

SECTION 5 PRODUCTION OF E-TEXT (SUPPORTED) THROUGH SCANNING AND OPTICAL CHARACTER RECOGNITION (WHAT IT IS AND SUGGESTIONS FOR DOING IT BETTER)

Introduction

The module Production of E-text (Supported) Through Scanning and Optical Character Recognition will address the use of the OCR program OmniPage Pro12 and a scanner. It will outline the scanning procedure and provide tips and strategies for efficient and successful implementation.

Description of Student Need

Students need to access all of the print resources that are used in the classroom, both textbooks and Teacher prepared materials. They will do so in a variety of formats including e-text which can be made available to the student in hard copy Braille and refreshable Braille.

Assumptions

Materials provided in hard copy will require that the student has an organized way of filing and retrieving paper Braille. Materials provided in e-text will require that the student can independently use (or is actively learning) the technology that receives and stores it. The student will benefit from the active process of reading, (rather than the passive process of listening), therefore the student who utilizes scanned text should, whenever possible, do so via hard copy (translated) Braille or a refreshable Braille display. Due to its graphical interface, OmniPage Pro 12 is not fully accessible to a blind user. Products by Arkenstone and Kurzweil are better suited to providing independent access to print, for blind users.

Prerequisites

The person providing the material (the Brailist) will have a good working knowledge of the OCR program, a word processor, and the operating system of the computer. The person receiving the material will have a good working knowledge of their device (computer or notetaker) for file storage, organization and back up).

Classroom Implementation Strategies

- The student will require storage space in each classroom if hard copy Braille is being provided.
- For e-text letter baskets can be placed in classrooms so print materials can be left by the teacher and e-text (on disk) can be left for the student.

Common Barriers or challenges that might occur

- Not receiving materials with sufficient time to produce it in e-text. This can be minimized if Brailleists are in daily communication with the classroom teachers.
- The complexity of the technology (both the computer operating system and the software that is used). This can be minimized by networking with others who use the same technology. It can also be minimized through relevant in-servicing and the ability of the person using the technology to become an independent learner (by using the Help files).
- The student may require ongoing instruction on the use of their technology and their ability to organize the materials that they receive.

Equipment Needed

- Current Pentium computer (512 meg RAM, Windows XP)
- Industry standard optical character recognition (OCR) software (e.g. OmniPage Pro 12)
- Supported scanner (e.g. Epson). Check scanner speed before purchasing
- Industry standard word processor (e.g. Microsoft Word 2000 or 2002)
- (Optional) Braille Embosser
- (Optional) Automatic Document Feeder is a useful add-on for high volume e-text production

FAQ

1. Why is it necessary for students to have all of this in e-text, when someone could just read it to them?

Students who are blind have a right to the same learning opportunities as their sighted peers. This includes access to written materials in a format that provides spelling, punctuation and the ability to retrieve it at their convenience.

2. Is it necessary to make the e-text look exactly like the print page?

No. Things like running headers and footers, with the exception of page numbers, are unnecessary as they add to the student's reading load. Students are reading for information and repeated text simply slows them down. Students generally prefer headings to be left justified as it saves them from searching across their Braille display (or across a hard copy page) for text.

3. Is it OK for the student to use my office as a home base?

No. Students who are blind spend a disproportionate amount of time with adults. Our job is to support them in the background and foster their independence.

Information and Instructions OmniPage Pro 12

Part A. Preliminary Information

- OmniPage Pro 12 was designed primarily as a business application. This module only elaborates on features of the program related to the production of e-text for students who are blind or low vision.
- Certain features of the program are beyond the scope of this document. These features include:
 - Creating and using Adobe PDF
 - Creating zone templates
 - Proofing OCR results
 - Verifying text
 - User dictionaries
 - Training OCR
 - Reading text aloud
- This document may be best utilized by having it running in the background. You can then access it at any time from the Windows Taskbar located at the bottom of the screen (make sure Auto Hide in the Taskbar properties menu is unchecked).

Why Scan?

- Students have access to any and all reading material, thereby encouraging them to read more.
- Material can be provided for blind and low vision students quickly and on-sight, thus providing them with written information at the same time as their classmates.
- The resulting e-text is not consumable, takes up no external space, can be easily and efficiently organized, and is permanently archived for future students' reading needs.
- Students equipped with a computer or Braille notetaker and a refreshable Braille display can carry disks rather than cumbersome multi-volume Braille books.
- By using the "Find" command in their word processor students can navigate e-text easier and faster than hard copy Braille.

Some Considerations

- Transferring files to your student is much easier if you and the student have the same operating system. In most cases this will mean that a Windows operating system is preferable.
- Newer versions of OCR and word processing programs are large. The faster the computer and the more RAM it contains will largely determine the speed at which the process can take place.

- Scanners vary greatly in the speed that they pass over the document. This is determined by the connection of the scanner to the computer (USB or parallel); the dots per inch (dpi) settings within the scanner software and whether the optical character recognition (OCR) is set to scan colour (slower) or black & white or grayscale (faster).
- Setting the scanner's own software (see scanner manual) to 300 x 300 dots per inch (dpi) will increase scanning speed and optimize the results for text. (Note: This is not the optimum dpi setting for colour photographs or pictures).

About the Source Document (Paper)

- OmniPage Pro 12 attempts to reproduce the letters and words on the page to match the original. Many factors can cause OmniPage Pro 12 to make mistakes which result in the need to manually correct the spellings of words. See "About the Clean Up".
- The better the quality of the source document, the less time will be spent doing clean up.
- Examine the text on the page to determine if there are parts that do not need to be scanned (e.g. running headers, footers). One quick way to eliminate these from the resulting document is to cover them with post-it notes before scanning the page.
- For close columns you can draw a pencil line between them. This eliminates the scanner joining two columns into one zone and creating lines of text that run across both columns. This applies mainly to older OCR programs.
- Books opened with two facing pages can be scanned at once (e.g. novels) and OmniPage Pro 12 will correctly order the two pages. Make sure that "Look for facing pages" is checked (Tools / Options / Process).
- It is unlawful under current copyright laws to reproduce print copies of books for students unless the original print copy has been purchased or permission has been received from the publisher.

Part B. The OmniPage Pro 12 Program

General Information

- Consistent with all Windows programs, any menu item ending with an ellipsis (...) will result in a dialogue box being opened when it is chosen.
- When you hover your mouse cursor over any box, button or icon in the program, a description of its function will be displayed in the Status bar at the bottom of the OmniPage Pro 12 desktop screen.
- Right clicking on any control in a dialogue box will present a tool tip screen that explains the options for that control.
- Help may be obtained three ways from the Help menu:
 1. An online manual broken into chapters
 2. A searchable index that displays all topics in the online manual
 3. A searchable index that displays only topics related to the search item
- Help may be obtained at any time by pressing F1. This help is context sensitive and will relate to the specific aspect of the program that you are currently working with.

The OmniPage Pro 12 Desktop

- The OmniPage Pro 12 Desktop has a Title Bar and a Menu Bar along the top.
- Below that is a Standard toolbar for common tasks (print, save etc).
- Below the Standard toolbar is the OmniPage Pro 12 Toolbox consisting of 2 rows of buttons. The first row contains: the Start (1 2 3) button; the Get Page (1) button; the Perform OCR (2) button; and the Export Results (3) button. Below each button is a drop-down list where settings for each button are made.
- The OmniPage Pro 12 Desktop has three main working areas, separated by splitters for resizing the pane. They are the Image Panel, the Text Editor, and the Document Manager.
- Along the bottom of the OmniPage Pro 12 Desktop there is a Status Bar where the appearance of the Desktop can be changed by hiding or showing the Image Panel, the Text Editor and the Document Manager.

The Image Panel

- The Image Panel, on the left, is divided into two sections, separated by a (vertical) Image toolbar.
- The Image Panel displays a thumbnail image of each page on the left side of the Image toolbar and the zones for the selected (thumbnail) page on the right side.
- After a second page has been scanned, you must click on each thumbnail to view the recognized text for it. The current page is indicated with an “eye icon”. You can re-order pages by dragging a thumbnail and placing it between other thumbnails.
- Various icons will appear in the thumbnail to indicate its status (see manual p.26 for details on icons). The most important ones are: the “glasses” icon indicating that the page has been recognized, and the “disk” icon indicating that the page has been saved.
- The Image toolbar in the middle is used to manually re-order, delete or re-draw the zones for the selected thumbnail page (see The Image Toolbar).

The Text Editor

- The Text Editor, on the right, displays the recognized text for the thumbnail that is selected.
- The Text Editor has a formatting toolbar at the top to change the attributes of the recognized text. To do this you must select the text first. Also on the formatting toolbar are buttons to:
 1. Add/remove bullets from selected text in the Text Editor
 2. Show/hide non-printing characters in the Text Editor
 3. Show/hide the Verifier (see owners manual for information)
 4. Find the next suspect (non-dictionary word)
 5. Show/hide the reading order of paragraphs, tables, graphics (see Changing the Reading Order)
 6. Change the reading order of paragraphs, tables, graphics (see Changing the Reading Order)
- If the formatting toolbar is not visible select it (View menu / Toolbars / Formatting). If only part of the formatting toolbar is visible (i.e. Reading Order buttons not showing) re-size the Text Editor pane by putting your cursor on the splitter line separating the

Text Editor and Image panel, click and drag to the left to make the Text Editor pane larger.

- The Text Editor has view buttons at the bottom to change the way recognized text appears for editing. The choices for viewing recognized text are:
 1. True Page (exact replication of the original page)
 2. Retain Font and Paragraphs (text retains attributes but is decolumnized)
 3. No Formatting (text is decolumnized and stripped of all attributes and format)
- The view buttons are only for viewing while editing and are independent of the way text will be exported to a file (e.g. viewing text in “No Formatting” does not mean it will be saved without formatting – see The Save Dialogue Box to change the way text is saved).

The Document Manager

- The Document Manager, located below the Image Panel and Text Editor provides details of your document in a table format.
- Each row represents one page and the columns give statistical or status information for the page.
- If it is not necessary to see the Document Manager it can be removed from the desktop view by clicking the corresponding button on the Status bar at the bottom of the OmniPage Pro 12 desktop.

The Process (The Big Picture)

- OmniPage Pro 12 turns paper documents into electronic documents containing editable text (e-text) through the use of a scanner.
- OmniPage Pro 12 turns non-editable text, such as Adobe PDF files (.pdf) into electronic documents containing editable text (e-text),
- It is done in 3 steps: capturing the page as a graphic; ordering the information on the page into zones (both pictures and text); and converting the words from a graphic into editable text (recognizing it).
- Once editable text has been created, it can be saved (in many common file formats) and given to the student (on disk or via email) or it can be saved and opened in a Braille translation program for the creation of hard copy Braille.
- The entire process outlined above can be done from within a word processor without opening OmniPage Pro 12, as OmniPage Pro 12 can run in the background. See Scanning directly into Microsoft Word.
- Pictures may be removed when the document is saved to a word processor (for Braille using students). To remove graphics when saving in OmniPage Pro 12 go to “Converter Options” in the Save As dialogue box and uncheck “retain graphics”.
- It is possible to defer zoning and recognition and have it done at a later time. This frees the Brailist from waiting for these processes to be completed before scanning the next page. To do this press the Get Page button for each new page to scan it. When all pages have been scanned press the Export Results button. OmniPage Pro 12 will ask you if you would like to have all pages recognized. Answer “yes” and it will finish the zoning and recognition processes for all scanned pages. (This feature is unnecessary with fast computers as zoning and recognition happens before a new page can be placed on the scanner).

The Process (The Details)

Using the OCR Wizard

- This is the simplest way to use OmniPage Pro 12 as it guides you through the various settings and commands in 5 steps, by asking you questions. When you are done the 5 steps OmniPage Pro 12 launches “Automatic Processing”.
- The OCR Wizard is a good way for new users to learn about the various settings and commands, however, it is not usually the method of choice once the user becomes familiar with the program.

Automatic Processing

- This is most appropriate for large jobs where pages are relatively simple (i.e. do not contain lots of sidebars and/or graphics with captions).
- First, choose the appropriate setting in the Select Image Source listbox located below the Get Page button. See the section: Choosing from the Image Source listbox
- Second, choose the appropriate setting from the Select Layout listbox located below the Perform OCR button. See the section: “Choosing from the Layout Description” listbox.
- Third, choose the appropriate setting in the Select Destination listbox located below the Export Results button. See the section: “Choosing from the Select Destination” listbox.
- Click the Start (1 2 3) button to have OmniPage Pro 12 automatically perform all processes according to the individual settings that you have made.
- Scans can be scheduled to occur automatically. This allows you to turn the page and have OmniPage Pro 12 scan the next one after a set time (Tools menu/Options /Scanner tab). Using this feature eliminates the need to press the Start (1 2 3) button to begin the scanning process for each page.

Choosing from the Select Image Source listbox

- Choose Load Image File if you do not want to use a scanner, but rather, have an image file (i.e. non-editable text such as a PDF file) located on your computer that you want to process (turn into editable text).
- Choose Scan B&W for crisp black text on white paper. It works best with original source documents (not photocopies). If it yields acceptable results, it is the fastest way to process the document.
- Choose Scan Grayscale for most situations. It typically provides the best results for both original black and white documents and those that have been photocopied.
- Choose Scan Color if you have a colour document. This will significantly slow down the scanning process and create much larger files. It is not recommended to choose Scan Colour unless you are supporting low vision students for whom colour is important.

Choosing from the Layout Description listbox

- This listbox tells OmniPage Pro 12 what to look for when it is creating zones. Zones are the separate areas on the page (text, graphics, tables). By dividing a page into zones, OmniPage Pro 12 knows how to reconstruct it during the recognition process.
- The latest version of OmniPage Pro 12 is much more sensitive to tables. Many options in this listbox now relate to the inclusion of tables on the page. This setting is useful in that once exported to Word, Duxbury will recognize the table and translate it appropriately into hard copy Braille.
- It is important to check this setting before starting a new scanning job. If it was set to single column and the new document has multiple columns, the resulting single zone text will join lines of text from both columns into one string.
- Choose “Automatic” if the document does not have a consistent layout (some pages are single column and some have multiple columns) or if it contains both columns and tables.
- Choose “Single Column, No Table” for simple layouts that do not contain columns or tables.
- Choose Multiple Columns, No Table for most scanning tasks. Multiple columns can be decolumnized which is most appropriate for reading with a screen review program or for conversion into hard copy Braille. Single column documents will not be affected by this setting (unlike the Single Column” setting which does affect multiple column documents).
- Choose “Spreadsheet” for export to a spreadsheet program.
- Choose “Custom” to pre-determine whether columns, graphics or tables will be detected. Make these settings in the Options dialogue box / Custom Layout section. This dialogue box is useful if you want to certain aspects (columns, graphics, or tables) detected, or not detected during the entire scanning job.
- Choose “Template” to create a consistent zone pattern for every page in the scanning job. This can be useful if you want to eliminate running headers and/ or footers that have a consistent location on every page.
- Do not confuse the zones that are created in this step with the output formatting level that will be chosen when you save. The choices made in this listbox simply tell OmniPage Pro 12 what to look for on the page during scanning.

Choosing from the Select Destination listbox

- Documents can be saved either as “images” (i.e. a picture of the word, therefore not editable) or as “recognized” (editable text). For most situations relating to student support scanned documents will be recognized before they are saved.
- Choose “Save As File” to have the document saved after it is recognized. If you have multiple pages to scan and do not want the Save As dialogue box to open immediately after each page is recognized, check “Automatically prompt for more pages” (Tools / Options / Process dialogue box). OmniPage Pro 12 will display an “Add more pages” dialogue box for about 10 seconds before opening the Save As dialogue box. If you can put the next page on the scanner within this time period and press the “Add more pages” button, the next scan will start. Otherwise the Save As dialogue box will open. If you do not want to save yet, close the dialogue box and the “Add more pages” dialogue box will appear. Continue scanning until all pages are scanned. When you are finished scanning and the Save As dialogue box opens on the last page, name the file and open the Page Range listbox. Choose “All pages”

from this listbox. The File Options listbox now becomes active. Choose “Create one page for all files”. Now all scanned pages will be saved in a single file (see “The Save As” dialogue box).

- Choose “Send as Mail” to have your default email program opened automatically and the file included as an attachment. A dialogue box will open asking the format of the attachment.
- Choose “Save to Clipboard” to have pages saved to the clipboard, rather than have the Save As dialogue box open after each page is scanned. Once all pages are scanned change the listbox to “Save as File” and click the button above it. The Save as dialogue box will now open.

Manual Processing

- Manual processing involves pressing the Get Page, Perform OCR and Export Results buttons on the OmniPage Pro 12 toolbox individually. During manual processing only the action related to the button will be performed until the next button is pushed.
- Manual processing is primarily used when you know that you will want to manually draw a zone (or zones), for recognition, from within the scanned page and leave the rest of the page unrecognized.
- Using the Image toolbar you can re-draw zones, select zones for processing, ignore zones and have zoning changes immediately made on the recognized page (see The Image toolbar).
- It is generally unnecessary to use manual processing as zones can be removed, re-drawn or rearranged during the automatic process.
- Any changes made to the zones on the page will be automatically made on the recognized page when “zoning on the fly” is enabled on the Image toolbar (see The Image toolbar). Otherwise the changes will be made when the Perform OCR button is pushed.

About Zones

- In order to reconstruct the page after it has been scanned, the program must organize it into its component parts and sequence them as per the original document. These parts may be headings, paragraphs, pictures, captions, sidebars or tables. Each part is enclosed in a border and becomes a separate zone. Only information in “process” zones will be recognized and transferred when saved and exported.
- Zones can be set to “process” or “ignore”. Unless they are manually drawn using the “ignore zone” drawing tools on the Image toolbar, all foreground zones will be processed (recognized)
- The background is also a zone which can be set to “process” or “ignore”. It is set to “ignore” by default
- Zones are automatically determined by the program (when Automatic processing is activated by using the Start 1 2 3 button), but zones can be altered using the drawing tools on the “Image Toolbar” (located in the Image View).
- Zones can be re-ordered so that the reading order will be changed in the final document. See “changing the Reading Order”
- When you are manually processing you must manually draw the zones to be recognized

- There are 7 zone types. They are: Text zone; table zone; graphic zone; process zone; process background; ignore zone; ignore background. The zone types are determined automatically when automatic processing is used. When manually processing is used they are determined by the drawing tools on the Image toolbar that you used to create them
- Zones may be ignored and therefore not recognized during processing. To ignore a zone select it by right clicking on the zone and choosing “clear” from the context menu that appears. This is an easy way to eliminate running headers and footers that may not be appropriate for e-text for Braille using students. Once a zone has been cleared you must press the Perform OCR button to have the changes displayed in the Text Editor if the “zoning on the fly” button is not pressed (see The Image Toolbar for information on this button).

The Image Toolbar

- This toolbar contains drawing tools which allow you to have complete control over all aspects of the zones on the page.
- If zones need to be moved or re-sized it is usually better to do it when using the manual process. Moving or re-sizing zones that have been made during Automatic processing can create confusing results in the Text Editor. The Undo button on the standard toolbar can be used to return zones to their original state.
- Hover your mouse over each drawing tool to get a description of it on the status bar.
- The 5 drawing tools on the Image Toolbar are:
 1. Process zone: Use this tool to draw an area where you want auto-zoning to occur. The program will automatically determine what kind of zone it is (and break it into separate zones if necessary)
 2. Ignore zone (gray colour): Use this to draw a zone that you do not want transferred to the text editor, and do not want to be recognized for export to a word processor
 3. Text zone (brown border): Use this to draw a text zone. It will decolumnize text in the Text Editor and in the exported document
 4. Table zone: Use this to have the zone contents treated as a table. Table grids will be automatically detected or can be placed manually in the zone. The manual creation of tables is beyond the scope of this document – see documentation that came with the program for information on this
 5. Graphic zone: Use this to draw an area that will be transferred to the Text Editor as an embedded graphic. That means it will not be recognized as text
- There is a button on the Image toolbar to have the background set as a process zone (so it will be recognized in the Text Editor and exported).
- There is a button on the Image toolbar to ignore the background. This is selected as the default.
- There is a button on the Image toolbar to select a zone. Once selected the zone can be moved or the size of the zone can be changed. If you wish to change the reading order of the text, it is best done using the Reading Order buttons on the Text Editor toolbar (see Changing the Reading Order).
- There is a button on the Image toolbar to rotate the zone. If you rotate a zone its properties and border will disappear i.e. it will no longer be a defined zone.

- There are two zoom buttons on the Image toolbar. One to zoom in on the text shown in the Image pane and one to zoom out. Note: once you zoom in you can't zoom all the way back to the original size.
- There is a button to have zone changes processed "on the fly". When it is pressed zone changes are made immediately in the Text Editor. During manual zoning pressing this button will cause the page to be recognized without pressing the Perform OCR button.
- There is a button on the Image toolbar to have zoning changes stored. Press the "Zone on the fly" button when you want to have them processed.

Changing the Reading Order

- Unlike previous versions of OmniPage Pro 12, the reading order is now changed from within the Text Editor (rather than by rearranging zones in the Image Panel)
- To change the reading order:
 1. Choose True Page from the buttons at the bottom of the Text Editor.
 2. Make sure the Reading Order and Change Reading Order buttons are visible on the Text Editor toolbar (if not resize the pane as described in the Text Editor).
 3. Press the Reading Order button. Notice how gray arrows appear in the Text Editor showing the flow of text, tables and graphics (i.e. the reading order). This step is optional. Its only purpose is to map the flow of text. It stays in effect for all pages (until pressed again to toggle it off). It is not necessary to press it before pressing the Change Reading Order button.
 4. Press the Change Reading Order button. Notice how numbered (blue) boxes outline the different sections on the page. Click the "Define New Reading Order" button (blue down-facing arrow). Notice how all of the numbers disappear. Click on the boxes in the order that you want them to be read. The boxes will be numbered accordingly. To cancel the Reading Order changes click on the red X button. To restore the previous reading order, click the blue undo button.
 5. When the correct reading order is established click the green checkmark button. The new reading order will not show up in the True Page view (even though it must be set to this view to make the changes). To actually rearrange the text in the Text Editor you must switch to either the "No Formatting" or the "Retain Font and Paragraph" view.
 6. The new reading order will be exported to a word processor only if you choose "No Formatting" or "Retain Font and Paragraph" when you save. Saving with the Formatting Level set to "True Page" will not cause the reading order to be changed.

The Dialogue Boxes

Note: Only dialogue boxes that relate to the production of e-text or hard copy Braille will be discussed.

Note: A tool tip description of any dialogue box item can be displayed by right clicking on it.

The Save As dialogue box

- Only the features unique to OmniPage Pro 12 will be explained.
- Open the “Files of Type” listbox to choose a file type. There are many file types that the document can be saved as. Some refer to specific destination programs (e.g. Word, Word Perfect), and others refer to more generic types of files (text, Unicode). For e-text choose the file type according to the destination program that the student will be using. For hard copy Braille choose Word as Duxbury will allow direct importing from Word (and will also preserve the formatting and style types). For use with a Braille notetaker, determine what file types are supported by the notetaker (Pacmates and Brailnotes will support Word, Braille Lites will only support text files).
- Open the “Formatting Level” listbox to select how you want the final document to be exported (i.e. how it will look in the destination program that will open it after scanning is finished). The choices are:
 - No Formatting, which means exported text will be decolumnized and all unique attributes (bold, italics etc) will be removed.
 - Retain Font and Paragraph, which means text will be decolumnized but different and paragraph styles will be preserved in the exported document. This is typically the choice used for both e-text and hard copy Braille.
 - Flowing Page which means all aspects of the text on the original page (i.e. columns) will be exported to the target program.
 - True Page which means all aspects of the original page (including text, columns and tables) will be exported to the target program.
 - For most situations involving visually impaired students (blind and low vision) either No Formatting or Retain Fonts and Formatting will be used. Note that if pages have been reordered, the new reading order will not be saved if True Page is used.
- Open “Converter Options” dialogue box (press Advanced button if it is not showing) to select or deselect various aspects of the document depending on how you want it to look in the destination program. These settings can be saved. This is useful if there are things you always want to remove from the final document (e.g. graphics, drop caps).
- Check Save and Launch if you want OmniPage Pro 12 to automatically run the destination program (determined by the Files of Type listbox). Typically you would keep this unchecked until the last page has been scanned.
- The Page Range and File Options listboxes become active only after a second page has been scanned. In order to save many pages as a single file choose “all pages” and “create one file for all pages” respectively (see Choosing from the Select Destination listbox).
- Note how thumbnails in the Image Panel have a disk icon beside them when they have been saved.

The Options dialogue box

- This is a multi-page dialogue box found in the Tools menu. Only the choices on each page that relate to the efficient production of e-text and hard copy Braille will be described.

OCR page

- Check the language of the source document. If the Braille translation program and/or the Braille notetaker supports foreign languages, the correct Braille contractions will be produced directly from the source document.

Scanner page

- Brightness and contrast can be adjusted here, depending on the quality of the source document and the resulting accuracy of the scan. If the characters in the source document are thick and touching, lighten the brightness and/or reduce the contrast. If the characters are thin and broken darken the brightness and/or increase the contrast. For large scanning jobs it is worth investing the time to get optimal accuracy from these settings. The payoff will be the reduced amount of cleanup necessary.
- If you are using an Automatic Document Feeder check this box to enable double sided scanning.
- If you are using a flatbed scanner you can have scans done automatically and set the time between scans here. This saves clicking the Scan 1 2 3 button to initiate each scan.

Direct OCR page

- Check “Enable Direct OCR” to be able to scan directly from, and into, your word processor (See Scanning directly into Microsoft Word). This may be preferable and more efficient than scanning in OmniPage Pro 12 and exporting the file to your word processor. OmniPage Pro 12 looks for common word processors and registers them within its program allowing you to enable Direct OCR.
- Check “Draw Zones Automatically” to have text recognized in your word processor. If it is unchecked then text will remain as a graphic (uneditable text).
- Check “Proofread OCR” to have the OmniPage Pro 12 proofreading window opened in your word processor. This is usually unnecessary as the word processor will have its own spellchecker.

Process page

- Settings here depend on the individual job that is being done. Each checkbox is a step or process OmniPage Pro 12 will have to take before it initiates a scan. Check only the boxes that are needed for the job as unnecessarily checked boxes will slow down the scanning process.
- Check “Automatically prompt for more pages” to delay the opening of the Save As dialogue box while you place the next page on the scanner and click “Add more pages”.

- If the source document is a single page and will be carefully oriented with the top of the page at the top of the scanner, check “Automatically correct page orientation” in the “Process Settings” dialogue box (Tools / Options / Process). Doing this will speed up the scanning process. However, if the page is to be put in the scanner sideways or upside down make sure “Automatically correct page orientation” is checked.

Proofing page

- Check “Automatically Proofread Results after OCR” if you want to do the clean up in OmniPage Pro 12. It is not necessary to do the clean up here as your word processor will have its own spellchecker.

Custom Layout page

- Using radio buttons on this page you can customize the text and objects that OmniPage Pro 12 will look for when it scans a page or pages. Graphics can be eliminated from the scan here. This will speed up the scanning process and reduce cleanup.

Typical Settings

- These settings assume that you are producing e-text for a Braille using student (clear text, right side up orientation, no zone modifications, and recognition to be done at the time of scan). They are:
 1. Retain Graphics – unchecked
 2. Image Source – Grayscale
 3. Layout – Multiple Columns, no table
 4. Save dialogue box – Retain Font and Paragraphs
 5. Direct OCR dialogue box; Enable Direct OCR – checked
 6. Direct OCR dialogue box; Draw zones automatically – checked
 7. Direct OCR dialogue box; Proofread OCR - unchecked
 8. Automatically Proofread Results After OCR – unchecked
 9. Process dialogue box; Automatically correct page orientation – unchecked
 10. Process dialogue box; Automatically Prompt for more pages – checked

About the Clean Up

- Cleanup involves the correction of spelling and formatting errors that OmniPage Pro 12 has made.
- Clean up may be done in either OmniPage Pro 12 or your word processor.
- If clean up is done in OmniPage Pro 12 check the “automatically proofread results after OCR” checkbox located in the Tools / Options / Process dialogue box. After the OCR process you will be prompted to correct the suspect words on each page.
- It is usually quicker correct the spelling and formatting of the entire document in the word processor (e.g. Microsoft Word), than to do it one page at a time in OmniPage Pro 12. This is because you can use the Find and Replace feature in the word processor to make global corrections to the document. Word will mark misspelled words with a red underline if “Check Spelling as You Type” is activated (Tools / Options / Spelling and Grammar). This makes them easy to spot. Uncheck

“proofread OCR” in OmniPage Pro 12 (Tools / Options / Direct OCR) if you are going to clean up in Word. It is redundant to have both the OmniPage Pro 12 and Word spellcheckers running at the same time and doing so may lead to system crashes.

File Management

- Make subject folders in Windows Explorer on your computer and have the student do the same. Avoid putting all finished files in “My Documents”. You can make subject folders within My Documents however, this is just another folder that must be opened to get to the subject folder. It is more efficient if subject folders (on both your computer and the student’s) are made directly off of the C:\ drive.
- Teaching students some DOS basics can speed up their ability to open and save files (i.e. to open a file from disk type A:\ in the “Look In” listbox of the Open dialogue box; to save a file to a subject folder type C:\folder name\file name and it will be saved directly in that folder.
- For students with their own laptops save the cleaned up document to disk (“Save As”) after saving it to your hard drive so that you always have a master document.

PART C: Scanning Directly into Microsoft Word ('98, 2000, XP)

General Information

- The OCR process can be done directly from within Microsoft Word, making it fully accessible for blind students to do their own scanning (however programs by Kurzweil and Arkenstone are designed specifically for independent access to print and have more features).
- The process may be slower using Word because you must wait for the entire process before proceeding to the next page. This will depend on the speed of your computer.
- “Enable Direct OCR” must be selected in OmniPage Pro 12 before scanning is possible from within Microsoft Word (or any registered word processor).
- Two items will be added to the File menu of the word processor. They are “Acquire Text” and “Acquire Text Settings”.
- Clicking on “Acquire Text Settings” will display an abbreviated OmniPage Pro 12 Options dialogue box. Not all choices are present. One choice that is not present is “Retain graphics”. There is no option to remove graphics so they will always show up when scanning in Word (even if graphics have been turned off in the OmniPage Pro 12 program).
- Clicking on “Acquire Text” will cause the scanning process to begin. If “Draw zones automatically” is checked (Direct OCR / Process Options) the incoming page will be zoned, recognized and displayed in the open Word document
- Source text may be in columns (make sure Original Layout is set to “multiple columns” in OmniPage Pro 12). The columns will be replaced with a properly ordered, continuous flow of text.
- Source material may be scanned upside down or sideways (make sure “Automatically correct page orientation” is checked in the OmniPage Pro 12 Options dialogue box (Tools / Options / Process)).
- For low vision students the font size can be adjusted (however, graphics will not be enlarged making them disproportionate to the text). It may be preferable to use a

screen enlarging program which will maintain proper proportions between text and graphics.

The Process

- Open a “New” Word document.
- To have Word underline words that are not scanned correctly, choose “check spelling as you type” from the Tools/Options/Spelling and Grammar dialogue box in Word. This makes it easier to identify misspelled words during clean up.
- In the “File” menu choose “acquire text settings”. Make settings changes if necessary (see **General Information**) and close the dialogue box.
- In the “File” menu choose “acquire text”. The scan, zone and recognize process will take place in Word.
- Subsequent scans using “acquire text” from the “File” menu will place text after the cursor location in the Word document.
- After all source material has been scanned into the Word document it may be easier to clean up if it has uniform size, font and justification. This will not be appropriate if retaining original text attributes is important for the student. To do this, however, choose “select all” in the “Edit” menu. This will highlight all of the text in the document. Press Control+Spacebar remove all attributes (bold, italics, etc.). Press Control+L to left justify all of the text and then choose a size from the toolbar that is easy to work with. Press the Up Arrow to remove the highlight and place the cursor at the top of the document.
- Save document regularly.

A “Related Materials” section is included in this resource book. Please refer to the Related Materials section for a list of narrated screen demonstrations that have been developed for each content section. All narrated screen demonstration files are included on your Accessing Literacy cd, also included with this resource book.