



Using Dragon NaturallySpeaking 7 in an Enterprise

Introduction

Dragon NaturallySpeaking® 7 comes with many tools designed to make integrating speech recognition into an enterprise productivity workflow easy and beneficial. By employing Dragon NaturallySpeaking in this workflow, organizations can save money on manual transcription costs, improve productivity, and optimize the contributions of employees with disabilities.

This white paper will describe the ways in which Dragon NaturallySpeaking can be installed and used in an enterprise environment. We will begin by describing the methods for using Dragon NaturallySpeaking as a “front-end” speech recognition system. “Front-End” means that the Dragon NaturallySpeaking application resides on the client machines, allowing users to instantly review and correct the results of the speech recognition. We will then describe how to integrate Dragon NaturallySpeaking into an existing workflow client application. Finally, this paper will conclude by describing how Dragon NaturallySpeaking can be used to create a “Back-End” transcription. “Back-End” means Dragon NaturallySpeaking resides on a remote machine in a network, and is designed to facilitate third party correction of transcripts generated by speech recognition. We will include a description of how to integrate Dragon NaturallySpeaking into a digital dictation workflow.

ScanSoft’s professional application engineering and training staff are available to assist developers, system integrators and IT professionals in implementing the workflows described in this white paper. Please contact ScanSoft sales for more information.

Front-End Speech Recognition

Installing Dragon NaturallySpeaking

Dragon NaturallySpeaking can be installed in one of two ways:

1. Directly on the client machine. This is done simply by inserting the Dragon NaturallySpeaking installation CD into the CD-ROM drive of the desired machine.
2. From a central network location using an MSI file. The MSI file allows Dragon NaturallySpeaking to be “pushed” to multiple client workstations on a network. This MSI file is available to all volume license customers of Dragon NaturallySpeaking software.

Creating A User Profile

To use dictation software such as Dragon NaturallySpeaking in a front-end implementation, each speaker must first actively “train” the system. This process prompts the speaker to read an enrollment text for about five minutes while speaking into a microphone, and results in Dragon NaturallySpeaking creating a unique “user profile.” User profiles are usually between 40 and 80 MB in size. After the five minutes of reading, Dragon NaturallySpeaking analyzes the style of existing documents (when such documents are available) representative of a speaker’s style and makes corresponding adjustments to the user profile.

Optimizing User Profiles

A key factor in the quality of the dictation experience is the quality of the user profile. Users can generally begin dictating immediately after the five-minute active enrollment. In many cases, however, the quality of the user profile can be further enhanced using tools available with Dragon NaturallySpeaking. All of these user profile-enhancing tools can be used at any time.

One option is to have users perform profile optimization individually:

- Users can perform additional active training by reading into the microphone for longer than the original five minutes.
- Users can add new words to their individual vocabularies.
- Users can create individual voice commands to match their workflow patterns.
- Users can teach Dragon NaturallySpeaking by correcting the recognition errors it makes during regular use. These corrections can be made by voice, or with the keyboard.

A second option is to optimize profiles centrally:

- An administrator can create custom commands, word lists and vocabularies as needed and “push” these enhancements out to all user profiles on the network. To do this, the administrator can create a script using nsadmin.exe. This command-line utility (found in the naturallyspeaking\program directory along with instructions) copies the files containing the custom data from a shared location into existing user profiles. Nsadmin can add functionality to a profile whether or not that profile is currently open.

ScanSoft and its Certified Dragon NaturallySpeaking Resellers offer full training and support services to assist in creating high quality user profiles. Additional details about how to optimize a user profile are also available in the *Dragon NaturallySpeaking User's Guide*.

Accessing User Profiles for Front End Dictation

The user profile resides on the user's client machine. When the user is on their dedicated client, the profile can be accessed simply by launching the Dragon NaturallySpeaking application. Frequently in an enterprise environment, however, users want to dictate into different workstations on the network. To accomplish this, user profiles must be accessible throughout the network.

There are several ways to achieve this.

- Install multiple user profiles on multiple machines in a network. Depending on the size and resources of a network, it may be possible to simply store everyone's user profiles on all the machines they regularly access. Using nsadmin.exe, vocabularies and commands can be updated centrally for all profiles on all machines.

It should be noted that this method does not allow for incremental user profile improvements that result from regular correction of misrecognitions to be shared among profiles for the same user that reside on different machines. For this reason, enterprises employing this method should be sure to properly optimize their profile before deploying to multiple machines. Of course, a user's master profile can be optimized on a regularly scheduled basis by performing additional training and analyzing new documents. This optimized profile can then be installed on all the network machines.

- Import user profiles from a central network location. Dragon NaturallySpeaking user profiles cannot be accessed directly from a networked location, nor can they be saved directly to a network location. Therefore, before dictating, each user must import his or her profile, using the Manage Users dialog of Dragon NaturallySpeaking. Opening Dragon NaturallySpeaking user profiles can be made part of a standard network log on procedure, so when a user signs into a network from a particular terminal, their user profile is automatically loaded. (More information on integrating Dragon NaturallySpeaking with workflow applications can be found in subsequent sections of this document.) The time required to import the user profile varies based on network conditions.

This method does allow for incremental user profile improvements that result from regular correction of misrecognitions. To take advantage of these improvements, a user must save their profile on the client machine after each use, then export the profile back to the central network location for later access from a different workstation. Similar to importing, exporting a user profile can be made an automatic part of the network log off process. Additional information on importing and exporting user files is available in the *Dragon NaturallySpeaking User's Guide*. It should be noted that if a user profile is sufficiently optimized, it may not be necessary to save and export the user file since the improvements are so minor as to be immaterial.

- In Dragon NaturallySpeaking 7, user profiles can be loaded directly from and saved directly to a read/writable device (such as a USB drive, CD/RW...). This method functions exactly the same way as having the profile installed on every machine in the network, with the advantage being incremental improvements from regular corrections of misrecognitions can be utilized, since the portable device acts as a master user profile and travels with the user from machine to machine. After each session, the user updates their profile, removes the device, and carries the device to the next workstation. More information on using portable storage devices is available in the *Dragon NaturallySpeaking User's Guide*.

Using a Third Party Correctionist to Correct Transcriptions Generated by Front End Speech Recognition

In the Professional, Legal and Medical Solutions editions of Dragon NaturallySpeaking, dictation in DragonPad, Microsoft® Word and Corel® WordPerfect® generates .dra files. .dra files contain the acoustics, results and transcription of the dictated text. .dra files can be sent via e-mail or over a network to a correctionist, who loads the .dra file in his or her own Dragon NaturallySpeaking installation and edits the transcription. The corrected transcription can be sent back to the dictator as corrected text, or routed directly to an enterprise content management system.

The .dra file allows the correctionist to play back the dictation with synchronized highlighting of text. If a document was created in Word (or WordPerfect), the correctionist must save the .dra file in the same directory as the Word (or WordPerfect) document. Then, while Dragon NaturallySpeaking is running on the correctionist's machine, open the Word (or WordPerfect) document, and initiate playback using the playback buttons on the Extras bar, via a voice command, or via the Sound menu. Playback can also be invoked via a hot key. There are hot keys for fast and normal playback speed. Foot pedals are available that can be programmed to generate keystrokes, so it is possible to use one of these foot pedals to control the playback functions and improve the efficiency of the correctionist.

If the correctionist doesn't have Word or WordPerfect, or the Word or WordPerfect document wasn't sent with the .dra file, the correctionist can simply open and edit the .dra file directly in DragonPad, the word processor that comes with Dragon NaturallySpeaking.

More information on third-party correction can be found in the *Dragon NaturallySpeaking User's Guide*.

Using Dragon NaturallySpeaking in a Thin-Client Environment

Users on a Citrix® or Windows® Terminal Server system can make use of Dragon NaturallySpeaking.

Certain features, such as Select-And-Say™ capability and Natural Language commands, can not be used with applications running on the server. These capabilities assume the application and Dragon NaturallySpeaking are running on the same client workstation.

Dictation without Select-and-Say does function properly, however. Commands that send keystrokes to the application also work. These commands must be application-specific to the Citrix or Windows Terminal Server client application, since that is the only application that Dragon NaturallySpeaking can detect.

To take advantage of Select-and-Say editing with an application that runs on the server, copy and paste text from the DragonPad editor or build custom commands to display dialog boxes for dictation and editing. More information on using the DragonPad editor is available from the ScanSoft applications engineering group.

Dragon NaturallySpeaking can be run on a server for transcription of .wav files.

The Dragon NaturallySpeaking user interface can be operated through the Terminal Server or Citrix client window for the purpose of doing transcription. If the audio output drivers are installed, the user can even play back the dictation.

Dragon NaturallySpeaking has no concurrent multi-user, multithreading or multitasking capabilities to support execution in a multiuser server environment. In other words, every user must run his or her own copy of Dragon NaturallySpeaking in its own address space. The server would have to have enough memory and processor speed to support multiple copies of Dragon NaturallySpeaking.

It is possible to use the Dragon NaturallySpeaking SDK to build a server-based transcription application that would take .wav files from multiple users and transcribe them sequentially. This is described in more detail in a subsequent section of this white paper.

Application Integration with Dragon NaturallySpeaking

When used in a front-end speech recognition context, it is often beneficial to have Dragon NaturallySpeaking work in conjunction with an application designed to facilitate a specific workflow. This may include Enterprise Content Management applications or Hospital Information Systems (HIS), for example. In this context, users will generally use speech in these applications to accomplish one or more of three basic tasks:

- Dictation (conversion of speech into text) – generally the most important function in a medical application because of its potential for boosting productivity
- Data entry (entering specific values into fields) – can also boost productivity if implemented well
- Navigation (moving from field to field, “clicking” menus and buttons) – typically used only if the user has his/her hands busy or is physically impaired

Dragon NaturallySpeaking, unmodified, provides these capabilities to some degree in any Windows-based application. Dragon NaturallySpeaking is optimized for use with Microsoft Office applications, Corel WordPerfect, and other popular applications. Dragon NaturallySpeaking also provides an application-programming interface (API) that gives software developers and system integrators a high degree of control over the speech capabilities of an application.

The following section of this white paper is aimed at the application developer. It provides some details of how an application can be written to take advantage of the built-in support for the above three tasks, and gives examples of how an application can use the Dragon NaturallySpeaking API to integrate speech capabilities when the built-in support is not sufficient. Complete documentation on the Dragon NaturallySpeaking API can be downloaded from <http://www.scansoft.com/naturallyspeaking/developers/>.

Dictation

Dragon NaturallySpeaking can do linear text entry into any Windows application; in other words, whatever the user says is typed as keystrokes into the application window.

However, dictation users typically want to be able to use speech to correct and revise the text they have just dictated, as well as have the recognizer learn from their corrections (“adapt”) and become more accurate. In Dragon NaturallySpeaking, these capabilities are collectively called “select-and-say”.

Most users of Dragon NaturallySpeaking are dissatisfied if Select-and-Say capability does not work in the application into which they are dictating. Depending on the application, the amount of programming effort to make Select-and-Say work can vary.

Applications explicitly supported

Applications in which Dragon NaturallySpeaking supports Select-and-Say without any customization include Microsoft Word and Excel 97, 2000, and XP, WordPerfect 10, Microsoft Excel 97, 2000, and XP, Microsoft Internet Explorer 5, 5.5, and 6, and Microsoft® Outlook® 2000 and XP.

Special consideration for Microsoft Word

Using Microsoft Word as its text editor can be a very effective way for an application to take advantage of the capabilities of Dragon NaturallySpeaking. Not only is Select-and-Say fully supported, but there is a powerful set of natural language commands that allow users to do formatting and editing tasks.

The application developer needs to be aware that if Microsoft Word is launched from or embedded in another application, Dragon NaturallySpeaking does not activate its Microsoft Word support unless the parent application is listed in a particular location of the Windows registry. In other words, it treats Microsoft Word as a nonstandard window unless this registry setting is made. Instructions for making this registry setting can be found at <http://knowledgebase.scansoft.com/view.asp?tnID=3512>.

Windows edit controls with built-in support

Dragon NaturallySpeaking supports Select-and-Say in these Windows edit controls:

Edit
RichEdit20W
RichEdit20A
RichEdit

Enabling select-and-say in edit controls that are derived from Windows edit controls

The application developer can sometimes enable select-and-say in edit controls that are derived from the controls on the above list. To do this, attach a dictation-edit control (a component of the Dragon NaturallySpeaking API) to each edit control in the application. This technique is very reliable but requires modification to the application (approximately one line of code per edit control in which dictation needs to be supported). See the Dragon NaturallySpeaking API documentation for more information.

Implementing Select-and-Say in nonstandard windows

Select-and-Say works in other applications and controls (called “nonstandard windows”) that do not fall into the above categories, provided the user employs only certain commands and avoids using the mouse or keyboard. See the Dragon NaturallySpeaking Help topic “Dictating in nonstandard windows” for details. It is not desirable to rely on this type of Select-and-Say support, since it does not work as reliably as users expect it to.

Reliable Select-and-Say support can be implemented by the application developer using the custom-dictation control in the Dragon NaturallySpeaking API. An application that uses the custom-dictation control is responsible for keeping the contents of the internal dictation buffer in Dragon NaturallySpeaking synchronized with the visible text in the application. (See the Dragon NaturallySpeaking API documentation for more information.) In particular:

- The application must handle notification from Dragon NaturallySpeaking every time the user dictates some text or uses a “select <words>” command to change the selection; in these cases the application must update the text that is visible to the user.
- The application must notify Dragon NaturallySpeaking every time the user types text into the application or changes the selection using the keyboard or mouse.

It may be simpler to convert the application to use Windows edit controls (Edit, RichEdit20W, RichEdit20A, or RichEdit) or Microsoft Word than to implement dictation support using the custom-dictation control; the developer should explore both options.

Data entry

Built-in support for data entry

Since Dragon NaturallySpeaking allows dictation into any field that accepts typing, it can be used for data entry virtually anywhere in an application.

Two of its modes can be used to improve accuracy:

- Numbers mode allows the user to enter only digits and some punctuation marks. It is useful for entering data that is purely numeric, such as vital signs.
- Spell mode allows the user to enter letters, digits, and some punctuation marks. It is be useful for entering non words such as part numbers.

The user says a command such as “switch to numbers mode” to place Dragon NaturallySpeaking in this mode, and “switch to normal mode” to turn it off. The user can also switch modes by using the Words menu of the Dragon NaturallySpeaking toolbar.

Integration using Dragon NaturallySpeaking API

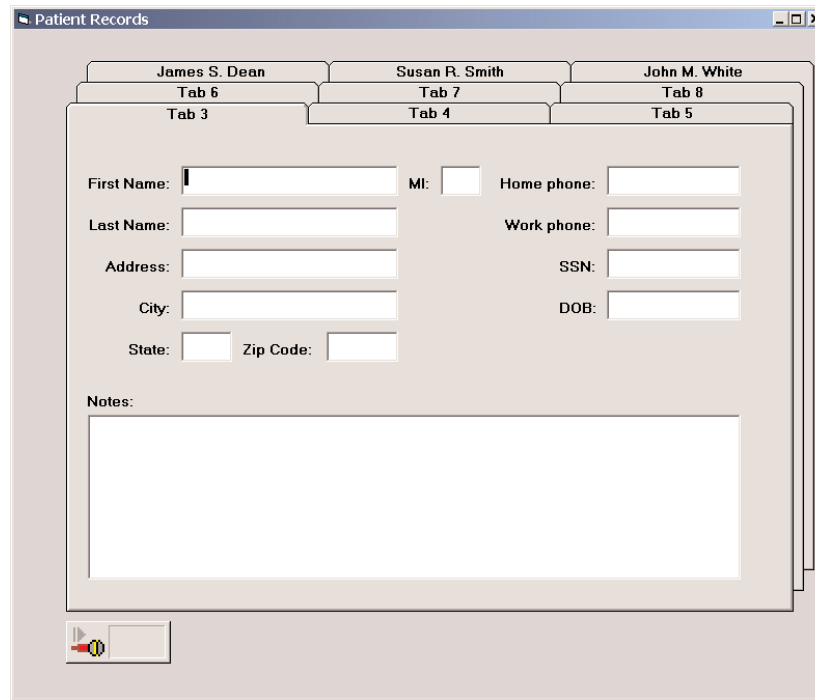
Data entry can be done with virtually perfect accuracy if the recognizer is able to accept only the values that are valid for a particular field. The software developer can use the voice-command and voice-menu objects of the Dragon NaturallySpeaking API to define precisely what words the user can say when the application is in a particular state.

The Dragon NaturallySpeaking API documentation includes a sample application called demo_de that illustrates this concept (see figure 1).

When the focus is on any field except First Name, Last Name, Address, City, or Notes, dictation is disabled. The user can say only navigation commands (commands that set the focus to a different field) or values that are valid for the field that has the focus; the recognizer rejects anything else the user says. Examples of valid values include:

- When the focus is on the State field, only state names and navigation commands are recognized.
- When the focus is on the Zip Code field, only a five-digit number or a navigation command is recognized.
- When the focus is on the Home Phone or Work Phone field, only a ten-digit number is recognized.

Figure 1: demo_de screen



Values for data entry can be determined at run time. For example, whenever the user places the focus on the Last Name field, the application can read the list of valid last names from a patient database and build a voice menu (a set of voice commands) that lets the user say any name on that list.

Navigation

As mentioned earlier, navigation by voice using built-in commands is typically slower than using the mouse or keyboard, and is used only if the user has his/her hands busy or is physically impaired. However, navigation by voice can be quicker than mouse or keyboard if custom commands are designed for navigation.

Also, in recent years, support for injured and disabled workers has become more and more commonplace in software applications, with speech input an important alternative for using the mouse and keyboard. Navigation commands are an important part of that support.

Built-in support for navigation

Dragon NaturallySpeaking provides basic navigation capabilities that can be used in most Windows applications.

- “Say what you see” – In an application that uses standard Windows menus, the user can say the name of any visible menu, submenu, or menu item to activate it. The user can also say the name of any button, radio button, or check box to activate it. The user can say the name of a static control to navigate to the corresponding text entry field. Furthermore, if the application uses Microsoft Active Accessibility (MSAA -- see the Microsoft web site for more information), Dragon NaturallySpeaking allows the user to say the name of any item exposed through MSAA. (In all cases, the name of the item must be pronounceable.)

- Keystroke navigation – If an application provides shortcut keys for its features, the user can use those features by using built-in voice commands that send keystrokes. For example, the user can say “press control n” to send the Ctrl+n keystroke combination, which in many applications causes a new document to be created.
- Mouse movement and clicking – If an application has features that can be activated only by mouse (or other pointing device), the user can use those features by saying mouse movement commands such as “mouse move right” and “mouse click.” Since moving the mouse by voice is slower than moving it by hand, this capability is useful only to physically impaired users.

If an application is designed so that “say what you see” is usable throughout the application, no additional programming effort is necessary for the user to navigate by voice.

Some capabilities of an application may not lend themselves to “say what you see.” For example, navigation through expanding lists of folders (similar to the folder pane in Windows Explorer) is not straightforward to do by voice; Dragon NaturallySpeaking does not automatically build voice commands for visible folders or other list items, let alone folders that are not visible but to which the user would like to go directly.

Special navigation capabilities for Internet Explorer

Dragon NaturallySpeaking has special capabilities for navigation in Microsoft Internet Explorer. For example:

- The user can say the name of any link instead of clicking it.
- If the focus is on a selection object, the user can say the name of any item on the list to select it.
- The user can say a command such as “click edit box” or “click list box” to place the focus on a particular item.

See the section “Working with Microsoft Internet Explorer” in the Working with Applications chapter of the *Dragon NaturallySpeaking User’s Guide* for a complete description of these capabilities.

These capabilities, along with the fact that Select-and-Say is fully supported in input areas, make Internet Explorer a very attractive mechanism for delivering a client user interface to a software package. No special programming effort is required to implement the speech capabilities. However, the Web developer must follow some common-sense guidelines, such as giving each link a unique and pronounceable name. Guidelines are in the white paper Guidelines for Speech-Accessible HTML for Dragon NaturallySpeaking, which is available in the downloads area of support.scansoft.com in the file [htmlguide3067.zip](#).

Integrating Navigation with the Dragon NaturallySpeaking API

If the built-in navigation capabilities are not sufficient, the software developer can use the voice-command and voice-menu controls of the Dragon NaturallySpeaking API to create additional commands for navigation.

With a well-designed set of navigation commands, the Dragon NaturallySpeaking user can be even more productive than someone using only mouse and keyboard. The efficiency comes when multiple operations are combined into a single, concise command. For example, a command could be built to allow a physician to say “New Progress Note for John Doe” (where John Doe is one of the existing patient names in the database), and as a result the software would look up John Doe’s record (otherwise the physician might have to scroll through a long list of patients to find John Doe) and open the screen that allows the physician to enter a new progress note (which might otherwise require several mouse clicks to get to).

The user can be still more efficient if the developer can combine navigation and data entry in a single command. For example, suppose a doctor needs to enter systolic and diastolic blood pressure into two separate fields of a screen. The developer can build a voice command called “BP <systolic> over <diastolic>” where “systolic” and “diastolic” represent lists of valid numeric values for blood pressure. The doctor would say “BP 110 over 70” and the application would respond by placing 110 in the field for systolic pressure and 70 in the field for diastolic pressure.

Application integration support and services are available from ScanSoft as well as from our international network of Certified Resellers.

Back-End Speech Recognition

An enterprise generally chooses to implement back-end speech recognition instead of front-end speech recognition because users are accustomed to a traditional manual transcription workflow and are reluctant to adopt a new process. Hence, back-end speech recognition attempts to create an environment whereby the person creating dictation is unaware that speech recognition is even part of the workflow.

Creating User Profiles Using Pre-existing Transcripts and Recordings

Back-end speech recognition still requires that a unique user profile be created for each user. The assumption is that users will not want to take the time to actively enroll. Profiles can still be created by using archived recordings and their associated transcripts, which were generated by manual transcriptionists. This can be accomplished by using a command line utility called `efenroll.exe`, which is located in the Programs directory of Dragon NaturallySpeaking. `Efenroll.exe` can be used to create a single user profile, or it can be configured to generate multiple user profiles based on batch processing. Details on how to use `efenroll` to create user profiles can be found in the appendix of this white paper.

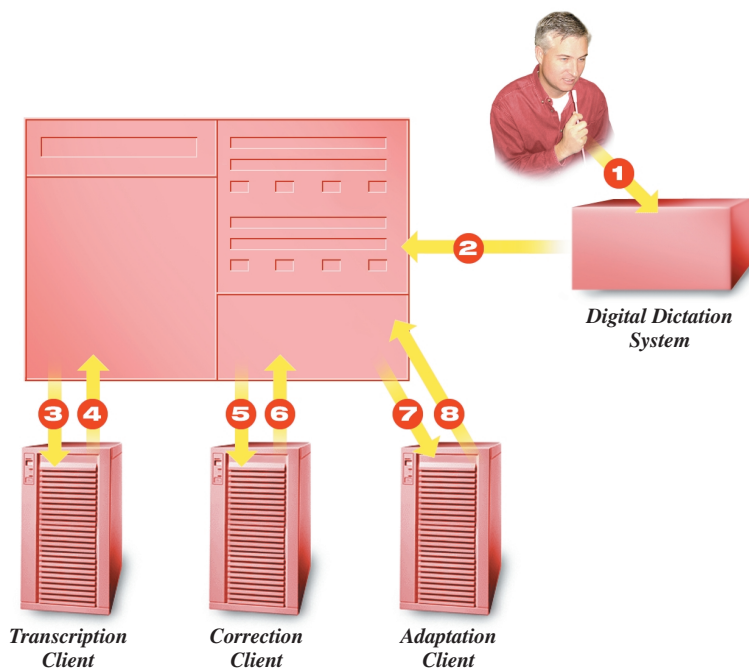
Integrating Dragon NaturallySpeaking Back-End Speech Recognition into an Existing Dictation Workflow

The diagram in figure 2 assumes a workflow application on the server that will dispatch various types of tasks to client machines that each run a specific recognition task. These tasks are:

1. Transcription
2. Correction
3. Adaptation

In the diagram, each Dragon NaturallySpeaking task is represented as running on a different machine, although in reality, this is not required. The workflow server application and all of its functionality would be designed by the system integrators or application engineers associated with the digital dictation system developers, as each system has unique configurations with specific design requirements.

Figure 2: Workflow Server Application



In step 1, a person dictates into a digital dictation system's input device, creating a recording such as a .wav file. In step 2, that .wav file is routed to the digital dictation workflow server application. In step 3, the digital dictation workflow server application sends three files to the transcription client: a user profile, wave data and a vocabulary. The user profile has been generated by the `efenroll.exe` utility as described above. The wave data (.wav file) has been generated by the digital dictation system. The vocabulary has

been created using NaturalVoc Tool and installed using the nsadmin.exe utility. When the digital dictation system receives a .wav file, it generally includes some metadata indicating the person who has made the recording and/or the topic of the recording. This metadata should be used to select the appropriate user profile and vocabulary.

An interface can be built using the Dragon NaturallySpeaking API that will initiate recognition of the wave data using Dragon NaturallySpeaking, which must be installed on the transcription client. This installation must be a Solutions Series edition of Dragon NaturallySpeaking (Professional, Medical or Legal). Specifically, the application needs to use the TranscribeFile method of the custom-dictation object to do the transcription, and the SessionSave method if the .dra file is to be preserved to allow playback during correction. For more information, see the Dragon NaturallySpeaking API documentation.

When recognition is complete, the transcription client sends back to the workflow server application: text, the original wave data and/or a .dra file, completing step 4.

The workflow server application then routes the text, wave data and/or .dra file to the correction client, illustrated by step 5. The correction client is staffed by a correctionist. The correctionist can then either listen to the wave data and read the text and make corrections using third party transcription and editing tools, or, if the correctionist has a Dragon NaturallySpeaking Solutions Series edition installed on their machine, use the .dra file to view a synchronized playback/highlighting of the wave data and text to facilitate corrections as described previously in this white paper. When the correctionist has completed editing and formatting, the correctionist sends the completed text back to the workflow server application, illustrated by step 6.

In step 7, the workflow server application routes the corrected text, the user profile, and the original wave data or .dra file to the adaptation client, which, using the efenroll.exe utility, updates the user profile to account for corrections made by the correctionist. Again, this is accomplished through the use of an interface that can be built using the Dragon NaturallySpeaking API. Once updated, the adaptation client sends the updated user profile back to the workflow server application, as illustrated in step 8. The adaptation client must be running a Solution Series edition of Dragon NaturallySpeaking. Once back at the workflow application, the updated user profile is ready for the next transcription assignment.

Appendix 1: Using efenroll.exe to create a user profile without active enrollment

To create a user profile without active enrollment, first create an untrained user. This can be done using the Dragon NaturallySpeaking API. Keep this user loaded as you proceed. Using the efenroll command line interface, then specify the archived wave file. The wave file needs to be in the “<drive letter>:\<installation directory>\<naturallyspeaking>” directory. To change the path and name of the wave file, the following variable needs to be written to options.ini (found in <natspeak>\users\<username>\current\options.ini) in the created user directory:

```
EFEnroll source file=c:\path\myfile.wav
```

The path of the wave file has to be fully defined. Modify options.ini only if the location and name of the wave data is different from the default. Using the above variable allows each user to have their own database of wave data and texts.

NOTE: To verify how this variable is set in options.ini, you can create a wave user and select a text for reading. At that point, you can inspect options.ini and you will see this variable in options.ini.

For custom training text change the file “enroll5.bin”, and list any additional texts. The texts themselves need to be copied into <drive letter>:\<natspeak>\training\<LNG>

Refer to the text with proper index into the file enroll5.bin. The full command line is:

```
efenroll fromfile textselected=n
```

The variable textselected is 0-based.

By default, efenroll requires 10 minutes of data or a 15 MB wave file. If this needs to be a lower value, the following change needs to be made in nssystem.ini (“<drive letter>:\<installation directory>\<naturallyspeaking>\program”)

[Settings]

Training Data Needed For Eyes-Free Adaptation=nYourChoice (define as milliseconds; default value is 600,000)

Change the following parameter in nssystem.ini to the following value.

[Settings]

Acceptance percentage for EFenroll=10

Once a user is trained, the user can continue to be trained with additional acoustic data and accompanying transcription using the same command line as above. The training will be incremental:

```
efenroll fromfile textselected=n
```

If additional training is not the preferred way of training, then all data and text needs to be carefully concatenated into one large wave file with an accompanying transcription that is consistent in order with the acoustic data.

The processes described here can be automated into a batch process. The following is a rough specification on how this can be accomplished:

Create a small utility that would read all the input data from a text file (user name, path to the wave file and associated transcription).

In order to create the user the program would have to:

1. Read input file to get data
2. Call the SpeakerCreate method of the engine-control interface in the Dragon NaturallySpeaking API. This will create a user based on a predefined speaker model. Call BaseSpeakerModelsGet to obtain the default.
3. Call TopicCreate, to create a predefined topic model. Call BaseTopicModelsGet, to select the default.
4. Since the user is created with the default dictation source, is necessary to modify options.ini to change it to a mobile source, so the following entries need to be added: “Dictation Source=<8>”, “Live Input User=0”, “EFenroll Source File=<wave>”.
5. Call SpeakerTopicSelect in order to load the user and topic just created.
6. Run efenroll using the command line.
7. Detect when efenroll has completed and return to step 1 to read the next data input from the file.

The ScanSoft logo, Productivity Without Boundaries, Dragon NaturallySpeaking, and Select-and-Say are trademarks or registered trademarks of ScanSoft, Inc. in the United States and/or other countries. All other company names or product names may be the trademarks of their respective owners.

Copyright © 2003 ScanSoft, Inc. All Rights Reserved. ScanSoft, Inc. does not provide any assurances on the accuracy of the product. The product can also be subject to changes without notice. ScanSoft, Inc. shall not be responsible for, or liable to, any customer for decisions made based on this document. The information contained in this white paper does not constitute legal advice.