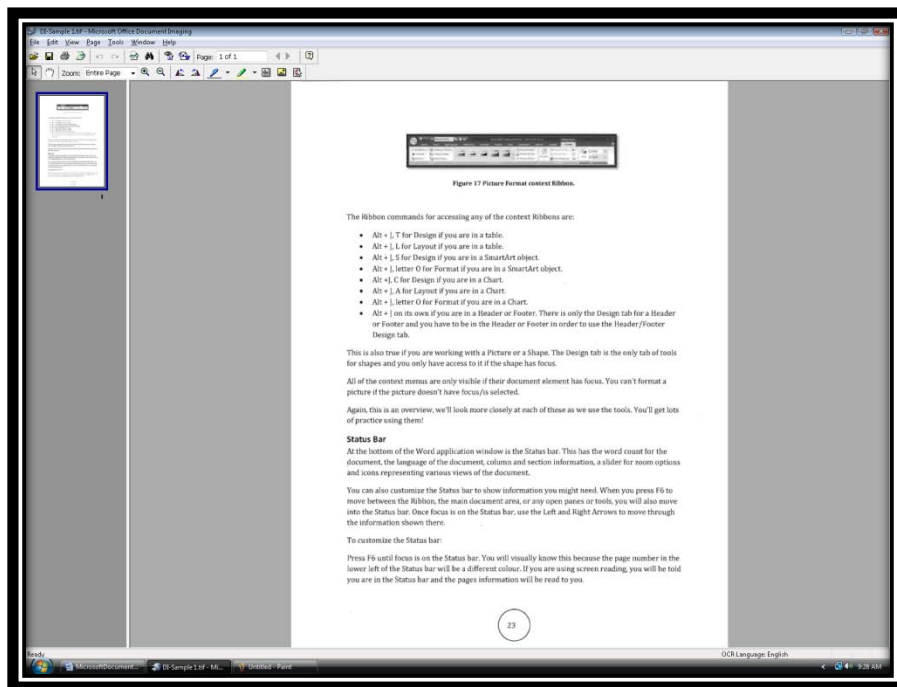


Figure 8 TIF file opened in Document Imaging tool.

5. Press Alt + T to open the Tools menu and choose "

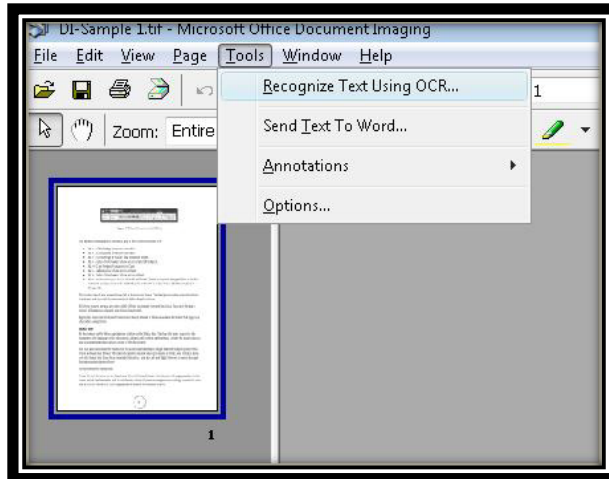


Figure 9 Tools menu showing option to perform OCR.

At this point, if you are using a screen reader you will not have access to auditory feedback in either the Thumbnail or Pages Panes. The best option is to send the text to Word.

Your screen reader will provide feedback on the OCR progress and you will know once the process is completed.

At that point press Ctrl + T to send the recognized text to Word. You can also return to the Tools Menu and choose to send the text to Word.

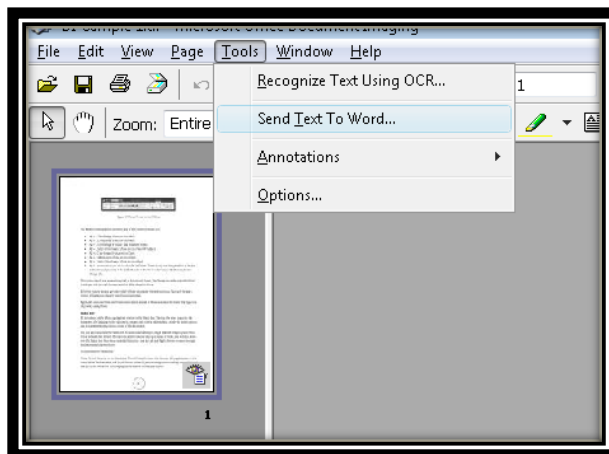


Figure 10 Send the recognized text to Word from the Tools Menu.

When the text opens in Word, it will not be formatted as it was in the scanned page. However, bulleted lists are recognized by screen readers as individual bulleted items which helps preserve list relationships. If an image has a caption, that text is present so someone knows an image is on the page. A heading text is larger and bold, so it retains the formatting but not the underlying heading structure.

Most importantly, we can read the page!

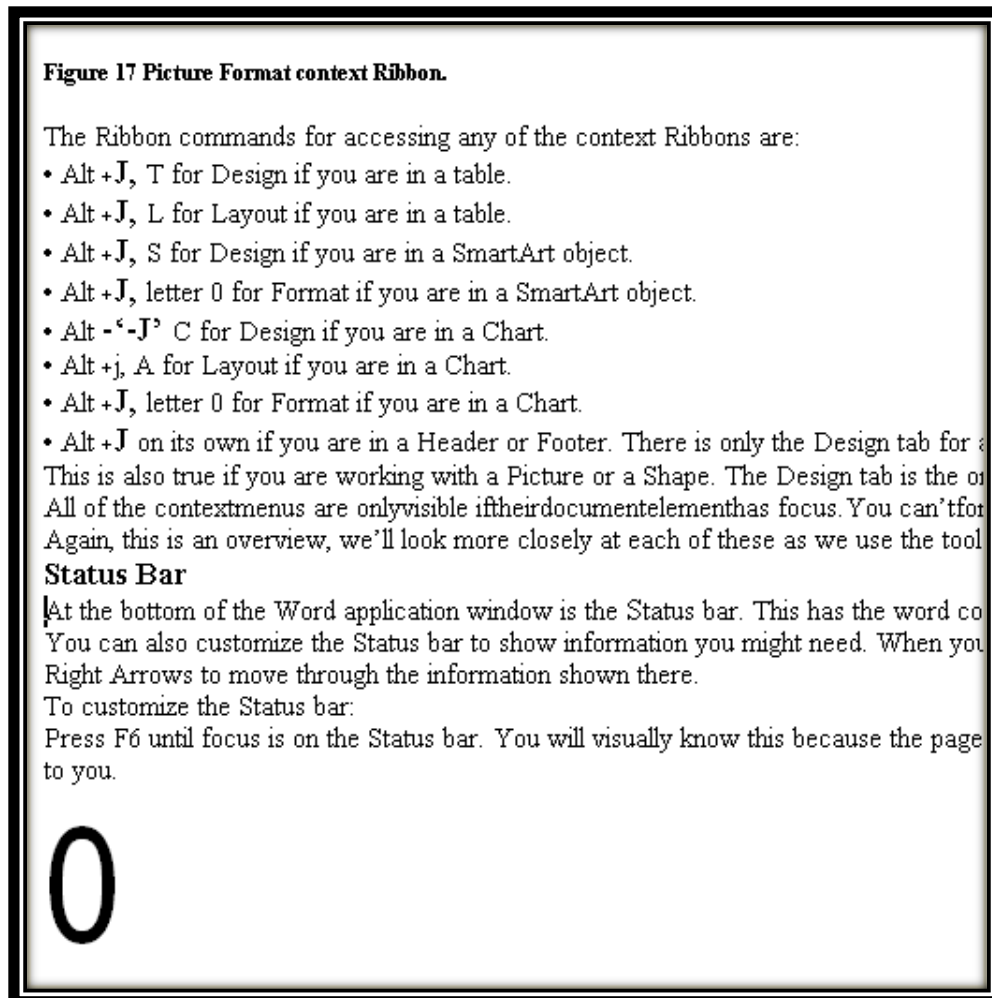


Figure 12 Closer view of recognized text in Word document.

The large zero at the end of the text is a page number. As you can see, the recognition is quite good making this a useful tool for reading images of text in TIF or TIFF format for those of us with print or visual disabilities.

We can then take this content to other devices.

OCR on Scanned Pages

You can scan a page directly into the Microsoft Office Document Imaging tool. For this sample I've chosen a different page from the book "Microsoft Word 2007 from the Keyboard" by Karen McCall. This sample includes footnotes, an image with a caption, paragraph text, and a list.

To read a scanned page:

1. Launch the Microsoft Office Document Imaging tool.
2. Insert your page into the scanner.

- From the File Menu choose to "Scan new document."

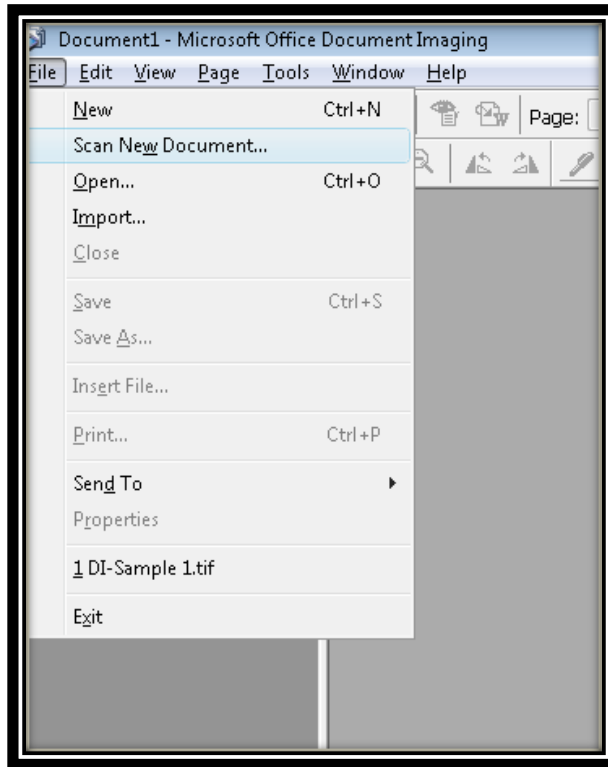


Figure 13 Can new document into Microsoft Office Document Imaging tool.

- You will be prompted to identify the scanner.

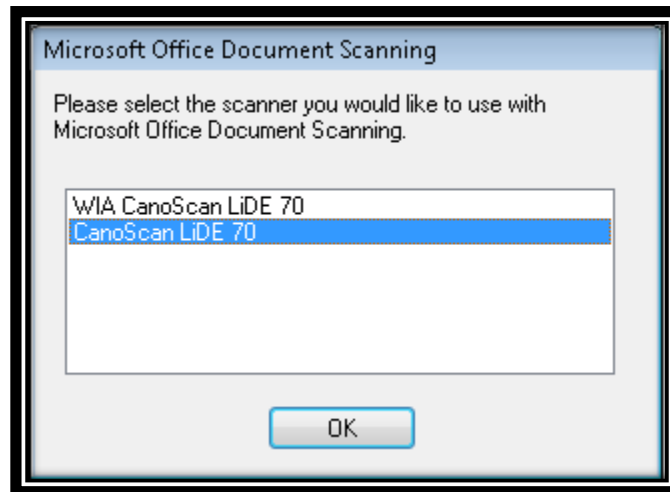


Figure 14 Identify the scanner dialog.

- Once you identify the scanner you are returned to the Scan new document dialog to make further choices. We didn't see this when we activated the File Menu item because we had to choose a scanner before we could make other choices about the page or pages to be scanned.

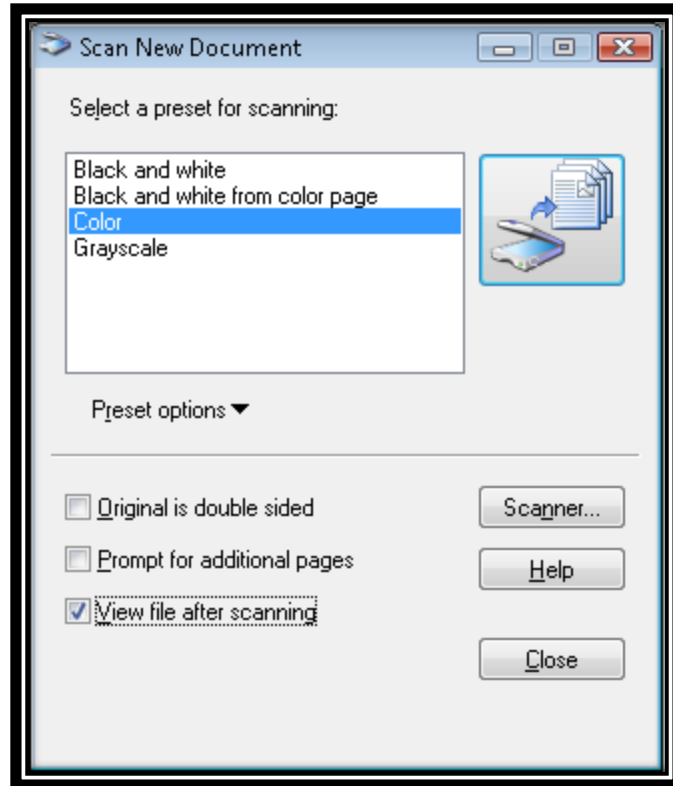


Figure 15 Options for scanning a page or pages into Microsoft Office Document Imaging tool.

6. For this sample I chose to scan in color.
7. A screen reader will identify the scan button which looks like a scanner with a document in it.
8. To scan the page or pages, activate the Scan button.
9. You can add other pages by reactivating this button once the current page has been scanned.
10. When you are finished scanning pages, move to and activate the close button in the Scan new document dialog.
11. Your page or pages will appear in the Thumbnail and Page panes of the Document Imaging tool.

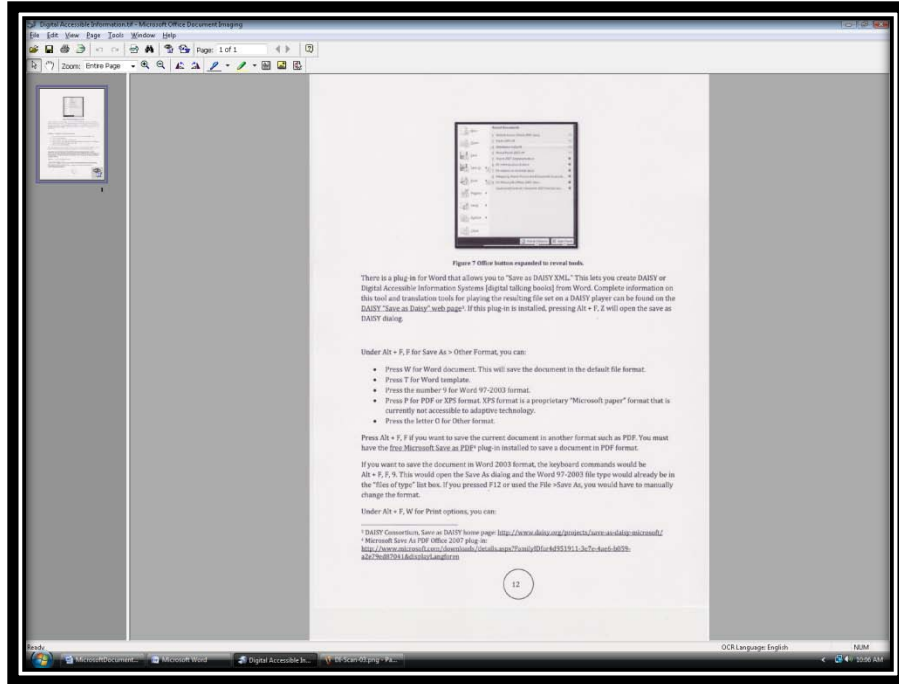


Figure 16 Page scanned directly into the Microsoft Office Document Imaging tool.

Since the document was scanned directly into the Microsoft Office Document Imaging tool, it has already been recognized using OCR.

Press Ctrl + T to send the text to Word.

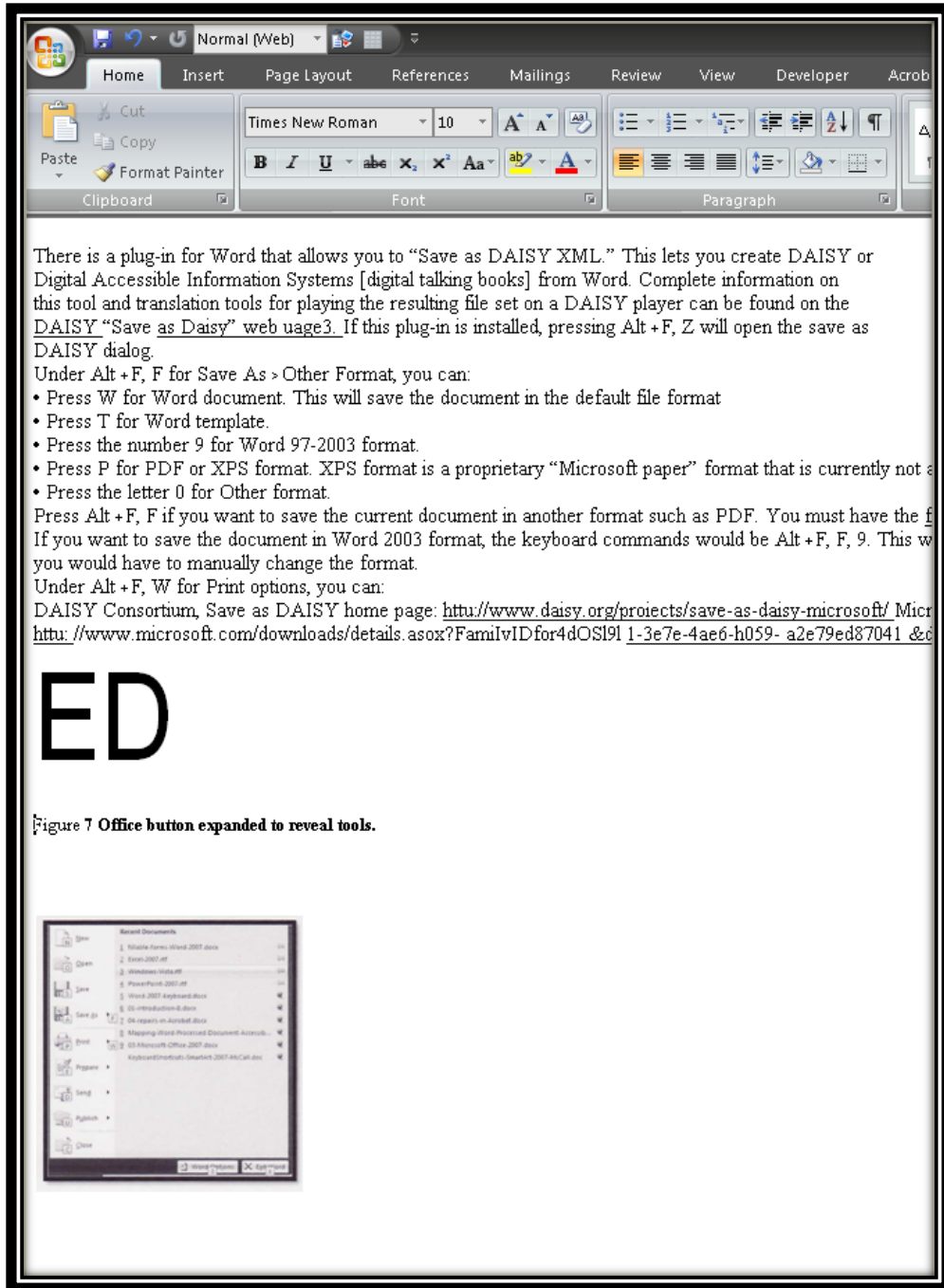


Figure 17 Text from Document Imaging tool sent to Word.

In this case, the footnotes are included but not specifically identified as footnotes. The image is included but is at the bottom of the page rather than the top or where it is located on the print page. The page numbers are represented by the "ED."

Once you get used to the quirks of this tool and the fact that it is a fast and free way to gain access to printed or scanned text, you will find this a valuable tool.

Scanning in French or Spanish

If your printed page is in French or Spanish and you've installed these languages, you will need to change the OCR recognition language before the page is scanned into the Microsoft Office Document Imaging tool.

To change the OCR language:

1. With the Microsoft Office Document Imaging tool open choose Tools, Options.
2. Go to the OCR tab.
3. Tab until you are on the language list.
4. By default the setting is for the Microsoft Office default language.
5. Press Alt + Down Arrow to open the list of languages installed.
6. You can press the first letter of the language you want [for example F for French or S for Spanish].
7. Move to and activate the OK button.
8. Now you can choose File, Scan New Document and the OCR language will be the one you've chosen.

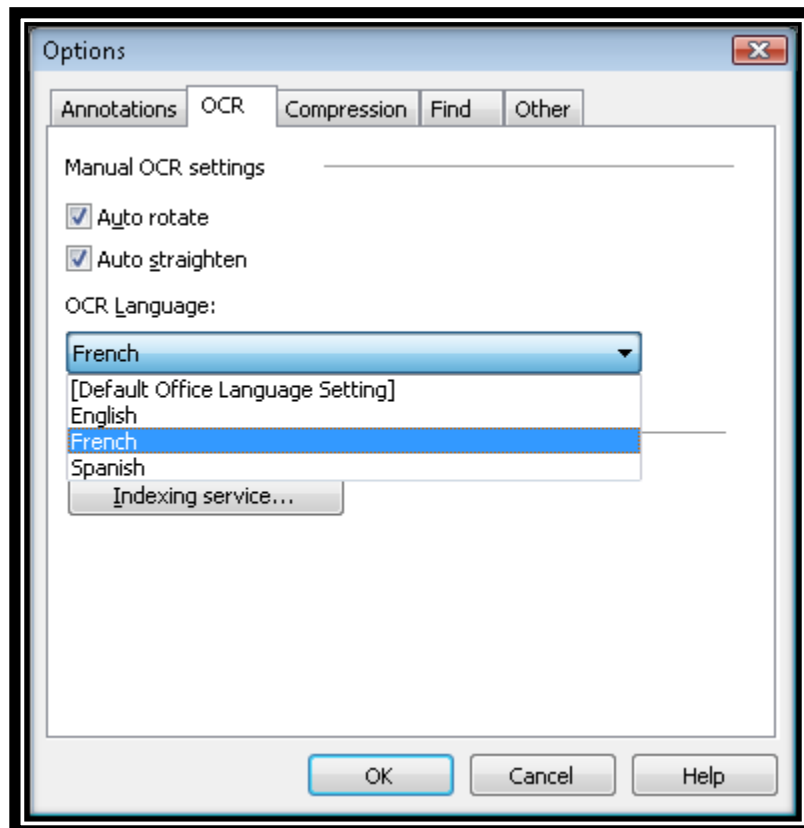


Figure 18 Languages options in the Microsoft Document Imaging tool options.

Remember to change the language back to the default once you've finished.

If the default language for Microsoft Office is Spanish or French, you will still see these languages in the list of available languages.

Microsoft Office Document Imaging Options

You do have options for annotations and OCR in the Microsoft Office Document Imaging tool.

To review the options press Alt + T, letter O. This opens the Options dialog.

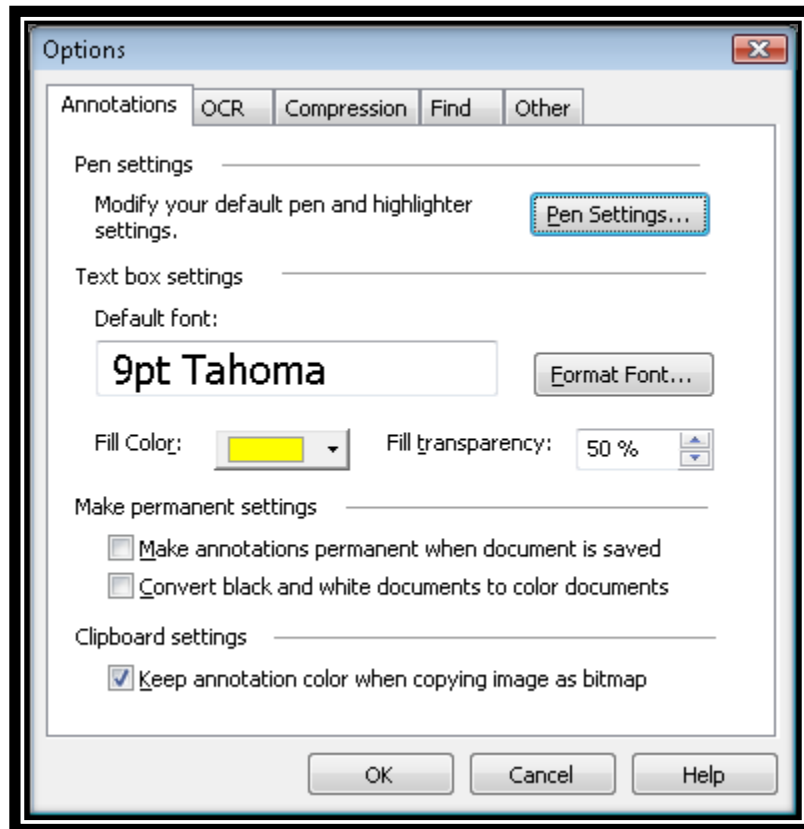


Figure 19 Annotations options for the Document Imaging tool.

This How To document does not cover how to use the annotation tool. Its purpose is to introduce you to the free Microsoft Office Document Imaging tool and how you can use it to gain access to text that is readable if you use a screen reader or TTS/Text-to-Speech tool.

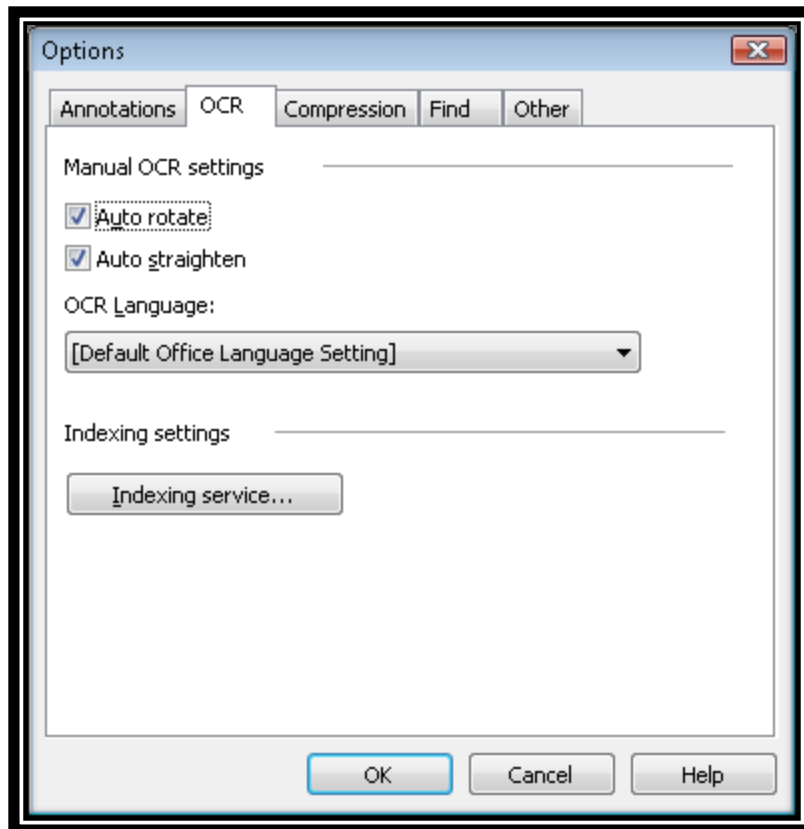


Figure 20 OCR Options for the Document Imaging tool.

Have fun! If you have questions, contact [Karen McCall](mailto:karen@karlencommunications.com).¹

¹ Contact Karen McCall: info@karlencommunications.com